

Working Paper 1
3 October 2007

ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION
and ECONOMIC COMMISSION FOR EUROPE**

**UNITED NATIONS DIVISION FOR THE
ADVANCEMENT OF WOMEN**

CONFERENCE OF EUROPEAN STATISTICIANS

UNITED NATIONS STATISTICS DIVISION

Expert Group Meeting on indicators to measure violence against women

(Geneva 8-10 October 2007)

Co-organized by UNDAW, UNECE and UNSD

in collaboration with ECA/ECLAC/ESCAP/ESCWA

INDICATORS TO MEASURE VIOLENCE AGAINST WOMEN

Invited Paper

Prepared by Sylvia Walby¹

The views expressed in this paper are those of the author and do not necessarily reflect those of the co-organizers.

¹ Sylvia Walby, Lancaster University, UK (email: S.Walby@Lancaster.ac.uk).

CONTENTS

- I. INTRODUCTION
- II. INDICATORS
 - A. What are indicators?
 - B. Purpose of indicators
 - C. General criteria for indicators
 - D. Specific criteria for indicators of violence against women
- II. OVERVIEW OF INITIATIVES ON INDICATORS
 - A. Policy proposals for indicators
 - B. Table: Summary of initiatives on indicators
 - C. Indicators embedded in survey statistics
 - D. Reflections on the relationship between policy proposals for indicators and indicators embedded in surveys
- III. OPTIONS IN DEVELOPING INDICATORS ON VIOLENCE AGAINST WOMEN
 - A. Introduction
 - B. Definition of the types of violence
 - C. Units for the measurement of extent: prevalence and incidents
 - D. Severity and impact
 - E. Time period
 - F. Population sub-groups: age and marital status
- IV. SUMMARY OF MAJOR OPTIONS AND RECOMMENDATIONS FOR INDICATORS
- V. BIBLIOGRAPHY
- VI. APPENDIX: DATA COLLECTION REQUIREMENTS AND CONSTRAINTS AND OPPORTUNITIES FOR OVERCOMING THESE

I. INTRODUCTION

1. This paper is prepared for the UN Expert Group Meeting on indicators to measure violence against women. The Expert Group Meeting is intended to support the work of the Statistical Commission and the Commission on the Status of Women in developing ‘a set of possible indicators on violence against women in order to assist States in assessing the scope, prevalence and incidence of violence against women’.

2. The paper includes: (i) an overview of existing major initiatives on indicators to measure violence against women; (ii) an assessment of the advantages and disadvantages of those initiatives; (iii) proposes criteria for the identification of a possible set of indicators on violence against women; (iv) summarises options and provides recommendations for a possible set of indicators to support countries to measure the scope, prevalence and incidence of violence against women; and (v) addresses the related data collection requirements and constraints and opportunities for overcoming these.

II. INDICATORS

A. What are indicators?

3. Indicators summarise complex data into a form that is meaningful for policy makers. The focus here is on indicators and statistics that measure the ‘scope, prevalence and incidence’ of violence against women. It does not address indicators of policy developments, which are considered elsewhere (UN DAW 2005b; Kelly 2007).

B. Purpose of indicators

4. Indicators constitute a key link between an evidence base and policy making. There have been many policy innovations to reduce and eliminate violence against women; much political good-will; and much rhetoric. In order to decide whether initiatives are having a positive impact it is necessary to know whether the situation is deteriorating or improving. There are many forms and types of knowledge about the nature of violence against women and the policies to stop this. Often these data are too complicated to support the decision-making of policy makers without the input of considerable time and expertise. The purpose of indicators is that they provide a simple summary of a complex picture, abstracting and presenting in a clear manner the most important features needed to support decision-making (Berger-Schmitt and Jankowitsch 1999; Luxembourg Presidency 2005; UN Economic Commission for Latin America and the Caribbean 2004; Statistics Canada 2002; Walby 2005a).

C. General criteria for indicators

5. Several criteria for the selection of the indicators have been developed (Berger-Schmitt and Jankowitsch 1999; UN Economic Commission for Latin America and the Caribbean 2004; Statistics Canada 2002). In general, indicators should:

- Summarise complex data;

- Be unambiguous and easy to interpret;
- Enable an assessment as to whether an improvement or deterioration has occurred;
- Be meaningful and relevant to policy makers, service providers and the wider public;
- Be capable of being supported by reliable and robust quantitative data;
- Be neither so many as to confuse, nor so few as to mislead;
- Be available at regular intervals and be comparable over time;
- Be comparable between countries and population groups.

D. Specific criteria for indicators of violence against women

6. In addition to the general criteria for any indicator, there are criteria specific to the field of violence against women. While there is consensus on the criteria, there has been varying resolution of the tensions contained within them. The best way to resolve these tensions is discussed below using academic, statistical and policy literature. The criteria include:

- Inclusive scope of the full range of types of violence, as noted in the UN definition, but not so specialised as to prevent comparison between countries, thereby balancing local specificity with international comparability;
- Meaningful measurement of the extent of the violence, using the appropriate balance between the concepts of prevalence and incidence;
- Meaningful measurement of severity of the violence, especially in relation to its impact.

7. There are two technical considerations for the indicators:

- Consistent identification of the time period: both a longer period (for example, life-time or since 16 years old) and a more recent period, (for example, last year);
- Consistent identification of the same population sub-set, e.g. age and marital status.

8. There are two further criteria relating to policy and data collection:

- Consistent with indicators in adjacent fields, so as to facilitate the mainstreaming of violence against women into mainstream data collection and policy development, while still being sensitive to the nuances in the specific field of violence against women;
- Practicality of data collection; availability of data and existing use of indicators.

III. OVERVIEW OF INITIATIVES ON INDICATORS

9. Proposals for indicators of violence against women have been developed by a number of policy oriented governmental and non-governmental bodies. In addition, surveys of violence against women often include 'headline' figures when reports are published, although their main purpose is usually a more rounded detailing of the violence than a focus on indicators. Further, various bodies are in the process of developing indicators, including the UN Rapporteur on Violence Against Women (Kelly et al 2006), Eurostat (European Commission 2006a), and MEASURE Evaluation (2007).

A. Policy proposals for indicators

10. This section focuses on proposals for indicators from: United Nations High Commissioner for Refugees (2003); UN Economic Commission for Latin America and the Caribbean (Alméras, Bravo, Milosavjevic, Montaña and Rico 2004); European Union (Presidencia de la Unión Europea 2002); Statistics Canada (2002); the US Centers for Disease Control and National Center for Injury Prevention and Control (Saltzman et al 2002); and the UK Home Office (2005). In most cases these bodies have proposed indicators not only on the 'scope, prevalence and incidence' of violence against women but also on policy development. Only the former are reported here, because this report's focus. The section starts with a discussion as to why criminal and administrative statistics are an inadequate basis for indicators.

Criminal and Administrative Statistics

11. The extent of violent crime reported to the police or where criminal convictions are obtained are sometimes used as indicators of violent crime. However, for two main reasons, this approach is rarely recommended for indicators of violence against women. First, in most countries data on violent crime does not include the gender of the perpetrator and the victim. Hence it is not possible to use either criminal statistics or reports to the police as the main basis for indicators of gender-based violence against women. Second, criminal statistics and administrative data are in general unsuitable for the measurement of the scope, prevalence and incidence of violence against women because the majority of women who are victims of violence do not report this to any authority. This means that the majority of incidents of violence against women are not recorded in any administrative database. Hence it is necessary to conduct surveys in order to ascertain the scope, prevalence and incidence of violence against women.

12. There are two types of possible exceptions to this analysis. One concerns crime categories that are predominantly concerned with gender-based violence against women. One example of this is rape; but this is seldom reported to the police and is only a small proportion of overall violence against women. A further potential exception is in those few, but increasing, jurisdictions where domestic violence is named as a specific crime (e.g. Sweden), but since most incidents of domestic violence are not addressed using this crime category in these countries, it is not a recommended basis for indicators. A second type of exception concerns homicide. Homicide is different from most crimes in that in most instances it comes to the attention of the police and is recorded. Since the number of deaths of women from gender based violence cannot be ascertained by survey, police and crime statistics are the only possible sources. However, while the sex of the victim is often reported, it is very rare (the UK is an exception) that the analysis is done and reported as to whether the death of the women is gender based (e.g. domestic or rape-murder, but not due to robbery).

13. The remainder of the report considers indicators based on statistical data, since this is the most important form of data to support indicators of violence against women.

United Nations High Commissioner for Refugees

14. The United Nations High Commissioner for Refugees (2003) reported on guidelines for prevention and response to sexual and gender-based violence (SGBV) against refugees, returnees

and internally displaced persons. The guidelines contained a requirement for effective documentation of the extent and nature of such abuse. It presented a set of definitions of different forms of SGBV, including sexual violence, physical violence, emotional and psychological violence, harmful traditional practices, and socio-economic violence, together with a set of reporting tools. A key indicator is that of the 'sexual and gender-based report rate', based on the number of incidents of SGBV in a population during a designated time period (month, year etc), expressed as a number of incidents per 10,000 persons during that time period. This proposal was not applied to data.

