Ch. III.N. Data processing

697. Data processing includes the transfer of data included in the questionnaire to data files ready for analysis, coding processes, data checking and data editing. The nature of data processing will depend on the way in which the data has been collected. Paper questionnaires will require considerably more data processing after the interview while computer assisted interviews include much of the checking process during the interview. This section reviews some of the key components of data processing including transferring data to files, security concerns, data checking and data editing.

Transfer to data files and security issues

698. Once an interview has been completed, secure arrangements are needed for the return of the interviews to the central office. Names and addresses of respondents should be returned separately to ensure confidentiality. There will also need to be a comprehensive checking-in system to ensure all the issued sample is accounted for.

699. Paper interviews require data be keyed or scanned into a suitable data entry program or directly into one of the main survey analysis programs. Commonly used survey analysis programs used for this purpose include the SPSS, STATA, and SAS (there are also freeware alternatives such as EpiData). Data entry should be checked either through double entry processes or using facilities within the programs. Some provision will need to be made for sampled cases which did not produce an interview so that response figures can be calculated.

700. Initial processes for CAPI and CATI (computer assisted telephone interviews) differ. Data collected from a telephone interview will already be held centrally (although they may require an internal transfer to a main data file). In contrast, data from a computer assisted personal interview will need to be transferred from the interviewer's laptop to a centrally held data file. Completed interviews using CAPI can be returned electronically by modem directly to the central office through an intranet arrangement or on some other media such as a computer disk. Security measures are essential for any type of electronic transfer. Unauthorized access to personal data is prevented through user name and password protection systems. In addition, all laptops and central computers should run anti-virus software.

701. All personal details of the respondent (e.g., name, address, telephone number) should be collected and stored in files that are encrypted and transmitted separately from the main data to provide additional security.

702. For both paper and computer assisted interviews a back-up copy should be created during each stage of the process and maintained separately from the main data file. This includes files on the interviewers' laptops. Archived data files should be created and stored separately from the main CAPI servers or the original paper interviews.

703. All office-based staff dealing with the survey should be required to abide by information security policies and standards. In addition, there should be contingency plans for major disasters.

Data checking

704. After the data are received, the next major step is data checking. Data checking includes range, logic and consistency checks. For paper interviews these can be done by hand or they can be conducted once the data has been entered into suitable software. The use of CAI (computer assisted interviewing) removes much of the requirement for post-fieldwork edit checks. As far as possible this type of checking can be built into the CAI program. This approach is methodologically preferable to post- fieldwork editing, since the interviewer can resolve any inconsistencies during the interview. However, because some checks require reference to external data sources, and issues of timing and interview flow means that there is a limit to the number of in-program checks, some in-office edit checks conducted in the central office are desirable.

705. Some examples of checks that can be conducted at various stages of the process are as follows:

- Check every question in the analysis data file against frequency counts that are produced directly from the CAI or initial data entry database. This is to ensure that the numbers on the analysis file match the numbers in the actual database and that no corruption or error has occurred during the creation of the analysis file.
- Check every question in the analysis data file to ensure that the base number is consistent with what you expect based on the filter for that question. Questions should also be checked to ensure that the "don't know" and the "refused" codes are correct and that the labelling is correct.
- Where there is more than one file per case (for example where there are separate files for each victimization experienced by the respondent) check the consistency between the files to ensure they contain records from the same cases.
- If the survey is repeated, compare the distributions of key variables to previous iterations. If any large changes in distributions are noted, a more detailed check to ascertain why these differences exist is necessary.
- In addition to standard consistency checks, additional checks relating to offense coding are advised. For example, if during a victimization, data reveals an arson but no fire recorded further checks are required. Similarly, if during a single incident, the data show a car theft but the data also show that nothing was stolen, additional checking is needed. Or if during a single victimization, a crime was noted as occurring at the respondent's work place, yet the respondent also is categorized as "not working" requires additional attention. Clearly this type of consistency checking is limited by resources. Even with extensive checking and vast resources, it is likely the case that the data set will never be fully consistent
- If the data has been weighted then checks will be required to ensure the weights have been computed correctly.

706. The edit checks discussed above also form a key part of ensuring there has been no loss of data gathered at interview.

Data editing

707. Data editing is concerned with the identification and, if possible, the correction of erroneous or highly suspect survey data.

708. Errors can arise in the data for a wide variety of reasons, such as:

Coverage errors (arising from the omission or duplication of responses)
- duplicate records
- loss of records
- incomplete canvassing

Content errors (arising from incorrect reporting or recording)
- errors in questionnaire design such as poorly worded questions
- misunderstanding of questions by respondents
- mis-keying during data entry
- out-of-range or invalid responses (i.e. age given as 130 years)
- incorrect units (i.e. response given in days rather than weeks)
- incorrect scale (numerical items expressed as thousands instead of millions)
- inconsistent data items (i.e. age of child provided when number of children given as zero)
- failing to follow correct skip patterns
- incorrect balances (i.e. sum of parts on numerical item does not equal total)

709. The effective and appropriate use of data editing procedures is designed to minimize these errors and so enhance the quality and accuracy of the survey data. Furthermore, consistent errors in the survey data might be identified and these could suggest problems with questionnaire design or interviewer technique, for example. Identifying the reasons behind these errors will improve future survey waves.

