

## Part IV. Comparability of data across countries

### What is comparability and why is it desirable?

830. The accurate comparison of crime statistics across countries is difficult and should always be approached with caution. It is necessary to be clear about exactly what is being compared and why. The reasons why it may be desirable to compare results across countries can include:

- Fostering the exchange of information relating to crime and to the functioning of criminal justice systems at regional and international levels;
- Transparency and accountability in crime prevention and the operation of law enforcement and criminal justice systems;
- The development of common benchmarks and indicators for assessing the nature, extent and public perception of crime.

831. For the purposes of victimization measures produced by surveys, it should be made clear that any comparison of results between different countries is essentially a comparison of the (combined) *experience* and *attitude* of persons towards crime in those countries. Such cross-national comparison of these characteristics may serve useful policy ends, enabling countries to learn from the experience of others and possibly to gauge the impact of regional or sub-regional crime prevention or criminal justice programmes.

832. When comparing results cross-nationally, it must be remembered that respondents to crime victim surveys perceive (and remember) that they have been the *victim of a crime*. The interpretation of what constitutes a crime may vary from country to country. For example, respondents may be more or less used to violent crime, which may occur in diverse contexts. Juvenile gangs may be more or less active and organized, the availability and use of weapons may largely differ from country to country. The same applies to property crime. While theft and burglary may affect citizens in affluent societies by depriving them of cash, jewellery, hi-tech equipment, art objects and furniture, the same type of crime in developing countries may entail theft of basic household items such as crockery, cutlery and linen. Even though the monetary loss may be small, the impact on the victim is significant. In many developing countries where cattle are a vital resource, theft of livestock is perceived by victims as one of the most serious victimization experiences. The same may apply to the theft of a bicycle in a context in which that is the main means of transportation. Questions on bribery and corruption for example, are very susceptible to the different interpretation of respondents from different countries as to what constitutes such events. To some extent, the identification of inherent differences in a target population – such as acceptance of petty theft for example, and its subsequent under-reporting in a crime victim survey – is a valid criminological finding of the survey, and does not prevent comparison with other surveys *per se*. Nonetheless, such differences must always be borne in mind and figures should never be compared in isolation from the qualitative context.

833. If comparisons with results from other countries are to be made, the context in which the surveys have been carried out must be made clear. Otherwise, results (such as very low reported levels of minor theft in particular countries) may be especially misleading. The blind ‘ranking’ of quantitative results from different countries, with little or no explanation as to the context or factors that may affect comparability, is generally not therefore appropriate. Overall, crime victim surveys may be considered most ‘comparable’ where the methodology used in each survey aims to consistently capture information about equivalent crime events, and population perceptions and cultural practices are recognised and clearly described.

### **Comparability challenges of crime victim surveys**

834. A key difference between the measurement of crime through official police statistics and through victim surveys is that the data collection process for victim surveys is significantly more *controllable*. Measurement using administrative data is reliant on existing processes, information systems, and data-recording protocols (or, frequently, absence of protocols). Either way, the administrative data collected is already historic, or past, data. Decisions as to mode of recording have already been taken. Where a decision was taken by the police not to record a certain event, nothing can later be done in the data collection process to influence this past decision, or even to identify the existence of the non-recorded event. As such, the comparability challenge in administrative statistics is to seek as much information as possible about the manner and time of police data recording, about operational definitions used, about counting rules and the ‘threshold’ that causes a record to be made, and to attempt to control for such differences so far as possible. In other words, the process involves trying to correct for earlier, existing, and unavoidable differences in administrative data recording between countries.

835. When it comes to crime victim surveys on the other hand, data that may ultimately be compared cross-nationally is not yet in existence before the survey; at least not in any ‘captured’ form. The crime victim survey organiser is therefore at an advantage. He or she, as the primary data gatherer, may control the primary data collection process with a view to introducing practices that standardise data collection and enhance the potential comparability of results with those from other crime victim surveys, carried out either earlier in time, or in different geographical territories.