UN Economic Commission for Latin America and the Caribbean

15. The UN Economic Commission for Latin America and the Caribbean produced a report on the importance of measuring violence against women, containing a review of existing practices in Latin America and recommendations (Alméras et al 2004). The report draws on a number of national surveys of violence against women, carried out in Latin America and the Caribbean. These use a variety of definitions. They focus on one sub-set of violence against women, that of intimate partner violence, proposing five indicators distinguishing various dimensions of the incidence of this sub-set of violence against women (of which those that fall within the report's remit are): the overall rate of violence; physical violence; psychological violence; sexual violence. They recommend disaggregating these measures of incidence by selected characteristics, including: geographical area; poverty status of households; age of women; activity status; pregnancy status; racial and ethnic group; and relationship with aggressor. The report presents proposals for indicators for future data collection, rather than providing indicators for which the relevant data already exists.

European Union

16. The Council of the European Union agreed in 1998 to develop a set of indicators and benchmarks in order to monitor the implementation of the 1995 UN Beijing Platform for Action, including violence against women. Since then several EU Presidencies (Spanish, Danish, Irish, Greek, Dutch), in association with the Commission, Council and a High Level Group on gender mainstreaming, have made proposals for the development of indicators on violence against women (Presidencia de la Unión Europea 2002; European Union 2004). There have been several further declarations from the EU about the need to develop indicators of violence against women, including the European Economic and Social Committee (2006) and the European Commission (2006b).

17. There are three proposed indicators on domestic violence, of which one is focused on 'scope, prevalence and incidence': the number of female victims. There are proposals for three indicators on sexual harassment in the workplace, of which one is focused on 'scope, prevalence and incidence': the percentage of employees who report incidents of sexual harassment. They do not include detailed specification (e.g. no definitions are provided). While some ad hoc data on violence against women has been collected by national level surveys, there is no regularly collected data at EU-level to support any of these indicators.

18. There are plans to develop indicators and data collection on violence against women as part of a strategy to measure crime and criminal justice in the EU. The Dublin Declaration (2003) recommends a 'comprehensive system of European crime statistics'; while the Hague

Programme (2005: 1) (OJ C53 of 3.3.2005, p.1) welcomes the Commission initiative ‘to establish European instruments for collecting, analysing and comparing information on crime and victimisation and their respective trends in Member States, using national statistics and other sources of information as agreed indicators’. The action plan (European Commission 2006a) includes the development of two sets of indicators: ‘Objective 12. Measuring violence against women’ and ‘Objective 13. Measuring domestic violence.’ There are discussions on the development of these indicators on violence against women and domestic violence for the EU involving Eurostat and HEUNI, but no proposals in the public domain.

Statistics Canada

19. Statistics Canada (2002) recommends indicators for six themes, of which two concern ‘scope, prevalence and incidence’:

- Severity and prevalence of violence against women
- Impact of violence against women

20. The report provides detailed data, drawn from two large scale national surveys. Much is focused on spousal violence, including both marital and common-law unions. The measures of severity include a modified Conflict Tactics Scale (CTS). Prevalence is based on any experience of physical or sexual violence since the age of 16 and in the last year. Impact measures include self-reported psychological impact and fear, physical injury, whether injuries were medically treated, whether there was time off work, and the economic cost of violence. There is a separate section on homicide, using data from sources other than surveys.

US Centers for Disease Control, National Center for Injury Prevention and Control

21. Saltzman et al (2002) provide careful and detailed uniform definitions of intimate partner violence in order to promote consistency in the use of terminology and data collection, in a report sponsored by the US Centers for Disease Control, National Center for Injury Prevention and Control. The report is concerned with intimate partner violence, not the full range of violence against women. They recommend data on both prevalence and incidents:

- number of people (and their characteristics) affected by intimate partner violence
- number and type of intimate partner violence episodes (and associated injuries and other consequences)

UK Home Office

22. The UK Home Office (2005) has performance indicators for domestic violence. In addition to several concerning policy, two indicators concern its scope, prevalence and incidence:

- Number of domestic homicides (data available annually from *Criminal Statistics*);
- Headline prevalence of domestic violence (supported with data from the annual British Crime Survey Inter-Personal Violence module).

23. This is one of the few instances in which policy proposals for indicators are also supported by regularly collected data. However, these are not strictly indicators of violence against women in that they include men as well as women victims.

Spain

24. Within a set of around 100 mostly policy oriented indicators from Spain, one set concern the scope, prevalence and incidence of violence against women (Federal-Provincial-Territorial Ministers Responsible for the Status of Women 2002). These are prevalence of domestic violence against women by partners or ex partners in the past 12 months and over the life-time, differentiated into three types: physical, sexual and psychological. In addition there are indicators concerning domestic violence and attempted murder reported to the police. There are indicators for the impact of this violence on social interaction, physical health, mental health, children and labour market participation.

B. Summary of Initiatives on Indicators

| Source | Indicators | Proposed or in existence |
|--|--|--|
| UN High Commissioner for Refugees | Number of incidents of sexual and gender-based violence (sexual, physical, emotional and psychological, harmful traditional practices, socio-economic) per 100,000 persons in a designated time period. | Proposed |
| UN Economic Commission for Latin America | Rate of intimate partner violence, distinguished by: Physical violence; psychological violence; sexual violence | Proposed |
| European Union | The number of female victims of domestic violence. Percentage of employees that report incidents of sexual harassment. Violence against women. Domestic violence. | Proposed |
| Statistics Canada | Severity and prevalence of violence against women (since 16; last year) (focus on spousal violence). Impact of violence against women (self-reported psychological impact, physical injury, whether injuries medically treated, whether time off work). | Applied to 2 national surveys |
| US Centers for Disease Control | Number of people affected by intimate partner violence (IPV) (and their characteristics). Number and type of IPV episodes (and associated injuries). | Proposed |
| UK Home Office | Number of domestic homicides. Prevalence of domestic violence. Annual. | Annual crime statistics & survey |
| Italy ISTAT | Physical violence. Sexual violence. Psychological violence. Lifetime and last year prevalence. | One survey; others planned every 4 years |
| Spain | Prevalence of domestic violence against women by partners or ex partners in the past 12 months and over the life-time, differentiated into: physical, sexual and psychological. | One survey |
| UNECE review of surveys | Numerous complex statistics and summaries including: prevalence, incidence and injury. | Surveys |

C. Indicators embedded in survey statistics

25. There are many national surveys that report a range of statistics on violence against women (Walby and Allen 2001; Martinez and Schröttle et al 2006; Schröttle et al 2006), as well as a series of comparative surveys including the International Violence Against Women Survey in Australia, Costa Rica, Czech Republic, Denmark, Mozambique and Switzerland) (HEUNI 2007; Nevala 2005), Demographic and Health Surveys in Cambodia, Colombia, Dominican Republic, Egypt, Haiti, India, Nicaragua, Peru and Zambia (Kishor and Johnson 2004), and the World Health Organization surveys in Bangladesh, Brazil, Japan, Namibia, Peru, Samoa, Tanzania and Thailand (Garcia-Moreno et al 2005). On publication, some surveys highlight a limited number of statistics. Although these 'survey highlights' are rarely called indicators, they are intended to be indicative in some way of the more complex and detailed data in the survey. For example the report on the Italian survey (ISTAT 2007) provides summaries and detailed data, though it uses the concept of 'measure' rather than indicator.

26. The best account of the indicators used in these surveys is found in a report of the United Nations Economic Commission for Europe Task Force on Measurement of Violence Against Women (2006), which is the most up-to-date and comprehensive review of surveys at a national and international level. This report provides information on the use of indicators and survey methods in 25 questionnaires used in 17 countries predominantly but not exclusively in North America and Europe: Australia, Canada, Denmark, Finland, France, Germany, Italy, Mexico, Norway, Poland, Moldova, Serbia and Montenegro, Spain, Sweden, Switzerland, the UK, and the USA. This report includes information on: the use of prevalence and/or incidence; the range of perpetrators; the distinctions made between different types of violence; the way severity of the violence was recorded; and the time period for recall that was used; the extent to which comparisons were made with police data; and the extent to which surveys were the main source of information on violence against women.

27. It finds that 84% of surveys provided data to support indicators on prevalence, 60% on incidence, 76% on injury and 44% on attitudes to violence. It finds that 92% of surveys could separately identify violence by current intimate partners; 84% by previous intimate partners; a smaller proportion could identify violence by other family members. Almost all the surveys were able to distinguish between different acts of violence, most of these using the full or modified version of the Conflict Tactics Scale. 92% of the surveys had detailed questions on sexual violence; 80% had data on psychological abuse; stalking was included in 80% of dedicated surveys, dropping to 40% in modules. Of these surveys, 80% had information on the physical injuries caused by the violence; most of these had additional information about the use or not of medical services; 84% asked victims about their search for help from the police. The time period used was most frequently life-time in addition to a shorter period, most frequently one year, but sometimes three or five years. Half (52%) of the surveys do not have an upper age limit, while 44% do. Most surveys are asked of adult respondents, though the age of adulthood varies. Only 3 of the 25 surveys include respondents from institutions as well as private households. Most surveys interviewed only women; though 36% also interviewed men. Of the surveys, 60% were dedicated violence against women surveys, while 32% were modules in general victimisation surveys and 8% were modules in multipurpose or health surveys. Most were ad hoc rather than regularly repeated. Seventeen countries reported to UNECE that they had carried out a national survey on violence against women. For 12 countries and 68% of the

surveys this was a one-off ad hoc event. None of the governmental sources cited the International Violence Against Women Survey as a source of statistics, even if it had taken place in that country.