710. The process of data editing occurs during and after the data collection phase, but usually prior to any data imputation or analysis. The activities associated with the data editing process include the following:

Micro-editing
- concerning the validity and consistency of individual data fields, fields within a record, and the relationships between individual records (i.e. are data items within acceptable bounds, ratio checks, comparison of data to previous survey waves).

Macro-editing
- concerns aggregated data (either whole survey or a sub-set of respondents) and assesses the importance of a suspected error.

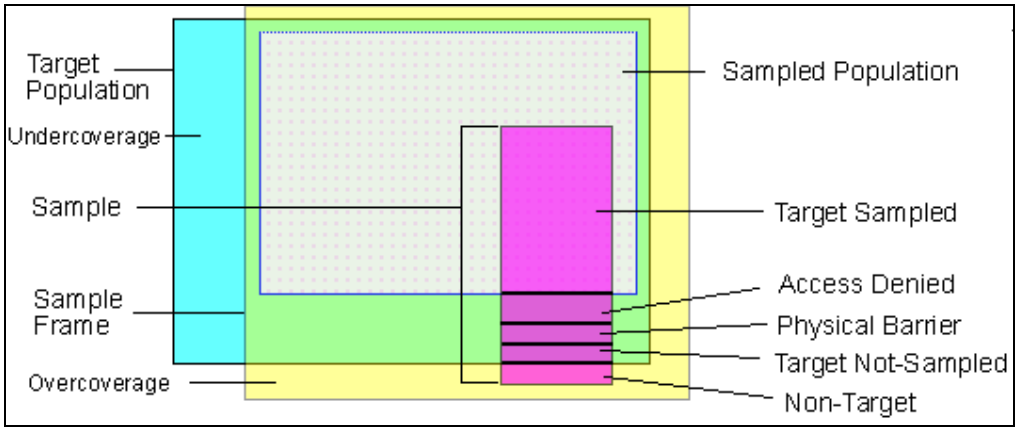
711. Data editing procedures range from the simple, such as checking for out-of-range responses, to the more complex, such as identifying inconsistencies in responses for particular groups of respondents. Although there are many generic elements to the data editing procedures that can be followed, there are also specific rules that need to be developed and implemented that relate specifically to the victimizations survey.

Ch. III.O. Data analysis; Estimation and presentation

Estimation and data weighting

712. The figure below taken from the U.S. Environmental Protection Agency, estimates based on Target Sampled units apply to the Sampled Population with no additional assumptions, but they can apply to the portion of Target Population within the Sample Frame only if it is possible to assume that the access denied, target not-sampled, etc., units occurred randomly and independently of units characteristics, especially those of interest. Furthermore estimates for the Target Population require assumptions above and also that portions of Target Population that are not included in the Sample Frame have same characteristics as the Sampled Population.

Figure: Target Population - Frame - Sample – Estimates: U.S. Environmental Protection Agency



713. Estimates from victimization telephone surveys are both totals (n° of victims), rates and proportions and enumeration of events (n° of experienced crimes). Considering a complex sample design as in the victimization survey, estimates are based on calibration estimators, which take into account the weight of each unit, considering also the stratification, in order to reproduce some significant total for the target population.

714. Weights, initially defined as the inverse of the inclusion probability of each unit, are corrected by a factor calculated ex post facto to take into account non response and list problems, and the main demographic and social characteristics of the target population as age and sex, level of scholarship, marital status, type and size of households, according to rates from the past census updated by offices of vital statistics or from other reliable face to face sample surveys. That is because telephone surveys are affected by bias due to non responses and under coverage of phone lists related to the above social and demographic population characteristics. Indeed people with higher level of scholarship are over covered, while divorced women, household s

of one person, and younger individuals are under covered, as those more likely own only a mobile phone.

715. The correction factor to multiply to the initial weight is then obtained by solving a problem of constrained minimum, where the function to minimize is the distance between the initial and the final weights, and constraints are the equalness of totals by the above population characteristics.

716. Telephone samples bias is then partially corrected by weighting as indicated.

717. Considering a total ${}_dY$ referred to the territorial area d, the final estimate will be:

${}_d\hat{Y} = \sum_{h=1}^{H_d} \hat{Y}_h = \sum_{h=1}^{H_d} \sum_{j=1}^{m_h} Y_{hj.} \cdot W_{hj.}$ where $Y_{hj.}$ and $W_{hj.}$ are respectively the value of Y and the weight for the individual sampled in household j of stratum h.