836. Despite such potential advantages, however, it should not be a foregone conclusion that results from crime victim surveys are automatically comparable between countries. Indeed, the use of victim surveys, whilst eliminating certain problems associated with administrative statistics, at the same time introduce a large number of new obstacles to effective comparability across countries and over time.

837. Whilst the crime victim survey organiser is at somewhat of an advantage, he or she also faces many challenges. The range of factors that may affect the comparability of results is great. From the method used to collect data, whether in-person interviews or telephone interviews, to the wording of survey questions, to sample composition and size, to the content of the questionnaire, it is no easy task to ensure that cross-national differences between each are as small as possible. Finally, the surveyor is interacting with a largely unpredictable entity; the individual person.

Quantitative data gathered from a crime victim survey only arises as a result of the (prompted) subjective recall of an individual person to particular events that they have experienced. The challenge is to ensure that the events persons are asked to recall are of close equivalence across cultures and countries, are recalled with equivalent accuracy, and that the persons asked are equally representative of the cultures and countries of which they are a part.

838. It should also be remembered that whilst cross-national comparison of surveys does present particular challenges, national surveys themselves may also have to deal with different languages, ethnic groups, social groups, and degrees of education within a single country. Surveys are commonly required to produce results for heterogeneous populations and comparability within and between population sub-groups should be a basic design component of any survey.

839. This part of the Manual addresses particular issues that may affect comparability. It examines why anomalies may arise between different surveys and steps that can be taken to enhance comparability.

### **Factors that affect comparability**

840. Problems of comparability arise from two principle sources:

- (i) Methodological differences in the conduct of the survey itself, such as sample selection and method of interviewing (these may be controlled for as far as possible); and
- (ii) Inherent differences between target populations, including different perceptions of crime or widespread cultural acceptance of criminal practices (difficult to control for).

841. Methods to increase the cross-national comparability of surveys can be divided into two types: '*output harmonisation*' and '*input harmonisation*'. *Input* harmonisation involves action taken to enhance comparability by considering aspects of the questionnaire design and survey implementation process itself. *Output* harmonisation involves selecting particular variables for comparison or making adjustments to the estimates after the data have been collected (such as performing age- and gender standardization, or using adjustments such as purchasing power parity). Choosing only survey results from countries that share common characteristics, such as geographic location, social structure and culture, or degree of human development for comparison is a form of output harmonisation.

842. Output harmonisation may be a valuable method of increasing comparability between victim surveys. However, its application is highly dependent upon the individual variables and characteristics of each survey.

843. In the following sections, specific issues that affect comparability are discussed. As a result of their inter-related nature, the problems of methodological design and population differences are dealt with jointly under each issue.

(i) *Data collection methods*

844. As set out in Section [ ] crime victim surveys may use a number of survey modes, including face to face interviews, telephone interviews, self-administered interviews, internet based questionnaires and mixed mode interviewing. Survey modes may also use a variety of methods for actually recording information supplied by respondents, including paper and pencil, computer assisted personal (CAPI) or computer assisted telephone (CATI) interviewing. Postal surveys and computer assisted self administered surveys are also growing in use. The use of different survey modes and information recording methods raises the issue whether this may compromise the comparability of results.

845. There are many theoretical reasons why different survey modes may do so. Conversation and prompting during a face to face interview, for example, may lead to better overall recall of events than during completion of a self-administered survey. Survey non-response rates – and resultant bias this might introduce – may also differ between survey modes. Persons may be harder to contact if face to face interviews are the only survey mode than if telephone interviews using landlines and mobiles are also used.