D. Reflections on the relationship between policy proposals for indicators and indicators embedded in surveys

28. In general, with rare exceptions, there is a gap between proposals for indicators from the policy world and the development of detailed statistics in the context of surveys. Many of the proposals for indicators have not been applied to actual bodies of data; they remain aspirational, indicative of the direction of development along which policy makers would like to go. Much of the intent of the work in the developing surveys has been an interest in developing a breadth of knowledge about the nature of violence against women, rather than indicators for policy evaluation purposes. This is a gap that needs to be bridged.

III. OPTIONS IN DEVELOPING INDICATORS ON VIOLENCE AGAINST WOMEN

A. Introduction

29. In order to evaluate indicators on violence against women it is necessary:
- to specify the separate **dimensions** of the indicator that need to be defined;
 - to identify the **criteria** by which these options are to be assessed;
 - to identify the **range of options** on this dimension;
 - and to identify the **best option** after due consideration of the relevant evidence.
30. The **dimensions** of the indicators on the scope, prevalence and incidence of violence against women require the following dimensions to be defined:
- Definition of the types of violence.
 - Units for the measurement of prevalence and incidence.
 - The measurement of severity, including as a threshold.
 - The time period for the indicator.
 - Whether any restrictions are placed on the population of women included.
31. The **criteria** for assessment of the options of indicators include:
- Definition of violence: Inclusive scope of the full range of forms of violence, but not so specialised as to prevent comparison between countries, balancing local specificity with international comparability;
 - Measurement of the extent of the violence: using the appropriate balance between the concepts of prevalence and incidence;
 - Measurement of severity of the violence: to be meaningful especially in relation to impact.
 - Time period: consistency and accuracy in the identification of the time period: both a longer period, and a shorter recent period.
 - Population sub-set: Consistent identification of as full a population sub-set as possible, in relation to age and marital status, so as to ensure only technically essential exclusions.

- Consistency with indicators in adjacent fields, to facilitate the mainstreaming of violence against women into mainstream data collection and the full range of policy concerns, while still being sensitive to the nuances in the specific field of violence against women.

32. While some of the proposed indicators include a full specification of the full set of dimensions, many do not. Many policy bodies indicate the broad scope of a proposed indicator without its technical specification. Typically the surveys contain a wealth of information that allow for support of a range of different options along these dimensions; sometimes these are prioritised.

33. Each of the dimensions is now considered in turn. The discussion is composed of two parts: the identification of the range of options (from proposed indicators, surveys, and the academic and policy literature); the assessment of these options against specified criteria.

B. Definition of forms of violence

34. *The major options for indicators*

- A single indicator that includes all of forms of gender-based violence against women and that does not separately specify them.
- A separate indicator for each form of violence;
- A few indicators, one each for the more major forms of violence against women;
- A single indicator confined to domestic violence.

Discussion of the application of the criteria

35. The indicators require a definition of the scope of the violence. In the UN Secretary General's report (2006: 12), "violence against women" is understood to mean any act of gender-based violence that is directed against a woman because she is a woman or that affects women disproportionately'. The United Nations (1993) definition of gender-based violence in the 1993 Declaration on the Elimination of Violence Against Women is: 'Any act of gender-based violence that results in, or is likely to result in, physical, sexual or mental harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or private life'. The UN Secretary-General's report (2006: 37-47) noted the various forms of gender-based violence against women: intimate partner violence, harmful traditional practices, sexual violence by non-partners, sexual harassment in the workplace, and trafficking. Intimate partner violence (often known as domestic violence), includes physical violence which 'involves intentionally using physical force, strength or a weapon to harm or injure the woman, sexual violence, which 'includes abusive sexual contact, making a woman engage in a sexual act without her consent, and attempted or completed sex acts with a woman who is ill, disabled, under pressure or under the influence of alcohol or other drugs', psychological violence which 'includes controlling or isolating the woman, and humiliating or embarrassing her', and economic violence which 'includes denying a woman access to and control over basic resources'. Harmful traditional practices include female genital mutilation/cutting, female infanticide, early marriage, forced marriage, dowry-related violence, and crimes against women committed in the name of 'honour'.

36. In developing the indicator, there is a tension between the separate naming of specific forms of violence against women and the use of more general categories. The greater the number of indicators, the wider the range of forms that can be separately identified. The detailing of the specificity of the form that violence might take has both the advantage of being close to the experiences of the women affected as well as the disadvantage of potentially being used to stigmatise more marginal communities. The separate identification of types of violence that are less common has the disadvantage that the small numbers that would be reported to a survey might not be statistically reliable for reporting and analysis. There is a data collection issue here: the larger the size of the sample in the survey, the more possible it is to make finely grained distinctions between the types of violence.

37. In response to these dilemmas, some proposals for indicators have focused narrowly on the more common forms of violence, such as domestic violence, in order to have sufficient cases each year to enable the development of a robust and comparable 'rate'. However, a focus on domestic violence only narrows the type of violence unduly; omitting non-domestic sexual violence skews the potential gendered policy implications in a troublesome way. So, the category of domestic violence is not a sufficient operationalisation of the concept of violence against women.

Recommended options

38. For international comparisons:

- A main indicator that includes all forms of gender-based violence against women and does not separately specify them (although they are separately named in the data collection instrument).
- Additional separate indicators for the major types of violence that are found in all countries: intimate partner violence; rape; sexual harassment in the workplace.
- Not to use domestic violence only as the basis of an indicator for violence against women.

39. For national purposes in addition:

- Additional separate indicators for those forms of violence that are common in that country, but which are not frequently found in all countries.

C. Units for the measurement of extent: prevalence and incidents

Major options

- Prevalence: rate (%) of violence against women in the female population
- Incidents: number of incidents of violence against women per unit (e.g. 100, or 1,000) of female population
- Variations: these may differ according to the time period of life-time or short recent period e.g. last year.

Discussion

40. There are two main approaches to the operationalisation of the extent of violence against women: prevalence and incidents (Hélie, Clément and Larrivée 2003). The first concept,

prevalence, refers to the proportion of the population that has experienced violence in a given period, usually either (adult) life-time or the previous year. The second concept refers to the number of incidents of VAW in a given population unit (for example x incidents per 100 or 1000 people). They are each expressed as rates. They are the same if each victim has just one incident in the given time period; however, if a victim is subject to repeat victimisation, then the rate of incidents per 100 people will be higher than the prevalence rate per 100 people. There are strengths and weaknesses associated with each of these approaches. These strengths and weaknesses vary very significantly by the time period under consideration; hence life-time (or adult life-time) will be considered separately from one year.

41. Life-time: In measuring the extent of violence against women over a life-time (or adult life-time), the concept of prevalence has the advantage over that of incidents because it would be unlikely that there would be accurate recall of each and every incident over such a long period. This figure, which ranges from around one quarter to one half of women in their life-time, has been important in the establishment of the scale of the problem (Krug et al 2002) and in raising consciousness about the issue (Garcia-Moreno et al 2005).

42. Annual: In measuring the extent of violence against women over one year (or similar short recent period), it is reasonable to expect that the victim might accurately recall the number of incidents, so there is a choice between the use of the concepts of prevalence and incidence. The advantages of the concept of prevalence are that it would map onto the life-time rate in a straightforward manner; and that it draws on the practitioner community preferred conception of domestic violence as a 'course of conduct' of repeated events. The advantages of the concept of incidence is that it uses the same concept of rate as that used in adjacent policy fields, especially that of crime policy, facilitating the mainstreaming of violence against women into crime policy; and that it avoids the underestimation of the proportion of violent crime that is domestic and gendered that is consequent upon the use of the concept of prevalence, which tends to generate a spurious gender symmetry.

43. The difference between the concepts shows up most clearly when the focus is on repeat offences, such as intimate partner violence, rather than on the kind of violent incident that is unlikely to be repeated, such as female genital mutilation/cutting (FGM), or less likely to be repeated, such as rape. Hence when the focus is on those forms that are not or unlikely to be repeated the use of one concept or another to capture the rate of the violence produces little difference. The location of the greatest difference is on violent actions that are often repeated, including especially, but not only, domestic violence. Hence the next section focuses on domestic violence, where the heart of the issue lies.