718. The better option would be rely on census data to correct weights, however, reliable estimates from other sample surveys, such as other face to face surveys, can be useful.

2.Data files

Determining whether and what kind of files will be made available for secondary analysis/releasing public use microdata files.

3. Tabular presentation

Determining how primary analytical variables will be classified.

Ch. III.P. Pilot survey

719. The pilot survey is a useful tool for preparing and verifying the selected survey methodology. A pilot survey is a survey that reproduces all survey phases on a small number of survey units or respondents. In general, about 800 respondents are utilized during a pilot study. It is not important that sample units, or respondents, are representative of the overall population of interest. Rather, it is important that they represent the various ways respondents will be contacted so they can reveal any difficulties that may emerge during the actual survey. For example, the pilot study should include segments of the population that could are anticipated to be problematic in contacting or interviewing (such as minority groups).

720. A pilot survey is far more complex than a pre-test. A pre-test focuses more on the questionnaire alone. In contrast, the pilot study deals with the *entire* study from the beginning to the end. Respondents are selected and interviews are administered. From this, aspects of the study are checked. The aspects checked include:

- What were the response rates?
- Was item non-response high?
- Did the interview take as long as anticipated?
- Did the skip patterns work as they should?
- Was the structure of the questionnaire appropriate?
- How many contacts to the respondent were required to actually complete the interview?

721. Aspects related to interviewers are also verified during a pilot study. Interviewer items verified include:

- Did the interviews and respondents develop the appropriate rapport?
- How many days for theoretical and practice training of interviewers were required?
- What did the training involve?
- What materials were required to train the interviewers?
- What did an evaluation of the training reveal?

722. Information flow is also investigated during the pilot study.

- How was address selection accomplished?
- What was involved in address substitution?
- How did the households respond to introduction letters and phone contact?

723. And finally, the pilot study addresses the data itself.

- What problems were encountered when recording the data?
- Were there issues when producing and reading data files?
- Were interviewers able to download electronically data?

724. In order to address these and other important questions raised during the pilot study, the thoughts of all members of the survey team are important to consider. For example, those conducting the interviews have great insight into the working of the

questionnaire and respondents' views of the survey. Clearly, maintaining a good relationship with open communication with interviewers is vital. One method to gather information from interviewers is via a meeting. These face-to-face meetings are useful for brain-storming and debriefing purposes. During these meetings items such as the following can be discussed.

- Did respondents have difficulties in understanding some questions? Did they request clarifications?
- Did respondents demonstrate reluctance in answering some questions? Why? What did you say to the respondents to encourage an answer? What were the results of your prompts?
- Did you find redundant questions that appeared to be useless or irrelevant?

725. When interviewers are located across geographical regions or when it is not possible to schedule a meeting that gathers all interviewers, it is recommended that they have another way to communicate as quickly and effectively as possible (mailings, chat/forum, telephone). Another option is to select some interviewers to participate in a focus group to further delve into particular areas of interest.

726. Supervisors have an important role as well. They offer a broader view of what is happening during the questionnaire and interview administration. In telephone surveys, where interviewers are gathered in a centralized facility, supervisors can listen to interviewers during the interview (according to laws defining the job characteristics) to get a better idea of what is happening during the process.

727. Regardless of the source, all thoughts and comments about the study should be forwarded to the study manager for review. An evaluation of these comments will determine whether additional changes in specific points in the survey process are warranted. In general, small changes made do not require another wave of the pilot survey. However, if changes are significant, it is advisable to conduct a second pilot study. All problems corrected during the pilot study phase minimize future problems during the actual survey.

Ch. III.Q. Analysis and interpretation

728. *Data analysis* is the process of transforming raw data into useable information that is often presented in the form of a published analytical article. The basic steps in the analytic process consist of identifying an issue, asking meaningful relevant questions, answering questions through examination and interpretation of data and communicating the key findings and results to the end-user.

729. Data analysis also has an important role as part of the survey development and revision process. It can have a crucial impact on data quality by helping to identify data quality issues and by influencing future improvements to the survey process, including question wording, questionnaire flow, approach and methodology. Analysis is essential for understanding results from previous surveys and pilot studies, for planning new statistical activities, for providing information on data gaps, for designing surveys, and for formulating quality objectives.

Principles

730. An organization undertaking the development of a victim survey should be aware of the relevance and usefulness of the information contained in the data. Analysis is the principal tool for obtaining information from the data. Analysis results may be categorized into two general types: (a) descriptive results, which are results relating to the survey population at the time that the data were collected - for example, the rate of the population reporting physical assaults for the year that the population was surveyed; and (b) analytical results relating to a survey population that often goes beyond the actual population surveyed – for example, the risk of someone being the victim of a sexual assault.