846. The question as to whether such effects exist in practice has been considered by a number of researchers. Results are rather inconclusive. On the one hand, researchers conclude that responses to questions on victimisation from telephone interviews are similar to those obtained face to face<sup>1</sup>. On the other hand, split sample tests with a national crime victim survey in the United States have demonstrated higher victimisation rates in CATI-based interviews than in either face to face or telephone interviews<sup>2</sup>. A study in Germany showed that face-to-face interviews recorded less incidents of victimisation as opposed to written (postal) surveys.<sup>3</sup> Reports also document rather unpredictable effects. When a switch was made from face to face interviewing to CATI in the international crime victim survey carried out in Spain and Northern Ireland in 2005, the results showed a substantial decrease of victimisation rates in Spain and an equally substantial increase in Northern Ireland. There is little way of knowing whether and to what extent the new interview mode has affected these changes<sup>4</sup>.

847. What is well known, however, is that survey mode effects are likely to be greater with respect to sensitive crime areas such as domestic violence and sexual assault. Higher victimisation rates for rape, for example, have been reported where the survey mode is self-completion as opposed to face to face interviewing, due to a reluctance to share information of a personal nature with the surveyor.<sup>5</sup> It is important to be aware that the choice of survey mode may affect comparability with other surveys in respect of such questions in particular.

848. As this Manual discusses in Section [ ], the choice of survey mode and information recording for any particular victim survey is dependent upon a number of

---

<sup>1</sup> Van Dijk, Mayhew, 1992; Lynch 2005; Catalano 2007.

<sup>2</sup> Lynch, 2006.

<sup>3</sup> Kury, 1993.

<sup>4</sup> Van Dijk, Van Kesteren, Smit; 2008.

<sup>5</sup> Johnson, Ollus, Nevala, 2008

factors, including cost, availability of resources, the goals of the survey and the possible sources of survey error. For the purposes of the future comparability of the survey with other crime victim surveys, possibly the most important factor is that the standard of data collection work is the same, whichever mode is used. Two surveys both using face to face interviews but where one has little or no training of interviewers and the other uses well-trained interviewers are likely to be significantly less comparable than two well organised surveys, one using face to face interviewing and the other telephone interviewing.

(ii) *Wording*

849. The basic principle behind standardised crime victim surveys is that they ask respondents about incidents that by and large accord with legal definitions of common offences, but using colloquial language. Burglary, for example, is captured with a question such as: “*Over [x year(s)] did anyone actually get into your home/residence without permission and steal or try to steal something?*”. The exact nature of the event to be captured is further defined by inclusion of a clarification such as: “*I am not including here thefts from garages, sheds or lock-ups*”. The interviewer may also receive additional instructions concerning the counting of events or how to deal with attempted crime. In the case of burglary, for example: “*Include cellars/basements. Don’t count burglaries in second houses. Exclude unsuccessful attempts, ie. Damage to locks, doors or windows*”.

850. Such careful wording and tight definition of the event to be captured may be considered the backbone of comparability between crime victim surveys. The question to be asked however, is whether comparability is greatest where crime victim surveys use identical wording for questions, or where wording is tailored slightly according to different national and cultural contexts.

851. Certainly, the starting point for comparability must be identical wording. Minor differences, for example, in the handling of attempted burglaries previously proved to make comparisons between American and British burglary rates derived from national crime surveys virtually impossible.<sup>6</sup> It is crucial for comparability of results from crime victim surveys that questions on common household crimes, at a minimum:

1. Use *recall time period(s)* that are either the same or equivalent.
  - For example, last calendar year or the 12 months immediately prior to the date of the survey.
2. Contain the same *core* elements to the offences described.
  - For example, for assault: (i) ‘personally attacked or threatened’, (ii) ‘in a way that really frightened you’, (iii) ‘either at home or elsewhere’.
3. Manage *counting of incidents* in a consistent manner.
  - For example, if a respondent was a victim of a particular crime more than once in the recall period, the survey wording should ensure that the number of events is accurately recorded but that details of the crime are obtained with reference to the *last time* that it happened.

---

<sup>6</sup> Van Dijk, Van Kesteren, Smit; 2008

4. Manage *attempts* to commit the offence in a consistent manner.
  - For example, many crime victim survey questionnaires count burglary and attempted burglary as separate crimes but deal with robbery and attempted robbery as one crime.