44. The concepts of prevalence and incidence relate to two different ways of conceptualising the repeated acts of violence and threats that constitute domestic violence. The notion of prevalence captures the particular and special nature of domestic violence as a coercive 'course of conduct', a series of related occurrences, rather than a one-off event, which engenders fear in between the acts of physical violence themselves. The conceptualisation of domestic violence as a course of conduct is often used by those who provide services for its survivors. While prevalence is often used by the practitioner community around domestic violence because of its alignment with the recognition of the linked nature of the incidents in a course of conduct, within the world of surveys and the adjacent policy field of crime it has two very problematic consequences. First, as compared with the use of the concept of incidence the use of prevalence

diminishes the visibility of gender inequality and gender asymmetry in domestic violence; second, the concept of prevalence and course of conduct is based on a specialist understanding that is harder than incidence to mainstream into adjacent policy and research fields such as crime policy.

45. The apparent finding of gender symmetry in gender-based violence in sample surveys of domestic violence (Straus and Gelles 1990) has led to long string of criticisms of particular aspects of methodology in the attempt to investigate why these rather surprising results have been found, including the nature of the introductory questions to the respondents, and the way that concepts of control, conflict and violence are absent or positioned in the survey (Dobash et al 1992; Johnson 1996).

46. The disadvantage of using the concept of prevalence is that there is no measure of the number of the individual acts. This is of immense policy significance as a consequence of its implications in those crime statistics that are gathered by surveys. If domestic violence enters crime statistics as a 'course of conduct', then it counts as just one crime, even though there are usually several events within this 'course of conduct'. In this way, the repetition and frequency of the attacks disappears from view; is made invisible. This leads to an underestimate of the extent to which violent crime in general and domestic violence in particular is gendered and the extent to which it is from men to women rather than vice versa. The concept of prevalence does not contain a measure of severity and can lead to a survey finding of spurious gender symmetry. In those countries where surveys of inter-personal violence are addressed to men as well as to women, the focus on prevalence obscures the extent of gender inequality in the use of violence. Such an approach produces spurious findings of near gender symmetry that have perverse effects on the policy field.

47. It is possible to examine this issue empirically, using data from a special module on inter-personal violence in the British Crime Survey, a large-scale national sample survey carried out in Britain by the government's Home Office (Walby and Allen 2004). Using the concept of prevalence, at some point in the 12 months prior to interview, 4.2% of women in Britain had suffered domestic violence as compared with 2.3% of men (p.16), with an estimated 657,000 female victims and 356,000 male victims (p. 24); meaning that women were 1.8 times more likely than men to suffer some domestic violence in the last year. Using the concept of incidence, the gender ratio appears rather different. On average, women suffered 20 incidents, while men suffered 7; so there were an estimated 12.9 million incidents of domestic violence against women and 2.5 million against men (p.24); meaning that there were 5.2 times as many incidents of domestic violence against women as there were against men. When the concept of prevalence is used the gender asymmetry is rather mild; two thirds of the victims appear to be women and one-third of the victims appear to be men. When the concept of incidence is used, the gender asymmetry is much starker; 84% of the domestic violence incidents are against women and 16% against men (see Table 1²). This means that the use of the concept of prevalence rather than that

² Table 1: Comparing gender ratios using prevalence and incidents of domestic violence

| | Women | Men | % against women | Ratio female:male |
|----------------------------------|--------------|-------------|-----------------|-------------------|
| Victims | 657,000 | 356,000 | 35% | 1.8 |
| Average no. incidents per victim | 20 | 7 | | 3.9 |
| Total incidents | 12.9 million | 2.5 million | 84% | 5.2 |

Source: calculated from British Crime Survey (Walby and Allen, 2004).

of incidence to underpin the operationalisation of the measurement of domestic violence will, on the same set of data, produce an appearance of only slightly gender imbalance, while the use of the concept of incidence to underpin its measurement will show greater gender inequality.

48. A second disadvantage of the use of prevalence to underpin measurement of the extent of domestic violence is that it is a specific measure developed in the specialised field of violence against women, and is hard to mainstream into adjacent policy domains. There are two major adjacent policy fields: crime and health. The crime field has been particularly important in relation to data collection on violence against women in North America and Europe, while perhaps, health is more important in the developing world. Crime, law and order are policy domains that are high on the political agenda and generally well resourced (Garland 2001); in many countries it is more prominent and better resourced than the specific field of violence against women (Kelly 1999). The mainstreaming of violence against women into the field of crime would be to include it in a more resource rich environment as compared with the specialised field of violence against women.

49. The mainstreaming of violence against women into the crime field would probably be assisted if the two fields were to use the same concepts underpinning the measurement of the extent of these crimes. However, the measurement of crime is underpinned by the concept of incidence not prevalence. Crime is counted as the number of separate crime incidents per unit of population, per unit of time (often one year). If a person is multiply victimised during the time period then these additional crimes are included in the crime count. The significance of violence against women within the crime field might be seen more easily if the measurement of violent crime against women and other forms of violent crime were to use the same basis. While the violence against women field uses prevalence and the crime field uses incidence, the field of violence against women is likely to remain more marginalised than if they used the same method of counting violent crime.

50. The significance of the crime field concerns not only policy development, but also data collection. While in many countries there has been a single one-off data collection exercise on violence against women, insofar as these are repeated, these have tended (but not exclusively) to be where there is an interest in collecting this data alongside crime more generally. Crime surveys are a significant possible source of resources for the routine repetition of surveys of violence against women. Most crime surveys where this integration has occurred, have asked questions about gender-based violence of men as well as of women. Insofar as there are substantial plans to conduct repeat surveys on gender-based violence, they tend to be crime surveys where men as well as women are asked the questions. This raises the issue of the advantages of incidence rather than prevalence with additional urgency.

51. There is a productive tension in the development of indicators between on the one hand the detailed appreciated of distinctive nuances within a framework specific to the expert VAW field, and on the other hand, the use of more general categories that facilitate addressing priorities within mainstream policy arenas. Early attempts at specifying indicators tended to the former, while later developments have (rightly) tended to move towards the latter.

52. The notion of prevalence captures the particular and specialised nature of domestic violence as a coercive course of conduct, a series of related occurrences, rather than a one-off event. However, prevalence is not a concept widely used in the related domain of criminal

justice. The counting of the number of incidents, rather than the prevalence rate, is the more usual approach in crime statistics. If domestic violence enters crime statistics as a 'course of conduct', then it counts as just one crime incident, even though there are usually several events. In this way, the repetition and frequency of the attacks disappears from view thereby leading to underestimates of the extent of violent crime and domestic violent crime in particular. It is important that the number of incidents is the main measure of extent of violence against women.

Recommended options

- life-time prevalence;
- annual rate of the number of incidents per 1,000 women.

D. Severity and impact

Major options

- the nature of the action;
- frequency;
- injury: whether or not there is an injury, and if so its seriousness.

Discussion

53. It is important to have an indicator that captures the severity of the violence, as well as its extent. Most of the literature on severity has focused on domestic violence, and within this on physical violence; but this is not the only type of violence against women. The measurement of severity in relation to domestic violence will be considered first, and then the issue will be broadened to other types of violence.

Actions :

54. In relation to domestic violence, the most frequently used scaling of the severity of actions is that of the Conflict Tactics Scale (CTS) (Straus 1979, 2007; Gelles and Straus 1990). First used in US surveys in 1975 and 1985, it has been included as an element, with minor or major modifications, in most later surveys that address domestic violence in detail including Statistics Canada (Johnson 1996), the World Health Organisation (Garcia-Moreno et al 2005), the International Violence Against Women Survey (Nevala 2005), DHS (Kishor and Johnson 2004) and many individual national surveys (Walby and Myhill 2001; UNECE 2006). The Conflict Tactics Scale (CTS) consists of a list of items (slightly varied in different versions) of increasing levels of severity. The list is slightly varied in different versions (Straus 1979, 2007; Straus and Gelles 1990); the one reported here is that in Straus and Gelles (1990). The first items on the list are not considered acts of violence, but forms of verbal reasoning (A-C): A discussed an issue calmly; B got information to back up your/his/her side of things; C Brought in, or tried to bring in, someone to help settle things. The next set (D-J) are considered verbal aggression, but not violence: D insulted or swore at him/her/you; E sulked or refused to talk about an issue; F stomped out of the room or house or yard; G cried; H did or said something to spite him/her/you; I threatened to hit or throw something at him/her/you; J threw or smashed or hit or kicked something. The violent items are: K threw something at him/her/you; L pushed, grabbed, or shoved him/her/you; M slapped him/her/you; N Kicked, bit, or hit him/her/you with a fist; O Hit or tried to hit him/her/you with something; P Beat him/her/you up; Q Choked

him/her/you; R Threatened him/her/you with a knife or gun; S Used a knife or fired a gun. These violent items are divided into minor violence (K-M) and severe violence (N-S). Some later versions of the scale additionally include sexual violence, but the early ones did not.

55. Several problems have been identified with the Conflict Tactics Scale and with its deployment. There is concern that its use produces spurious gender symmetry. This is for at least three reasons including: the exclusion of sexual violence and stalking; the significance of meaning and context; the lack of congruency of behavioural actions with impact on the victim.