731. To be effective, the analyst needs to know the audience and the issues of concern (both current and those likely to emerge in the future) when identifying victim topics and suitable ways to present results. Study of background information allows the analyst to choose appropriate data sources and statistical methods to influence the analysis. Any conclusions presented in an analytical study, including those that can impact on public policy, must be supported by the data being analyzed.

Ch. III.R. Language and cultural considerations

Languages: Will the questionnaire be administered in more than one language?

732. As borders become increasingly open, countries around the world are becoming more diverse. In designing a victim survey questionnaire, one has to take into account the diversity of the population if the results are to be representative of the population. One factor that is important to consider is the linguistic profile of those who will be included in the sample. Ensuring that the questions are relevant, clear and easy to understand for all individuals in the population, translating involves a number of challenges because not all languages have words to describe the same concepts.

733. The first challenge is deciding which languages the survey will be administered in to be representative of the population. This can be determined by consulting censuses of the population or other research which details which languages are most often spoken at home by the survey population. Efforts should be made to include the languages most frequently spoken at home.

734. The second challenge is translating the survey into the other languages so that the meaning is maintained and the entire sample is being asked the same questions. The degree of difficulty of this challenge varies depending on the language. Some languages do not have words or phrases equivalent to other languages. Consistency in question wording is critical to ensure comparability across populations. To facilitate this task, expert groups should be formed to assist in question wording and to identify other issues that may arise given the ethnic diversity of the population. Translation of the questionnaire should not be viewed as a one-way process, as it the process may also involve making adjustments to the original language version in order to find the terms with best equivalence in the different languages.

735. Following content development qualitative studies involving focus groups and face-to-face interviews with feedback should be conducted. If feasible, these studies should be carried out multiple times in various languages. Each series of studies will help better understand the way people responded to the questions, identify certain problems and improve the questionnaire. During a large pilot test additional problems that may not have been noticed before can be caught. This test will also confirm whether the non-response rate due to language barriers has been addressed.

736. Another challenge is that not only must the survey be translated properly, but individuals who speak the various languages in which the survey will be administered must be found to conduct the survey. This can pose problems since translators will often have a firm grasp of the other language orally, but not in writing, or vice versa. A final challenge is finding the resources to translate the survey responses back into the official language of the country in order that these results can be analyzed and included in the study.

Cultural issues/differences

737. In addition to problems that can arise through language differences, there are a range of further cultural issues that may impact on the success of a survey and quality of data collected. A range of these other cultural considerations may be relevant depending on the purpose of the survey, populations being surveyed, and the environment in which the survey vehicle will be enumerated. These situations are most likely to arise when the population of an area to be surveyed (be it a city, state, region, canton, province or nation), is comprised of persons from multiple ethnic or cultural backgrounds. A survey's aim may be to specifically understand crime victimization of a particular cultural or ethnic group. Alternatively, the aim may be to gather a representative sample of an area, and the challenge may be obtaining the support of a minority cultural group in order to achieve these goals. As previously mentioned, these challenges can be exacerbated by language differences. As a general rule, the statistical organization, in these instances, will represent a dominant cultural group, and will need to consider the requirements, mores and attitudes of other ethnic and cultural groups. Understanding, being able to define the relevant cultural groups, and identify persons from these groups where necessary, may be essential to fulfilling the survey goals.

738. It is vital that any respondent understand the importance of the data they are providing, and how this data will be used. If a specific cultural group is to be included in the survey, it is important to determine how they may specifically use or access the information that is collected. This is important throughout the survey process, from identifying information needs and who in communities may use the output data, through to ensuring that outputs can be access and understood. Depending on the situation this may simply mean promotion or translation into other languages, through customized outputs and information delivery. The latter may be necessary when interviewing persons living in remote communities who may not have access to technology required to view statistical outputs disseminated through electronic means, or where literacy levels may restrict the utilization of more sophisticated statistical output. The data released from the survey needs to be meaningful and visible to the communities who have contributed towards its collection. Communicating the use of the survey can facilitate both a current survey under development, and be a wise investment for agencies looking to gain the trust and cooperation of cultural groups and communities in future surveys.

739. Consultation with the communities to be surveyed is essential in gaining support for the survey endeavour – particularly if endorsement can be gained from community leaders. Persons with an in-depth understanding of the cultural group of interest will be able to provide expertise in relation to the content of collections. In some cultures, it may not be permissible to discuss certain topics entirely (particularly if sensitive, such as domestic violence or sexual crimes), or perhaps there are restrictions upon who may answer certain questions. In some instances, for example, a head of a household may wish to represent the views of the whole household, or there may be reluctance to let some members of the family speak independently. It may be necessary to design some of these factors into the survey instrument, instructions to interviewers, or seek to otherwise make respondents comfortable in participating in the survey.