852. It may be the case that respondents in different countries have different cultural thresholds for defining certain behaviours as crime. Careful control of the four question wording elements above, however, should greatly increase cross-national comparability through prompting recall of the same, tightly defined, event wherever it has occurred.

853. In light of this approach, variation of question wording in a crime victim survey designed to be cross-nationally comparable, should probably be seen as the exception rather than the rule. As always, however, a number of exceptions may still exist. These are likely to relate to questions that seek further detail on the core crimes, rather than the basic screener questions themselves. Terms used to describe the perpetrator of a crime, such as 'spouse/partner', 'intimate partner', 'boyfriend/girlfriend', 'relative', or 'family member' may contain slight variation depending upon the most commonly accepted form in a particular cultural context. Questions concerning corruption may also select particular officials, whether customs officers, police officers, judges, building inspectors, court officials, or magistrates, depending upon whether the victim survey wishes to find out about particular groups of public officials.

854. Where survey questionnaires need to be translated, this must be done particularly carefully in order to preserve the meaning of terms in local languages. As set out in Section [ ] of the Manual, qualified professional translators should be engaged in order to ensure that translated versions are as accurate as possible. An independent back-translation into the original language is also considered necessary to ensure that the meaning of concepts is unambiguous.

855. In summary, crime victim survey questionnaire wording should follow standard event definitions that include standardised recall time periods, core offence elements, and equivalent counting of incidents and treatment of attempts.

*(iii) Sequence of questions*

856. As set out in Section [ ] of the Manual, a major source of error in crime victim surveys is the problem of non-recall and telescoping – the tendency of respondents to place an event within the recall period when it in fact occurred before the recall period (forward telescoping) or to mentally place an event that did occur within the recall period at an earlier point in time (backward telescoping). These effects are shown in the Figure ...<sup>7</sup>:

---

<sup>7</sup> Table derived from Schneider, A.L., *Methodological problems in victim surveys and their implications for research in victimology*. Journal of Criminal Law and Criminology, Vol 72, No. 2, 1981.

**Figure**

		True Categorization	
		Victim in recall period	Not a victim in recall period
Survey Estimate	Victim in recall period	Correct recall	Forward telescoping Exaggeration or lying
	Not a victim in recall period	Non-recall, lying, underestimation of situation	Correct recall
		Backward telescoping	

857. The sequence of questions in a crime victim survey plays a key role in minimising recall error as it is crucial in assisting respondents to *contextualise* their crime experiences. Some surveys begin with screener questions that ask about victimisation during the past *five* years. Where victimisation is reported, follow-up questions typically ask whether the event occurred during the last *calendar year* (or 12 months immediately prior to the survey), followed by detailed questions about the last, or most recent, event.

858. In view of obtaining comparable results, it may be desirable to focus on the one year prevalence rates as the main result reported from crime victim surveys. Adhering to this sequence of questions is key to minimising recall error, particularly in the form of backward or forward telescoping. The five year reference period first assists the respondent in contextualising the one year period. A 1992 crime victim survey in Japan, for example, omitted the five-year screening question. When comparing results with the same survey carried out in 1989 with the screening question it was observed that one-year victimization rates had trebled.<sup>8</sup> This was most probably due to the telescoping effect.

859. Completion of all screener questions first, before moving on to detailed questions, is also important in preventing respondents with many incidents of victimisation from avoiding positive responses to an initial question about victimisation in order to prevent further questioning about details. Indeed, more generally, the interviewer must build rapport with the respondent prior to asking questions that may be particularly sensitive or personal. Perhaps the most widely known error in victimisation surveys is the relationship between failure to recall incidents of assault and the relationship of the victim to the offender. Whilst it is known that approximately half of assaults involve a known offender, victimization which involved family members, persons who knew each other, or juveniles is not as likely to be reported as where the perpetrator is a stranger.<sup>9</sup> Sensitivity in this respect may vary from country to country. As a result, cross-national comparison of crimes such as assault perpetrated by a spouse/partner or family member is especially difficult. If such results are desired, a very detailed and possibly standardised question sequence must be rigorously adhered to.