56. First, the early version of the scale was confined to physical domestic violence, excluding sexual violence and also stalking from an ex-partner; so it is not comprehensive (Dobash et al 1992; Johnson 1996). Although later versions of the scale did include an item on sexual violence, additional scales need to be used to capture the extent and range of sexual assault, stalking and other forms of domestic violence. Since women are almost all (though not entirely) the victims sexual violence (Walby and Allen 2004), its exclusion tilts the findings away from gender asymmetry. Its inclusion would show greater gender asymmetry in domestic violence.

57. Second, it does not take meaning and context sufficiently into account (Brush 1990; Dobash et al 1992; Smith 1994; Johnson 1996; Dobash and Dobash 2004). This criticism has several overlapping aspects. One is a preference for differentiating the intentions behind the action; it may be an initiation of aggression, or retaliation in response, or self-defence. It has been suggested that self-defence should be treated differently rather than identically with the initiation of aggression (Saunders 1988; Nazroo 1995; Dobash and Dobash 2004). A similar but slightly different aspect is concerned with the differential impact of the violence on women and men that is not sufficiently captured by the description of the act. For example, women are much more likely to be frightened and stay frightened than men as a consequence of acts of domestic violence (Mirrlees-Black 1999). This may mean that women are more likely to be controlled by the violence than men. The conclusion drawn here is that it is important to identify the seriousness of the impact of the act, not only the behavioural description of the action.

58. Third, the use of the actions of the perpetrator as the measure of severity may not be congruent with the severity of its effects on the victim; in particular, the injurious effect of the same action is greater when the perpetrator is a man and the victim a woman than vice versa (Schwartz 1987; Dobash et al 1992; Walby and Allen 2004). For example, the British Crime Survey reports that a minor act of domestic violence leads to an injury in 36% of instances in which men are the victims and 49% of instances where women are the victims; a severe act of domestic violence leads to an injury in 56% of the instances in which men are the victims and 77% of those where women are the victims (Walby and Allen 2004).

59. It is widely known that women suffer more injuries than men from domestic violence (Dobash and Dobash 2004). But this view can sometimes conflate the more frequent use of more severe acts by men than women with the issue as to whether the same act produces more injuries if committed by a man rather than a woman. Three issues need to be separated: frequency; severity of act; and seriousness of the injury from the same act. First, concerning frequency, women are subject to more incidents of domestic violence than men. The more frequent use of violent acts by men than women is a reason not to use a simple prevalence measure; since prevalence would treat multiple acts as equivalent with a single act. Second, the use of more severe acts by men than women is a further reason not to use a prevalence measure;

since prevalence would treat severe acts as equivalent with minor acts. Third, the more injurious effects of any given act by men than women (Walby and Allen 2004) is a reason not only not to use a prevalence measure, since prevalence would treat any act as equivalent with any other, but also for using injury as the measure of severity.

Injury

60. The major alternative measure of severity to the CTS is injury: whether, and if so, the seriousness of the injuries. This is a victim-focused measure of the severity of the impact of the violence. The existence or not of any physical injury is part of such a scale. A further distinction may be made as to the seriousness of the injury; many countries recognise in their law some distinction between more or less serious violent crimes.

61. The use of injury as an impact measure has three advantages: it is immediately meaningful; it avoids the problem of the differential gender impact of the same action; and it enables easier linkage to mainstream policy concepts, especially those of crime and health.

62. To be comprehensive, indicators would ideally include forms of injury beyond physical injuries. This especially concerns the injury constituted by sexual violence and injuries to mental health. Rape and other forms of penetrative sexual assaults are very serious injuries in their own right, perhaps not needing to be further translated. Other forms of sexual assault might be considered minor injuries in this scaling. These might be considered as separate scales from those of physical injury or (if numbers are small) added into the categories of minor and major injuries. There are some forms of violence that are common in some countries and rare but not unknown in others, such as: female genital mutilation/cutting. These should be separate named categories of violence against women in those countries where this is a common form of violence against women, but in those where it is much less common they should be included in the category of serious injury, or homicide. They should be separately named in surveys in those countries where these practices are known, and the decision as to whether to aggregate into the generic category of serious injury or not should be taken when it is known whether the numbers are statistically significant or not. A more difficult issue is that of the mental health injuries that are a common consequence of violence against women but which might be difficult to operationalise in form suitable for a survey. More complex surveys might include measures developed in the mental health field, but this may go beyond reasonable expectations for international standards for a VAW population survey. The most severe form of injury is death. Obviously this data cannot be collected by a population survey. Homicide is, however, widely regarded as the most robust of the administrative, police or criminal justice system statistics in relation to violent crime. Statistics on homicide are made available in internationally comparable gender disaggregated form by the World Health Organisation (Krug et al 2002). However, there is no widespread breakdown as to whether the homicide was domestic or not. Since there can be no assumption that the homicide of women is intrinsically gender-based, it would be inappropriate to use 'femicide' as an indicator of violence against women.

63. One advantage of 'injury' as a measure of severity is that it facilitates the mainstreaming of violence against women into adjacent policy fields such as the criminal justice system and health. This is important since these fields are well established and generally more resource-rich than the policy fields of gender equality or violence against women.

64. There is a tension in the development of scaling the severity of domestic violence between on the one hand a specialised rating system that is sensitive to the nuances of this particular form of violence and on the other hand the use of a more generic rating system that is used in a wider range of policy fields. The Conflict Tactics Scale is an example of a specialised system of categories that is, so far, uniquely used in the field of domestic violence. Since it is unique to domestic violence, it is harder to mainstream violence against women into other policy fields that use different forms of scaling of severity than if the same scaling of severity is used. Since the scale is unique to domestic violence, it makes comparisons with other forms of violence, such as violent crime, more difficult. For example, estimating the extent of violent crime that is domestic is hindered by its distinctive conceptualisation and measurement.

Frequency

65. A further possible indicator of severity of violence against women is that of the frequency of the attack. The number of times that the same person is subject to violence is a measure of severity. This is different from the number of incidents per population unit, being rather the average number of incidents per person subject to the violence. The main advantage of such an indicator is that it is readily meaningful. However, there is a major disadvantage in that frequency overlaps with the concept of the extent of violence against women. If it is used in the measurement of the extent, it should not also be included in the measurement of severity.

66. The main measure of severity should be indicated by injury. The simplest level is injury or no injury. A more differentiated measure would be: no physical injury but fear alarm or distress; minor injury; severe injury; death.

Recommended option

- Injury (not action or frequency)
 - Severity of injury:
 - no physical injury, but fear, alarm or distress;
 - minor injury: from bruising to bleeding, non-penetrative sexual assault;
 - major injury: broken bones or teeth, attempted strangulation, rape and other penetrative sexual assaults, FGM;
 - death

E. Time-period

Major options

- Lifetime
 - Or since adult e.g. 15, 16
- Last year
 - Or last 3 or 5 years.

Discussion

67. There is a question as to the time period over which to gather information on violence against women. Most surveys considered by UNECE (2006) had information on two time periods: lifetime and also a short recent period, usually one year, but sometimes three or five years.

68. The life-time measures are important in establishing the scale of the problem, for consciousness raising and raising its priority (Garcia-Moreno et al 2005). It is important for those forms of violence that can occur only once in a life-time, and those that are unlikely to occur more than once in a life-time. The life time measure is perhaps the easiest time period to address in a survey and in the early stages of development of this field, where techniques are not sophisticated and resources not well developed. However, it is of little use in assessing the impact of policy change, which requires a shorter time period.

69. Evaluation of innovative policies requires regular measurements of the amount of violence against women over time, for which the 'last year' rates are more appropriate. Most data in other policy fields uses a one-year time period, and it would be advantageous to match this. However, reliable data over the shorter time period requires a much larger sample size, because of the lower proportion experiencing violence last year as compared with over a life-time, and thus entails a more expensive survey. This issue highlights the practical resource issues involved. While many countries have now conducted one-off surveys of the life-time prevalence of violence against women (Garcia-Moreno et al 2005; Krug et al 2002; Presidencia de la Unión Europa 2002a, 2002), the collection of adequate data on an annual basis is rare, with annually collected data being almost entirely confined to methods, such as generic crime surveys, which are known to underestimate the levels of violence against women (Walby and Myhill 2001). The challenge is to develop methods to generate adequate data that has a one year reference period, using a realistically available amount of resource. One option here is for a survey to be conducted not every year, but once every two, three or even five years, but to use a one-year period of recall. A shorter period of recall can aid the accuracy of recall, although there are other techniques available to reduce the 'telescoping' effect (Muratore 2006). However, the shorter the period, the larger needs to be the sample size, so this is more costly.

Recommended options

- life-time, and
- 'last year'.

F. Population sub-groups

70. There are practical issues in survey methodology, which lead to the surveyed population being restricted in some ways, in particular concerning age and marital status. These differences have resulted in a serious lack of comparability of surveys (Schröttle et al 2006).

Age

Major options

- All adult women
 - Adulthood starting at 15 or 16
- All adult women up to an upper age limit
 - 45 or 59 or 65 or 70

Discussion

71. There is a distinction between restrictions on the age at which the violence took place and restrictions on the age of the respondent being questioned. The age at which the violence took place has been addressed above. This discussion concerns the age of the respondent: There are separate issues for lower age and upper age cut-off. The most important issue is consistency between surveys, rather than the exact ages selected. Most surveys have a lower cut off around adulthood, though the actual age varies, predominantly between 15 and 18. Many but far from all surveys have an upper cut off. The criteria for this include reproductive age of 49 (in some health surveys), to ability to use a computer unaided by interviewer of 59, with further ages representing various expectations of competence.