740. Other potential methodological issues can also be pre-empted through community consultation and liaison. Access to cultural groups or communities living

in non-mainstream arrangements or environments can be a challenge to representatives of statistical agencies conducting data collection. The establishment of community liaison points of contact can provide invaluable information about the best ways to approach persons from different cultural groups, and assist in access. This can be a particular challenge when conducting interviews or seeking to have questionnaires filled in by respondents in remote geographical locations, or where language and literacy levels may be variable. Employment of persons from different cultural groups on survey staff can be a way to gain the support and trust of different communities – particularly if face-to-face interviews are part of the survey methodology. However, there may also be disadvantages to this if both the respondent and the interviewer are from the same, relatively small group. This may lead the respondents to fear for the loss of anonymity, or pressure them to give answers (e.g. regarding their activities) that are more socially desirable in their culture. Different cultures can have different rules about seating, etiquette, touching, eye contact and communication between men and women. Facilitators or persons from the cultural background of interest may be able to assist representatives of the statistical agency in overcoming these potential cultural barriers.

741. Trust can sometimes be lacking between different cultural groups, and persons from some ethnic backgrounds may have an inherent distrust of governments or bureaucracies based upon previous personal or historical experiences. Depending on the groups involved, this may include histories of war, dispossession, stolen children, police harassment, incarceration, etc. Distrust may be found also in a country without such history, as respondents' attitudes towards officials may be based on their earlier experiences from other countries. It is vital that an agency that may be identified with such historical activities or part of a broader 'Government' be aware of these possibilities, and seeks to gain the trust of the groups involved. As with the issues raised above, this is best done through discussion with community leaders and a careful and considered approach of communication and exposure.

742. Provider load can arise as an issue for a statistical agency conducting surveys that are particularly focused on one smaller cultural or ethnic group (particularly a survey program that extends beyond crime victimization surveys). There are two primary elements to provider load: where a small community is repeatedly sampled for different purposes, and where a survey takes a long period of time for a respondent to complete.

743. In order to gain reliable statistical estimates at lower levels of disaggregation or geography, it may be necessary to sample a high proportion of a population sub-group. Should this be the case, consideration needs to be given to the balance between the needs and benefits of the information, and placing an undue burden upon respondents. Undue respondent burden can be one way of eroding cooperation between a statistical agency and a particular group – particularly if there are many cultural differences, and the agency is in the process of building trust and a strong liaison network.

744. Normally, as the length of an interview increases, respondents tire, and in some cases may begin to give responses that are less accurate or which they feel may end the interview more quickly. If data is collected through interviewers, long survey interviews can also result in 'short-cuts' being taken, especially if they notice the

respondent is becoming fatigued. Additionally, there may be a temptation to conduct long surveys with persons from difficult to sample or reach groups, but this can lead to a loss of cooperation, and persons responding in certain ways purely to end the interview quicker.

745. Disparate cultural groups may respond differently to contrasting methodologies, and some research will need to be conducted to determine whether or not this may be an issue in a particular survey situation. For example, a scripted interview may not be effective with all cultural groups – in some instances other techniques may be more appropriate to elicit data. In some more remote or traditional cultural areas, written questionnaires may not be as successful, and there may be different levels of support for phone or face-to-face interviews.

746. In instances where there are wide differences between cultural groups to be included in the survey, content and questionnaires may also need to be tailored to allow for different understandings of concepts, or ways of thinking about the world. For instance, there may be different concepts of time employed, and asking about events that have occurred in the past 6 months may not have relevance. However, if there is a known significant event within a particular community that occurs around the reference period date, this may be a more effective substitute. For example, "Since Easter, has the following occurred…?"

Table: Difficulties and solutions

Survey difficulties that can arise	**Solutions**
• communication barriers • cultural sensitivity • mistrust of government agencies • respondents are reluctant to participate • conceptual difficulties • high costs	• cultural awareness • consultation with relevant communities/organizations • involvement of people from the cultural group(s) of interest • modifying survey questions • containing respondent load • extensive testing and refining of survey strategies

Reference

Bizier, V., Kaddatz, J., & Laroche, D. (2004). Methodological Challenges in a Survey on the Ethnic and Cultural Diversity of the Canadian Population. *Symposium 2004: Innovative Methods for Surveying Difficult-to-reach Populations*. Catalogue no. 11-522-XIE. Ottawa: Statistics Canada.

Ch. III.S. Ethical considerations

747. Because survey researchers are asking respondents to provide information about themselves, their experiences, and possibly their attitudes, it is imperative that the researcher act responsibly regarding this information. This entails two primary tasks: Protecting the rights of respondents and ensuring the validity of the survey results. This section reviews some of the ethical considerations related to protecting the rights of respondents and ensuring the validity of the survey results.