860. Question sequence is one factor that may affect comparability. Crime victim surveys almost always begin with the least emotive crimes, such as auto theft and

<sup>8</sup> Oda, 1993.

<sup>9</sup> Schneider, 1981.

household crimes before moving to personal crimes, such as assault and sexual assault. This gives the interviewer time to gain a measure of the respondent's trust, giving the respondent time to become accustomed to speaking about unpleasant experiences. It can be expected that surveys asking sensitive question at the beginning may capture less incidents than those in which the interviewer has the time to gain the confidence of the respondent.

(iv) *Sample composition and size*

861. As set out in Section [ ], sample design consists of a number of elements: the population used to represent the target population (the sample frame), the number of respondents chosen, and the method of selection of respondents. When it comes to cross-national comparability, each of these elements plays an important role.

(v) *Sample frame*

862. In general, two samples, each from a different country, may be considered 'comparable' if they are *reasonably equivalent in the extent to which they represent the target population*. Where this is the case, any differences in results between the two surveys can confidently be attributed to 'real' differences between the target populations as opposed to arising from bias introduced by one sample or the other.

863. One major difficulty in cross-national comparisons of crime victim survey data is that results are frequently presented simply as quantitative data, with little or no information about the extent to which the sample frame used is representative of the country population as a whole.

864. As discussed earlier in Section [ ], probability sample frames are typically chosen through use of a proxy for the target population. This may be the population using landline telephones (sampled through random digit dialling or random selection from the telephone book), the registered population (sampled from a population register, electoral register, birth register or medical register), or the population residing at a known address (sampled from a postal address file). A further (multi-stage) technique is the use of a list of known population conurbations, from which a sample is drawn and 'random walk' instructions from a defined starting point (such as the centre of the conurbation) are used to select households for interview.

865. Whichever method is chosen, the sample frame – and hence the sample drawn from it – is unlikely to be representative of the country population as a whole. When comparing across countries, the margins of error become greater, also taking into account that limits of the sampling method may affect different countries in different ways.

866. It can be seen that the relationship between survey *sample frame*, *representativeness* and comparability is not necessarily straight forward. Figure [ ] shows a number of possible such situations:



**Figure [ ] Sample design and comparability challenges**

Country characteristics	Target Population	Comparability
Significant population of hard to reach groups (such as IDPs, ethnic minorities)	Defined to include hard to reach groups	Legitimate comparability <i>vis-à-vis</i> other country populations but potential comparability difficulties due to greater non-response or non-random sampling
	Defined to exclude hard to reach groups	<i>Prima facie</i> ‘methodological’ comparability with surveys in which the sample frame is similarly defined, but low comparability <i>vis-à-vis</i> other country populations
Comparatively small population of hard to reach groups	Defined to exclude hard to reach groups	<i>Prima facie</i> ‘methodological’ comparability and reasonable comparability to other countries with similar characteristics

867. As noted above, the trap to be avoided is the presentation of surveys as representative comparisons of country situations where the sample frame does not represent the whole country population. At the very least, it must be made clear exactly which sample frames are being compared. The use of a postal address file in a developing country as a proxy for the entire country population for example, would exclude significant internally displaced and rural populations. The resultant survey may be ‘methodologically’ comparable with an identical survey carried out in a second country, but the serious limitations in comparing results at the country level must be made clear.

(vi) *Sampling method*

868. Having considered the issues of sample design and comparability in a broad sense, it is also instructive to briefly consider the comparability of different specific sampling approaches. As noted at the beginning of this section, two samples may be considered ‘comparable’ if they are reasonably equivalent in the extent to which they represent the target population.