Recommended option

- Age at which violence took place: to include all years.
- Age of respondent: 16-65 years old.

Marital and Cohabiting Status

Major options

- Adult women regardless of marital status
- Currently married or partnered women
- Ever married or partnered women

Discussion

72. Some surveys, especially where the focus is domestic violence, are restricted to women who are currently or ever have been married or cohabiting. This is unduly restrictive, since violence against women can take place outside of marriage and cohabitation.

Recommended option

- No restrictions on marital status.

IV. SUMMARY OF MAJOR OPTIONS AND PROPOSALS FOR INDICATORS

73. This final section first lists the major and recommended options for each of the dimensions of indicators: definitions of types of violence; units for the measurement of extent; severity and impact; time period; and population sub-groups. Next it integrates the recommended options on each dimension into proposed indicators that are specified on each of these dimensions.

Major and recommended options on each dimension of the indicators

Definition of forms of violence

The major options for indicators

- A single indicator that includes all forms of gender-based violence against women and which does not separately specify them.
- A separate indicator for each form of violence.
- A few indicators, one each for the more major forms of violence against women.
- A single indicator confined to domestic violence.

Recommended options

74. For international comparisons:
- A main indicator that includes all of the types of violence and does not separately specify them (although they are separately named in the data collection instrument).
 - Additional separate indicators for the major types of violence that are found in all countries: domestic violence; rape; sexual harassment in the workplace.
75. For national purposes in addition:
- Additional separate indicators for those forms of violence that are common in that country, but which are not commonly found in all countries, such as female genital mutilation/cutting.

Units for the measurement of extent: prevalence and incidents

Major options

- Prevalence: rate (%) of violence against women in the female population
- Incidents: number of incidents of violence against women per unit (e.g. 100, or 1,000) of female population

Recommended options

- life-time prevalence;
- annual rates of number of incidents of violence against women per 1,000 women.

Severity and impact

Major options

- the nature of the action (Conflict Tactics Scale and its variants);
- frequency;
- injury: whether or not there is an injury, and if so its seriousness.
-

Recommended option

- Injury (neither action nor frequency)
 - Severity of injury:
 - no physical injury, but fear, alarm, distress or pain.
 - minor injury: such as bruising, scratches, bleeding, non-penetrative sexual assault.
 - major injury: broken bones or teeth, attempted strangulation, rape and other penetrative sexual assaults, trafficking, FGM.
 - death

Time-period

Major options

- Lifetime
 - Or since adult e.g. 15, 16
- Last year
 - Or last 3 or 5 years.

Recommended options

- life-time, and
- ‘last year’.

Population sub-groups as survey respondents

Major options

- All adult women
 - Adulthood starting at 15 or 16
- All adult women up to an upper age limit
 - 45 or 59 or 65 or 70

Recommended option

- 16-65 years old

Marital status

Major options

- Adult women regardless of marital status
- Currently married or partnered women
- Ever married or partnered women

Recommended option

- No restrictions on marital status.

Proposed indicators integrating all dimensions

76. Two indicators are proposed:

- 1). Life-time prevalence of any form of gender-based violence against women, differentiated by level of severity of injury (no injury, minor injury, major injury, death), expressed as a percentage of the total female population;
- 2). Annual number of incidents of any form of gender-based violence against women, differentiated by level of severity of injury (no injury, minor injury, major injury, death), expressed as a rate per 1,000 women, aged 16-65 years old.

77. Additional indicators:

- 3). In some instances, it will be both desirable and possible to make further differentiations within these indicators as to sub-types of violence against women, in particular, intimate partner violence; rape; and sexual harassment. This is not proposed as an alternative to the two main indicators, but as an addition.

V. BIBLIOGRAPHY

Alm eras, Diane, Rosa Bravo, Vivian Milosavljevic, Sonia Monta o and Maria Nieves Rico (2004) *Violence Against Women in Couples: Latin America and the Caribbean. A Proposal for Measuring its Incidence and Trends*. *Mujer y desarrollo* 40. (Santiago, Chile: UN Economic Commission for Latin America and the Caribbean).

Berger-Schmitt, R. and B. Jankowitsch (1999) *Systems of Social Indicators and Social Reporting: The State of the Art*. EUReporting Working Paper No.1. Mannheim: Centre for Survey Research and Methodology.

Brush, Lisa D. (1990) 'Violent acts and injurious outcomes in married couples: methodological issues in the national survey of families and households', *Gender and Society*, 4, 1:56-67.

CAHRV (2007) *Coordination Action on Human Rights Violations* <http://www.cahrv.uni-osnabrueck.de/reddot/>

Dobash, R.P., R. E. Dobash, M. Wilson and M. Daly (1992) 'The myth of sexual symmetry in marital violence', *Social Problems*, 39, 1:71-91.

Dobash, Russell P. and R. Emerson Dobash (2004) 'Women's violence to men in intimate relationships: Working on a puzzle', *British Journal of Criminology*, 44, 3, 324-349.

Dublin Declaration (2003) Dublin Conference on Organised Crime: Doc. 16302/03 CRIMORG 96 of 19 Dec 2003.

European Commission (2006a) *Developing a Comprehensive and Coherent EU Strategy to Measure Crime and Criminal Justice: An EU Action Plan 2006-2010*. Eurostat/F/06/DSS/10/12/Annex 1/EN.

European Commission (2006b) *Road Map for Equality between Women and Men, 2006-2010* (SEC 2006 275) http://eur-lex.europa.eu/LexUriServ/site/en/com/2006/com2006_0092en01.pdf

European Economic and Social Committee (2006) *Opinion on Domestic Violence Against Women* Soc/218 Brussels 16 March 2006 http://eescopinions.esc.eu.int/EESCopinionDocument.aspx?identifier=ces%5Csoc%5Csoc218%5Cces416-2006_ac.doc&language=EN

European Union (2004) *Council Resolution on the review of the implementation by the Member States and the EU institutions of the Beijing Platform for Action*. 1502/04. SOC 558. JAI 497. Brussels.

Fals-Stewart, William, Gary R. Birchler and Michelle L. Kelly (2003) 'The timeline followback spousal violence interview to assess physical aggression between intimate partners: Reliability and validity', *Journal of Family Violence*, 18, 3, 131-142.

Federal-Provincial-Territorial Ministers Responsible for the Status of Women (2002) *Assessing Violence Against Women: A Statistical Portrait*. (Spain).

Garcia-Moreno, Claudia, Charlotte Watts, Henriette Jansen, Mary Ellsberg and Lori Heise (2005) *WHO Multicountry Study on Women's Health and Domestic Violence against Women*.

Garland, David (2001) *The Culture of Control: Crime and Social Order in Contemporary Society*. (Oxford: Oxford University Press).

Hague Programme (2005) *Official Journal of the European Communities* OJ C53 of 3.3.2005, p.1.

Heiskanen, Markku and Minna Piispa (1998) *Faith, Hope, Battering - A Survey of Men's Violence against Women in Finland*. (Yliopistopaino, Helsinki: Statistics Finland).

Hélie, Sonia, Marie-Eve Clément, and Marie-Claude Larrivée (2003) 'Epidemiological considerations in the conceptualization and utilization of "prevalence" and "incidence rate" in family violence research: A reply to Brownridge and Halli (1999)', *Journal of Family Violence*, 18, 4, 219-225.

HEUNI (2007) *International Violence Against Women Survey* <http://www.heuni.fi/12859.htm>

Home Office (2005) *Domestic Violence: A National Report* (London: Home Office)
<http://www.crimereduction.gov.uk/domesticviolence/domesticviolence51.pdf>

ISTAT (2007) *Violence and Abuses against Women Inside and Outside Family, 2006*. (Italy: ISTAT).

Johnson, Holly (1996) *Dangerous Domains - Violence Against Women in Canada*. (Canada: Nelson Canada).

Johnson, Michael P. (1995) 'Patriarchal terrorism and common couple violence: two forms of violence against women', *Journal of Marriage and the Family*, 57, 283-94.

Johnson, Michael P. (2006) 'Conflict and control: Gender symmetry and asymmetry in domestic violence', *Violence Against Women*, 12, 11, 1003-1018.

Kelly, Liz (1999) 'Violence against women: a policy of neglect or a neglect of policy', in Sylvia Walby (ed) *New Agendas for Women*. (Basingstoke: Macmillan).

Kelly, Liz with Lorna Kennedy and Miranda Horvath (2006) *The Next Step: Developing Transnational Indicators on Violence Against Women*. Report prepared for the UN Rapporteur on Violence Against Women; presented to the UNECE Task Force on Measurement of Violence Against Women.

Kishor, Sunita and Kiersten Johnson (2004) *Profiling Domestic Violence: A Multi-Country Study*. (Calverton US: Measure DHS+, ORC Macro).