Protecting respondents

748. Protecting respondents (and potential respondents) requires attention to many details. First, the researcher must develop survey protocols that inform respondents of the survey's purposes and procedures. Second, the research must be conducted in a fashion that respects the respondents' rights to privacy. Third, the research must be sensitive to existing cultural norms for social interaction. Fourth, the research must avoid embarrassing or harming the respondent in any way including physical, psychological or emotional harm. And finally, in the case of voluntary surveys, the respondents must be free to chose to participate or not without fear of repercussion.

Informed consent

749. Informed consent means that respondents are made aware several things. First, the respondents must be informed about the purposes of the research. Second, the procedures used in collecting the data must be revealed. Third, any risks or benefits to the respondents must be conveyed. And finally, the respondents should have full information regarding the survey sponsor. Groves et al. (2004) expand on this idea by identifying eight essential elements of informed consent:

- A description of an explanation and the purpose of the research. This includes how long the respondent will be needed to participate, a description of the procedures including any that are experimental
- A description of anticipated risks or discomfort
- A description of any anticipated benefits
- Disclosure of any alternative procedures (if applicable)
- Disclosure of how the anonymity or confidentiality of the data will be maintained
- A description of compensation or treatment available if injury occurs (if applicable).
- Disclosure of whom to contact with any questions
- Disclosure that participation is voluntary and the respondent may discontinue participation at any time

750. Only after considering this information can the respondent decide whether or not to participate in the survey. Informed consent can be obtained verbally or in writing. Written consent forms generally provide the information about the survey and

the essential elements outlined above and include a place for the respondent to sign to acknowledge that he or she understands the survey purposes and protocols.

751. For a victimization survey, a written consent form may be in order if the survey contains sensitive questions such as those that ask about any criminal acts the respondent may have committed. In this case, potential harm could come to the respondent should the information provided become public. In general a written consent form allows potential respondents to evaluate the possible harms that might be associated with participating and make an informed decision about whether to participate. The form also provides the survey researcher with tangible evidence that that respondents participated willingly and were knowledgeable of the risks when they agreed to participate.

752. Generally, a verbal consent is acceptable when there is minimal risk of harm to the respondent. Nonetheless, surveys that obtain consent verbally should ensure that the respondent understands the potential risks, if any, for participating and also be provided information about the purposes of the survey and for whom it is being conducted. Telephone surveys generally use some form of verbal consent procedure.

753. Informed consent implies that the respondent is legally, mentally, and emotionally capable of providing such consent. The survey researcher must, therefore, create special procedures for obtaining such consent for participants who are underage or are unable to make their own decisions. In these cases, consent must be obtained from a parent or guardian (or the parent or guardian is used as a proxy respondent).

754. There are additional ethical concerns that must be addressed for interviews involving persons residing in institutions to ensure their safety and allow for informed consent. For example, a victimization survey of persons in a home for the elderly might focus on abuse by aides or employees of the home. The researcher must develop protocols to ensure that respondents are protected from potential retribution should they participate, as well as ensuring that pressure from others will not influence the respondents' responses.

Protecting respondent privacy

755. Because surveys seek to obtain information from respondents, the survey researcher has a moral (and possibly legal) obligation to protect the information. An aspect of informed consent is notifying the respondent about how the information will be used, and who will have access to it. Information provided to respondents about any legal guarantees of confidentiality of the data they provide should be accurate.

756. For face-to-face surveys, protecting the respondent's privacy entails ensuring that the interview is conducted away from other household members. This can be especially important for victimization surveys that measure violence between intimate partners[1], violence among family members, or violence by caregivers. Respondents

[1] For a more detailed guidelines for research on intimate partner violence against women, see *Putting Women First: Ethical and Safety Recommendations for Research on Domestic Violence Against Women* (WHO, 2001; http://www.who.int/gender/violence/en/womenfirtseng.pdf)

must be assured that the information they provide will not be shared with anyone, including other members of the respondent's household. To insure respondent privacy and safety, surveys related to violence against women or family violence or similar subjects should include instructions for interviewers to reschedule the interview if a household member is present and arrange for a time and place that will guarantee the respondent is alone when responding to questions about offenses committed by intimate partners or family members. . It is important to remember that such additional protections for respondent privacy are likely to have also budgetary consequences, and they should therefore be taken into account at an early stage of project planning.

757. Interviewer and staff training is a key component for ensuring the privacy of the information provided by survey respondents. Interviewers and others with access to survey data must be trained about their legal and ethical responsibilities to prevent disclosure to anyone who should not have access to the information.

Respecting cultural norms

758. Sound survey practice, as well as common sense, dictates that the researcher avoids any behaviors that will violate the respondents' beliefs or cultural norms. One example of a behavior that might be offensive and therefore avoided is interviewing on a religious Sabbath or holiday. Another example is using inflammatory, derogatory or foul language. And in some cultures, using interviewers of the opposite gender would be offensive.