869. In respect of samples taken using random digit dialling of telephones or random selection from a telephone book, the primary comparability issue derives from telephone penetration. In particular, the emerging trend among specific population groups to exclusively use mobile phones – notably young people – must be taken into account. In such countries (of which Finland is a good example<sup>10</sup>), the use of fixed telephone lines as a sampling frame introduces a serious problem of under coverage. Where young people are underrepresented in an original, fixed line, sample, it may be possible to draw an additional sample of persons owning exclusively mobile phones. The two samples then require combining and re-weighting for age, gender,

<sup>10</sup> source

geographical area and mobile-only ratio in order to restore comparability with (for instance) a country where mobile-only ownership is not a problem.

870. It is important in any registration-, telephone- or address file-based method to understand the coverage of the sample frame with respect to the target population. In principle, a telephone-based sample frame could be comparable with a registration-based sample frame if both represent equivalent coverage of the target population: ie. if the population of persons appearing on the civil register can be expected to be broadly equivalent to the population of persons owning a landline telephone - *vis-à-vis* experience of and attitude towards crime. Indeed, as discussed earlier, evidence from a comparison of CAPI and CATI surveys suggests that samples drawn in these two ways can be compared cross-nationally with some success, at least in developed countries. Comparison of samples from developing countries may be more problematic. Greater numbers of persons are likely to be missing from civil registration lists. Postal address files are also more likely to be incomplete, particularly in rural areas or unregulated municipal suburbs with informal, self-constructed housing. In order for surveys in developing countries to be comparable with other countries, it is likely that a method such as random-walk from a defined starting point must be employed to select household respondents. Such a method stands a better chance of producing a representative sample of the target population than reliance upon existing administrative lists.

871. Sampling designs may also make use of clustering and stratification. The aim of cluster sampling is usually to make data collection easier through the selection of a reduced number of data collection points ('clusters') where the population is expected to be relatively homogenous. The aim of stratified sampling is to increase the precision of estimates by dividing the sample frame into categories that are expected to show some similarity (such as urban/rural). In so far as stratified sampling aims to improve the overall precision of the sample, it will usually have no effect *per se* on the comparability of the survey. For example, an un-stratified random telephone survey in one country could be compared with a stratified random telephone survey in another country (given the same degree of telephone penetration in each). The confidence limits attached to results from each may differ. However, this only affects survey comparability in so far as such limits must be taken into account when considering the results side by side.

(vii) *Sample size*

872. As set out in Section [ ], crime victim surveys require a certain sample size in order to obtain an adequate level of accuracy. When it comes to comparability, the sample size is important primarily insofar as each survey compared must have used a sample of sufficient size to achieve statistical relevance (at least in respect of probability samples).

873. An important point that is frequently forgotten, however, is that the rule of 'statistical relevance' must remain true in respect of *all* variables that are to be compared. Two surveys may be able to generate statistically relevant figures for overall victimization. However, experience of a particular low-prevalence crime may not have achieved statistical significance. In particular, when it comes to specific

victim groups – such as women, or ethnic minorities – it is possible that both sample designs do not have the necessary statistical power to produce significant results *for these groups*. Where this is the case, comparison of results becomes meaningless.

874. An important further point is that sample size can have a significant effect on confidence limits. In general, the larger the sample, the greater the accuracy of the estimate and the smaller the confidence limits. Confidence intervals are an important factor to be taken into account in the interpretation of comparisons. Overlapping confidence intervals for example, indicates that a difference in estimates between two countries is not statistically significant.

875. In summary, sample design is crucial to the cross-national comparability of crime victim surveys. Broadly speaking, two samples may be considered ‘comparable’ if they are reasonably equivalent in the extent to which they represent their target populations. Where a sample is not representative of the whole country population, this must be made very clear in any attempt to carry out cross-national comparisons.

*(viii) Non-response*

876. Closely related to the issue of sample design is the problem of survey non-response. Survey non-response is a problem due to the increased potential for bias that it introduces. Low response rates raise the issue as to how far respondents who are successfully interviewed differ from those who refuse to co-operate, or who cannot be reached.