Krug, Etienne G., Linda L. Dahlberg, James A. Mercy, Anthony B. Zwi and Rafael Lozano (eds) (2002) *World Report on Violence and Health*. (World Health Organization: Geneva, WHO).

Luxembourg Presidency (2005) *Report of the Luxembourg Presidency on the Progress made by the Enlarged EU following the Beijing Platform for Action (Beijing +10)*. Prepared by: Anne Marie Theisen, Nadine Spoden, Mieke Verloo and Sylvia Walby. Luxembourg: Luxembourg Presidency.

Martinez, Manuela and Monika Schröttle, with Stephanie Condon, Marianne Springer-Kremser, Greetje Timmerman, Carol Hagemann-White, Hans-Joachim Lenz, Corinne May-Chahal, Bridget Penhale, Jolanta Reingardiene, Petra Brzank, Päivi Honkatukia, Maryse Jaspard, Eva Lundgren, Minna Piispa, Patrizia Romito, Sylvia Walby and Jenny Westerstrand (2006) *State Of European Research On The Prevalence Of Interpersonal Violence And Its Impact On Health And Human Rights*. Report by the Co-ordination Action on Human Rights Violations (CAHRV) to the European Commission, 6th Framework Programme, Project No. 506348.
[http://www.cahrvi.uni-osnabrueck.de/reddot/CAHRVreportPrevalence\(1\).pdf](http://www.cahrvi.uni-osnabrueck.de/reddot/CAHRVreportPrevalence(1).pdf)

MEASURE Evaluation (2007) *Violence Against Women: A Compendium of Indicators for Monitoring and Evaluating Programs*
http://www.cpc.unc.edu/measure/leadership/violence_against_women

Millennium Development Goals (2005) *Millennium Development Goals*. Accessed at:
www.developmentgoals.org

Muratore, Maria Giuseppina (2006) 'Results from the questionnaire on Victim Surveys on telescoping, non response and methodological challenges'. Paper presented to UNECE-UNODC Meeting on Crime Statistics, Vienna, 25-27 January 2006.

Nazroo, J. (1995) 'Uncovering gender differences in the use of marital violence: the effect of methodology', *Sociology*, 29, 3, 475-94.

Nevala, Sami (2005) *International Violence Against Women Survey*. Paper presented to UN Expert Group meeting, Geneva. <http://www.un.org/womenwatch/daw/egm/vaw-stat-2005/docs/expert-papers/Nevala.pdf>

Presidencia de la Unión Europea (EU Presidency) (2002) *Good Practice Guide to Mitigate the Effects of and Eradicate Violence Against Women*. Paper prepared by the Institute of Women, Spanish Ministry of Labour and Social Affairs and by María Diaz Aguado and Rosario Martínez Arias of the Universidad Complutense de Madrid. Accessed at
<http://www.mtas.es/mujer/guiaing.pdf>

Saltzman, Linda E., Fanslow, Janet L., McMahon, Pamela M., and Shelley, Gene A. (1999) *Intimate Partner Violence Surveillance Uniform Definitions and Recommended Data Elements, Version 1.0*. (2nd printing with revisions 2002). (National Center for Injury Prevention and Control, Centers for Disease Control and Prevention: Atlanta, USA).

Saunders, D.G. (1988) 'Wife abuse, husband abuse, or mutual combat?' in Yllo, K. and Bograd, M. (eds) *Feminist Perspectives on Wife Abuse*. (California: Sage).

Schröttle, Monika, Manuela Martinez, Stephanie Condon, Maryse Jaspard, Minna Piispa, Jenny Westerstrand, Jolanta Reingardiene, Marianne Springer-Kremser, Carol Hagemann-White, Petra Brzank, Corinne May-Chahal, Bridget Penhale, with Eva Lundgren, Greetje Timmerman and Hans-Joachim Lenz (2006) *Comparative Reanalysis Of Prevalence Of Violence Against Women And Health Impact Data In Europe – Obstacles And Possible Solutions: Testing A Comparative Approach On Selected Studies*. Report by the Co-ordination Action on Human Rights Violations (CAHRV) to the European Commission, 6th Framework Programme, Project No. 506348.
http://www.cahrv.uni-osnabrueck.de/reddot/D_20_Comparative_reanalysis_of_prevalence_of_violence_pub.pdf

Schwartz, M.D. (1987) 'Gender and injury in spousal assault', *Sociological Focus*, 20,1,61-75.

Smith, M.D. (1994) 'Enhancing the quality of survey data on violence against women: a feminist approach', *Gender & Society*, 8, 1, 109-27.

Statistics Canada (2002) *Assessing Violence Against Women: A Statistical Profile*. Federal Provincial Territorial Ministers Responsible for the Status of Women and Statistics Canada. (Ottawa: Statistics Canada).

Straus, Murray A. (1979) 'Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales', *Journal of Marriage and the Family*, February, 75-88.

Straus, Murray A. (2007) 'Conflict Tactics Scales', in N. A. Jackson (ed.) *Encyclopedia of Domestic Violence*. (New York: Routledge, Taylor and Francis). Pp. 190-197.

Straus, Murray A. and R. J. Gelles (eds) (1990) *Physical Violence In American Families*. (New Brunswick, New Jersey: Transaction Publishers).

Tjaden, Patricia and Nancy Thoennes (2000) *Full Report of the Prevalence, Incidence and Consequences of Violence Against Women*. NCJ 193781. (US Department of Justice, National Institute of Justice: Washington DC, USA).

United Nations General Assembly (1993) *Declaration on the Elimination of Violence Against Women*. UN General Assembly 20 Dec 1993. UN Doc. A/RES/48/104. (New York: UN).

United Nations Division for the Advancement of Women (2005a) Report of Expert Group on 'Violence against women: a statistical overview, challenges and gaps in data collection and methodology and approaches for overcoming them'. 11-14 April 2005, Geneva Switzerland.

United Nations Division for the Advancement of Women (2005b) Report of Expert Group on 'Good practices in combating and eliminating violence against women'. 17-20 May 2005, Vienna Austria.

United Nations Economic Commission for Europe Task Force on Measurement of Violence Against Women (UNECE) (2006) *Analysis of National Surveys Carried out by Countries of the*

Conference of European Statisticians to Measure Violence Against Women Presented to conference of UNECE, Statistical Commission and Conference of European Statisticians. ECE/CES/GE.30/2006/6.

United Nations High Commissioner for Refugees (2003) *Sexual and Gender-Based Violence against Refugees, Returnees and Internally Displaced Persons*. (Geneva: UNHCR).

United Nations General Assembly (2006) *In-Depth Study on all Forms of Violence Against Women: Report of the Secretary-General A/61/122/Add.1*
<http://daccessdds.un.org/doc/UNDOC/GEN/N06/419/74/PDF/N0641974.pdf?OpenElement>

Walby, Sylvia (2004) *The Cost of Domestic Violence*. (London: Department of Trade and Industry Women and Equality Unit).

Walby, Sylvia (2005a) 'Measuring women's progress in a global era', *International Social Science Journal*, June, 184, 371-387.

Walby, Sylvia (2005b) 'Improving the statistics on violence against women', *Statistical Journal of the United Nations Economic Commission for Europe*, 23, 3-4.

Walby, Sylvia and Jonathan Allen (2004) *Domestic Violence, Sexual Assault and Stalking: Findings from the British Crime Survey*. Home Office Research Study 276. (London: Home Office). <http://www.homeoffice.gov.uk/rds/pdfs04/hors276.pdf>

Walby, Sylvia and Andrew Myhill (2001) 'Comparing the methodology of the new national surveys of violence against women', *British Journal of Criminology*, 41, 3, 502-22.

VI. APPENDIX: DATA COLLECTION REQUIREMENTS AND CONSTRAINTS AND OPPORTUNITIES FOR OVERCOMING THESE

Introduction

1. The reliability of these indicators depends upon the data collected to support them. This requires the development of high quality surveys, which use sophisticated methods, to agreed international standards, and which are carried out regularly.
2. There are now many nationally representative sample surveys of violence against women (UNECE 2006). There is very considerable experience in the development of survey methodology for collecting data on violence against women; many national surveys have been carried out in the North and the South (Alméras et al 2004; Johnson 1996; Tjaden and Thoennes 2000; Walby and Allen 2004), as well as developments in comparative work using the same instrument in groups of countries (Garcia-Moreno et al 2005; Kishor et al 2004; Nevala 2005). The methodology has reached a high level of sophistication (UN 2005a), though there remain a series of methodological dilemmas and challenges (Straus and Gelles 1990; Johnson 1996; Walby and Myhill 2001; Garcia-Moreno et al 2005; Walby 2005b).
3. During the development of statistical data on gender based violence, there has been much reflection on and refinement of methodological dilemmas and challenges. These include: the importance of population surveys; sample size; free-standing survey or add-on module; the achievement of a comprehensive sampling frame; the best method of delivery of the questionnaire and the achievement of a high response rate; and whether confidentiality or rapport with the interviewer is more likely to facilitate disclosure of sensitive events.