Avoiding harm to respondents

759. In medical research, the potential for harm to participants can be obvious. Even though surveys such as victimization surveys do not involve physical contact or medical procedures, the potential for harm still exists. This potential for harm can come both during and subsequent to the survey interview.

760. During the interview, potential harm is associated with the possibility that the questions or survey procedures might cause physical, emotional or psychological damage to the respondent. Victimization surveys often contain questions related to sensitive subjects, including rape, sexual assault and intimate partner and family violence. Some victims of sexual assault (or other crimes) may suffer traumatic reactions when reminded of the incidents during the survey. While it is not possible to avoid all such situations, it is imperative that interviewers be trained to address these situations if they arise. One approach can be to provide victims of crime at the end of the interview information about local victim support services.

761. Survey developers should identify potentially sensitive questions and explore the creation of special procedures for asking such questions. Sensitive questions are questions that respondents are uncomfortable answering. These questions can be perceived as intrusive or embarrassing to the respondent. Often, a respondent will refuse to answer these questions or even provide false answers.

762. There are methods for avoiding refusals while asking sensitive questions. One technique for face-to-face interviews is to use self-administered questions for the topics deemed potentially emotionally taxing. During an in-person interview, the interviewer excuses themselves and allows the respondent to self-administer the questions. The British Crime Survey uses such a technique for their violence against women survey. Not having the interviewer present increases the likelihood of responses to sensitive questions.

763. The randomized response technique is another method for achieving responses to sensitive questions. In this technique respondents answer either the sensitive question or a different, innocuous question. Which of the two questions asked is randomly selected. The notion is that respondents are more comfortable answering a sensitive question when the interviewer does not know what question they are answering. The probabilities of the question being the sensitive question or the innocuous question are used to estimate the sensitive question characteristics.

764. The potential for harm to the respondent following the interview generally involves a situation in which someone else learns that the respondent participated in the survey. For example, a victim of violence by an intimate partner could face retribution from the offender if the perpetrator learns of the respondent's participation (such retribution may be triggered already by advance letters sent to the household or any material left to the household after the interview). Harm to the respondent can also come from breaches of information if survey respondents are identified in the data. Breaches in confidentiality can occur in many ways, therefore researchers must have protocols for protecting the information provided by respondents and for removing identifiers from survey responses, securely storing information, encrypting files, and reporting data.

765. Tables and data tapes must be examined to determine whether there is any possibility that they may enable identifying individual respondents. When only a few responses are represented in a survey table, there is a risk that the identity of the respondent could be determined when the survey results are reported. To account for that, the research may suppress the data by blanking some cells that contained only a few cases. To ensure that the data cannot be estimated through subtraction from other data in the table, additional data is often suppressed, too. This method does not change the data totals. Any blocking or blanking the cell data must be noted.

766. Other methods for addressing the confidentiality in tables that report only a few responses include using broadened response ranges and collapsing data over several categories. Although the data are not affected by these methods, information is still lost.

767. For data tapes, protecting the respondent's identity involves removing any information that could be used to identify individuals; scrambling identification codes, and examining outliers for key variables. For survey research, the risk of harm to respondents resides primarily in the possibility of their identity being revealed. Research organizations must strive to ensure confidentiality to protect respondents.

Respecting peoples' rights to not participate in voluntary surveys

768. There is clearly a tension between statistical theory and the concept of voluntary participation in surveys. For survey results to be valid, nonresponse must be low. Yet "voluntary" means that persons selected for the survey can choose whether or not to participate without any adverse impact to them. While survey procedures can be designed to persuade reluctant respondents, interviewers must be trained to respect respondents' right to decline to participate.

Balancing confidentiality and laws on reporting crimes

769. Many nations have laws protecting the confidentiality of information collected for statistical purposes. Many nations also have laws mandating reporting to the police of some serious crimes such as child abuse. Clearly, there can be a conflict between these laws. An interviewer is forbidden by law to reveal anything about a person who participated in a survey, but is also required by law to report an offense learned about during the survey. This creates both an ethical and legal conflict. In many nations this conflict has not been resolved. For example, in Canada, the a Statistics Act protects the confidentiality of information collected for statistical purposes, which means interviewers are not permitted to report an offense to the police.

Ensuring sound professional and scientific process

770. Conducting a survey ethically is not solely a matter of protecting respondents and the information they provide. It also entails adhering to sound professional and scientific methods and established codes of conduct for survey research. In the United States, the American Association for Public Opinion Research (AAPOR), an organization of academic and professional survey researchers, has created a "Code of Professional Ethics and Practices" for its members. The code, available at http://www.aapor.org/default.asp, covers several topics, including the principles of professional practice in conducting survey research, the principles for working with people, including respondents and the public, and the standards for minimal disclosure of essential research information.