877. The issue is not straightforward. Although the possibility that low response rates introduce bias in victimisation counts is real, the effect could operate in either of two directions. Where low response is due to high rates of non-contact, people may be omitted who are more liable to victimisation because they are residentially more unstable or simply away from home more. Victims could therefore be under-represented in the sample, with the effect that victimisation rates in countries where non-contact is high are underestimated. Non-contact may be a particular problem in developing countries where socially marginalised groups, especially those residing in informal housing may be difficult to contact for face to face interviews. This factor may contribute to underestimation of victimisation rates in developing countries.<sup>11</sup>

878. On the other hand, the concern is often raised that refusals to cooperate give rise to the selection of respondents who ‘have more to say’, with the result that victimisation rates are said to be overestimated in countries with low response rates due to non-cooperation. At least one study has shown that there is no statistical relationship in developed countries between the number of attempts needed to reach a respondent and overall victimisation rates.<sup>12</sup> This suggests – at least in so far as initial refusal may be taken as a proxy for eventual refusal – that this effect may not have a serious impact on cross-national comparability.

---

<sup>11</sup> Kury, Oberfell-Fuchs, Wurger, 2001.

<sup>12</sup> Gallup, 2005.

879. Non-response remains however a tricky problem for comparing victim surveys results. By definition, there is almost no way of telling whether non-respondents have different experiences of and attitudes towards crime than respondents, potentially introducing bias into the sample. Where samples in countries suffer from significant non-response (particularly non-contact), the possibility of bias is increased. As a result, the confidence with which cross-national comparisons can be made is significantly reduced.

880. Where surveys are desired to be internationally comparable, it is important, as with survey mode, that the standard of data collection work is as similar (and high) as possible. This includes with respect to efforts made to reduce non-contact. Enhancement of response rates may be achieved with thorough preparation, careful selection of interviewers, and consistent use of contact forms and call-backs.

*(ix) Timing of interview*

881. The time of year in which crime victim surveys are undertaken may, have an effect on comparability due to both its impact on the availability of respondents and the tendency to telescope events during recall. Interviewing during the summer months for example, may lead to increased non-response due to persons being away from home, either on holiday, travelling or staying with relatives. Any increase in the level of non-response will have implications with respect to potential bias in the survey estimates. When comparing across countries, timing of the interview should also be taken into account. For example, two surveys carried out at the same time in different hemispheres may not be comparable because of the different season in which they were conducted in the two countries.

882. With respect to recall and telescoping, interviewing in the latter half of a year, when asking respondents to recall events in the previous *calendar* year, may lead to greater memory decay and more forward time-telescoping. Where surveys that have been carried out at different times in the year are compared, this may lead to reduced comparability of one year victimisation rates, although researchers report, at least in respect of developed countries, that there is no concrete evidence of major distortions due to memory decay and/or forward time telescoping.<sup>13</sup> Comparability may be higher if surveys ask about incidents which occurred in the 12 months preceding the survey. The possibility of bias may nonetheless be minimised by instructing interviewers to assist respondents in accurately placing the time of events, such as with reference to significant times of year, including festivals, holidays or relevant personal life events.

*(x) Content*

883. Crime incidents included in victim surveys are generally understood in most cultures. Wording of questions reflects definitions of behaviours related to common offences, but using colloquial language. A core group of offences can be found in most surveys, typically related to theft and assault. Various definitions may apply (see

---

<sup>13</sup> Van Dijk, Van Kesteren, Smit; 2008.

Ch. III.I for a core set of issues). A number of key issues related to non-crime content may also be identified.

884. If comparability is desirable, the wording used should, subject to limited exceptions where necessary for cultural or social reasons, be formed in very simple language that does not allow for ambiguity in different contexts and/or languages.

885. This is not to say that all crime victim survey questionnaires should be identical. It is possible to consider a “module” containing the issues for which comparability is sought, while additional questions may easily be added according to local needs by way of additional modules.