Sample Size

4. In order to be able to deliver statistically significant data on changes in the scope, prevalence and incidence of violence against women it is necessary to have a large sample size. This is especially important if there is interest in separate identification of different forms of violence, and if there is interest in evidence of the extent of annual change. A sample size of 10,000 is the absolute minimum to be recommended, and much larger samples are needed to deliver statistically significant data on changes in specific forms of violence.

Free-standing or add-on module?

5. Data on gender based violence is positioned in surveys in three ways. One practice is to include a few questions within a mainstream survey. A second is to conduct a free-standing dedicated survey in its own right. A third is to add a relatively self-contained module onto a more mainstream survey. There is a tension here between creating the best possible survey environment and the need to keep costs down.
6. The inclusion of a few questions on violence against women within an already existing survey has the advantage of least cost, but the disadvantage of collecting a restricted amount of information in an interview context that has priorities other than violence against women.

7. The advantage of a dedicated survey is that it can be tailor made for the methodological needs of investigating violence against women; the disadvantage is the practical one that the resource base for an annual dedicated survey is less likely to be available than for a survey that also gathers information on mainstream matters.
8. There are a number of issues about interviewing where the priorities of a generic and a dedicated survey may diverge. For instance, an interviewer who is both female and specially trained has an advantage in eliciting responses in such sensitive areas, however, the prioritisation of such practices in interviewing is less likely to be achieved when there is a generic rather than dedicated survey. Further, a private context for the interview results in higher rates of disclosure than when there is someone else present in the room (Walby and Allen 2004). The achievement of this private context for the interview may be more likely to be prioritised in the context of a dedicated than a generic survey, but there is no intrinsic reason why a generic survey context could not deliver this.
9. A compromise solution is the use of a relatively self-contained module on gender based violence, which is attached to a mainline survey on a related topic. This uses an existing well-resourced annual national survey that has an established base of expertise and resources and attaches to this a self-contained, specially introduced, specialised module of relevant questions. The cost of the additional questions is modest as compared with the establishment of a dedicated annual survey. The UK BCS has experimented with this form (Walby and Allen 2004) and is now committed to regularly attaching special modules on domestic violence, sexual assault and stalking (Home Office 2005).
10. While the ideal is the dedicated survey, it is likely that cost restrictions would prevent regular repetition. Thus, it is probable that a self-contained module on gender based violence, attached to a related survey may be the most pragmatic solution.

Sampling Frame and Response Rate

11. The ideal sampling frame is one that includes all members of the population. This is most closely approached in the census, but many other national surveys may fall short of this ideal to varying degrees, often for reasons of time and resources. The ideal response rate is very high. A comprehensive sampling frame and a high response rate are probably of greater importance in surveys of violence against women than in many other surveys, since it is likely that those who fall outside the sampling frame or are not reached or do not respond are more likely to have been subject to violence than those who have not. It is the more 'marginal', excluded and disadvantaged groups of women who are most likely to have been subject to violence, especially in the near past, and these are precisely the groups that are most likely to be omitted if short cuts or economies are taken with the development of the sampling frame and survey instrument. While for many other types of surveys the omission of this section of the population from the sampling frame may not be considered sufficiently important to be worth the expense and effort to include them, for surveys on violence against women this is a potentially significant omission.
12. For example, women who have fled to refuges, to temporary residence with friends and kin, to emergency bed and breakfast or hostel accommodation, or who are homeless in the immediate aftermath of a domestic assault are most likely to be omitted from sampling frames

and to have low levels of response to the survey. Samples based on women who have gone to refuges and shelters have consistently shown much higher rates of frequency of abuse than those from national surveys (Dobash and Dobash, 1979; Straus and Gelles 1990).

13. The different profiles and especially the differing gender asymmetry of the profiles of victims of domestic violence drawn from population samples and from refuge samples has fed considerable controversy over the 'real' profile and its implications for theories of domestic violence (Straus and Gelles 1990; Dobash et al 1992). One proposed solution suggested that there were different types of domestic violence: both 'patriarchal terrorism', where men terrorised their intimate partners with repeated severe violence, and 'common couple violence', where women also engaged in violent acts, but where the seriousness of the violence was much lower (Johnson 1995).

14. The omission of the most heavily abused section of the population is a problem for a survey attempting comprehensive coverage and accurate estimates. It may lead to spurious conclusions being drawn. There are two main approaches to addressing this issue. One concerns the response rate; the second the sampling frame itself.

15. There are many techniques available to ensure a high response rate. In an analysis of crime surveys, Muratore (2006) finds that the most successful strategies were those of training programmes for interviewers, using female interviewers, and using personal visits. Other strategies included a series of post-hoc responses such as repeated contacts, monitoring interviews and using senior researchers to address non-response.

16. There are ways of supplementing the sampling frame to include marginalised populations, which could enhance future surveys. These include drawing up additional sampling frames based on lists of hostels, refuges, and other temporary accommodation that could be provided by those who fund and run such accommodation. However, this is hard to achieve; though UNECE (2006) reports that three surveys have included institutions as well private households in the sampling frame. A further possible solution would be to revise the procedure for sampling the person within private residential households so as to include all who are actually staying there, including temporary visitors, not merely those who are permanently resident.

17. The recommendation is that considerable priority is given to obtaining a comprehensive sampling frame and a very high response rate, while recognising that this can be expensive.

Mode of delivery of questionnaire

18. Surveys have been carried out using: postal questionnaires, telephone, face-to-face interviewing, and by self-completion on a computer. While some suggest that there is little evidence that it makes much difference, others have argued for particular methods, especially either telephones, or telephone or face-to-face, or for self-completion by computer (Walby and Myhill 2001).

19. Postal questionnaires usually have the lowest response rate of all methods, so are usually considered inappropriate for those surveys where this is important, as is the case in of surveys of violence against women. However, Statistics Finland used a postal questionnaire and obtained a

surprisingly high response rate of 70% (Heiskanen and Piipisa, 1998). This might be explained in terms of the unique features of Nordic society.

20. In some countries landline telephones have been used to make contact with respondents. Whether this provides a comprehensive sampling frame depends on two major factors that vary between countries. In some countries, such as Canada, coverage is nearly comprehensive, while in others such as the UK, telephone ownership rates in private households in Britain are particularly low among the poor heads of lone adult households (Beerton and Martin, 1999), and these are likely to include disproportionate numbers of women who have fled a violent home.

21. A further problem with telephone surveys today is the use of mobile phones for which random dialling techniques are not available, thus excluding those who have only mobile phones from the sampling frame. These are likely to be disproportionately young, probably leading to a skewing of the results, since younger people tend to have higher rates of violence.

22. The mode of enquiry is relevant because of its potential to not only reduce the response rate but to omit key sections of the population that have been found to be most at risk. The best approach will vary between countries at different levels of development of communication infrastructure. It may that sampling houses is the best way to obtain a comprehensive sampling frame, although this has implications for cost.

Self-completion: Rapport or confidentiality?

23. Is rapport or confidentiality more conducive to disclosure of events that may be sensitive? On the one hand there is the possibility that face-to-face interviewing can build up more rapport and support disclosure of sensitive events, while on the other hand, confidentiality engendered by strategies such as self-completion by computer or by postal questionnaire may increase the likelihood of respondents divulging sensitive information.

24. The British Crime Survey provided an opportunity to compare the outcome from face-to-face interviewing with that of the confidentiality of self-completion. In the first part of the interview, face-to-face interviewing is used; in the second, there is self-completion of specialised modules, with the computer handed over to the respondent (Computer Assisted Self-Interviewing, CASI). The respondent reads the questions from the computer screen and enters their responses into the computer. Only the respondent can see the questions and the answers they have given. The prevalence of domestic violence is five times higher when the more confidential self-completion methodology is used, as compared with the more traditional face-to-face interviewing. While there are other differences between the two questionnaires that may contribute to the differences, nonetheless, the scale of the difference suggests that confidentiality is more important than rapport in facilitating the disclosure of domestic violence (Walby and Allen 2004). One disadvantage of this method is that it depends upon literacy, which may be unevenly distributed within some populations. The model described here requires lap-top computer- based interviewing, which, while increasingly common in developed countries is expensive.

25. It may be that other forms of self-completion may also have similar effects in facilitating disclosure. UNECE (2006) report that six surveys used some form of self-completion: Computer Assisted Self-Interviewing (CASI) in two cases; postal questionnaires in four.

26. The recommendation is to deliver the questionnaire in a manner that prioritises confidentiality in order to facilitate disclosure.

Summary

27. In the development of survey methodology the priorities are: a large sample size; the use of a comprehensive sampling frame and the use of modes of enquiry that generate a high response rate, and so as to include the more marginal groups of women who are most likely to have been at risk. In the delivery of the questionnaire, the methodological priority is to prioritise confidentiality in order to facilitate disclosure; this is confidentiality not only from other household members, but also from the interviewer. The need for annual surveys in order to generate data about experiences last year, not only across the life-time, brings a sharp focus to the issue of cost and the advantages of mainstreaming. Hybrid surveys, in which special modules are attached to mainline surveys, enable annual collection of data, while maintaining specialist framing and modes of questioning, especially that of self-completion.