886. It may be desirable for a national victim survey to include all key content elements, in view of increasing its cross-national comparability. The essence of comparability of victim surveys is that standard indicators may be calculated and presented. The content of the questionnaire is crucial to ensuring that sufficient data for such indicators is collected and every effort should be made to ensure that crime victim surveys contain these standard minimum question elements.

### **Summary – Steps towards enhancing comparability**

887. As the preceding discussion has demonstrated, achieving full cross-national comparability between crime victim surveys is no easy task. Crime victim surveys are usually conceived locally, with national needs often (and somewhat understandably) prioritised over international comparability. Moreover, as the state of the art develops and survey questions evolve, the standard which crime victim surveys must achieve for comparability amongst each other can sometimes seem like a constantly shifting goal.

888. Making survey design adjustments for cross-national comparability should not prevent national goals for a victim survey from being met. Rather, elements of cross-national comparability should fit with, and indeed enhance, the effectiveness of a national crime victim survey.

889. It is possible to identify a number of key areas in which a relatively small investment has the potential to produce significant returns, both in terms of cross-national comparability and the technical standard of the victim survey in its own right.

890. Finally, it should be noted that comparability may also be obtained if presentation of the results provides adequate information on the context in which data were collected, including all details and metadata which may facilitate understanding by the reader.

### **References**

Catalano 2007. Methodological change in the NCVS and the effect on convergence, In J. Lynch and L. Addington (Eds), *Understanding crime statistics*, Cambridge: Cambridge University Press 2007.

Gallup, 2005. *Do lower response rates create response (estimation) bias? Effect of repeated recalls on crime experience estimations.* Powerpoint presented at the EU ICS workshop in November 2005 at Gallup Europe headquarters.

Kury, 1993. *The influence of the type of data collection method on the results of victim surveys: A German research project.* In *Understanding Crime: Experiences of crime and crime control*, edited by A. Alvazzi del Frate, U. Zvekic, and J.J.M. Van Dijk, Rome: UNICRI.

Kury, Obergfell-Fuchs, Wurger, 2001. *Methodological problems in victim surveys: The example of the ICVS.* *International Journal of Comparative Criminology* 2(1): 38-56

Lynch 2005;

Lynch, 2006.

Oda, 1993. *Japan* In *Understanding Crime: Experiences of crime and crime control*, edited by A. Alvazzi del Frate, U. Zvekic, and J.J.M. Van Dijk, Rome: UNICRI.

Schneider, A.L., *Methodological problems in victim surveys and their implications for research in victimology.* *Journal of Criminal Law and Criminology*, Vol 72, No. 2, 1981.

Van Dijk, Mayhew, 1992; *Criminal victimisation in the industrialised world. Key findings of the 1989 and 1992 International Crime Surveys.* The Hague: Ministry of Justice, the Netherlands.

Van Dijk, Van Kesteren, Smit; 2008. *Criminal Victimisation in International Perspective. Key findings from the 2004-2005 ICVS and EU ICS.* WODC: Netherlands.

Johnson, Ollus, Nevala, 2008. *Violence against Women: An International Perspective.* Springer: New York.

## Table of contents

<b>Part IV. Comparability of data across countries</b> .....	229
What is comparability and why is it desirable?.....	229
Comparability challenges of crime victim surveys .....	230
Factors that affect comparability .....	231
(i) <i>Data collection methods</i> .....	232
(ii) <i>Wording</i> .....	233
(iii) <i>Sequence of questions</i> .....	234
Figure .....	235
(iv) <i>Sample composition and size</i> .....	236
(v) <i>Sample frame</i> .....	236
Figure [ ] <i>Sample design and comparability challenges</i> .....	237
(vi) <i>Sampling method</i> .....	237
(vii) <i>Sample size</i> .....	238
(viii) <i>Non-response</i> .....	239
(ix) <i>Timing of interview</i> .....	240
(x) <i>Content</i> .....	240
Summary – Steps towards enhancing comparability.....	241
References.....	241