771. Sound professional and scientific processes involves using established statistical sampling methods for drawing samples, working to develop unbiased questions and testing them prior to implementation, providing adequate training for interviewers and other staff, and providing users of the data enough information about the survey and its methodology to enable understanding the limitations of the data as well as the strengths.

772. When publishing survey results, the researcher should present information about the sample design, sources of potential bias in the sample (i.e., under coverage of a particular subpopulation,) sources of nonresponse error, response rates, and measurement error associated with survey variables.

773. The findings should be accurate reflections of the data and should not be manipulated to create desirable results. It is often desirable to separate the agency or office that conducted the survey and analyzed the data from agencies related to government policy in order to ensure that the survey results are free from political influence. This separation can also serve to increase public acceptance of survey findings.

774. In the long run conducting surveys ethically by protecting respondents, protecting survey data, and presenting results completely and fairly, serves to promote trust in survey results. Respondents will be more willing to participate in future surveys if they are not abused by survey researchers. The public as well as policy makers will trust survey data if releases are honest and present everything one must know to interpret the results. Taking shortcuts or otherwise failing to adhere to sound professional and scientific practices may seem beneficial or appear to avoid difficult situations, but ultimately, such practices serve to cast doubt on the entire field of survey research.

References

American Association for Public Opinion Research. Code of Professional Ethics and Practices. http://www.aapor.org/default.asp.

Eurostat. European Statistics Code of Practice. 2005.

Groves, Robert M., Fowler, Floyd J., Couper, Mick P., Lepkowski James M., Singer, Eleanor, and Tourangeau, Roger. *Survey Methodology*, John Wiley & Sons. 2004.

Backstrom, Charles H. and Hursh-Cesar, Gerald, *Survey Research*, Macmillan Publishing Company, 1981.

Ch. III.T. Data dissemination and documentation

775. The way in which an agency approaches data dissemination and documentation critically influences the ability of users to access, use and understand the data produced from the victimization survey. In order to publish the findings from the survey in a statistical product several steps are involved:

- developing the product concept and output requirements
- preparing the product
- printing paper-based publications and duplicating electronic output (if necessary)
- release and distribution of the product.

776. These steps outline a dissemination plan, which aims to ensure that the statistics generated by the crime victimization survey are accessible and used. Outputs need to be provided in ways that take into account the variety of information needs expressed by users, and also the capability of users to deal with different forms of statistical information. Statistical output may be presented in tabular fashion as a general rule, but users may require simplified outputs with more visually stimulating presentations via graphs and charts. Conversely, a more sophisticated statistical audience may look for a greater degree of technical information and specialized reports to bring nuances in the data to the surface.

Dissemination policies

777. Most statistical agencies have some form of a formalized dissemination strategy which regulates the general types of outputs that are produced from a survey. There are four general types of products or outputs produced. The first is the *general interest publication.* This publication is prepared in anticipation of general user demand and contains statistics or information relating to statistics from the survey. The second type of output is the *special interest product.* These publications are produced to satisfy the anticipated demand and requirements of a limited range of specialist users on a strictly commercial basis. The third sort of output is *customized products and services* that are made to order for specific customers. And the final type of output is termed *other services* which involves a multitude of things including the provision of statistical advice.

778. Statistical information may be delivered to users in many ways including through direct access to datasets, through remote-access terminals, or as information provided on CD-ROM or other removable media. Some services may incur a fee for users.

779. Regardless of the form of output, the statistical agency should be clear about the manner in which the outputs will be disseminated. For example, four options may be applicable to the release of a crime victimization survey. First is the *general release.* Publications released in this way are simply released to the general public at a nominated time on release day, and provide the main findings of statistical collections and compilations, and more detailed statistical and related information. A second option is to *release under embargo.* In this scenario, access to statistical and related

information cleared for publication *prior* to general release may be granted to pre-approved parties. The purposes of such a release may be in order to prepare special briefings for Government Ministers responsible for the topic covered by the survey, or other significant stakeholders. The third option is to *release for peer review.* This method provides statistical information, prior to its clearance for publication, to approved parties for review, provided that such release will ultimately add value to the statistics or analyses to be released. This may be particularly important if the survey manager and staff do not have backgrounds in criminal justice or victimization statistics, and there is a need to seek external expertise. The fourth option is to *release for purposes of joint publishing.* In this case, the statistical agency may give joint publishing partners access to statistical and/or related information not otherwise released, including information from confidential unit record files. Depending on the arrangement between partner agencies, this exchange of data may be quite limited, and result in an endorsement and joint-badging[2] of the final product, through to a detailed collaboration on analysis and composition of the final output.

Standards for data dissemination

780. Standards for data dissemination are a key consideration in determining the dissemination of statistical output. Standards are a primary means of ensuring that data quality principles are followed and can be relied upon by users. It also allows the statistical agency to quality assure the statistical outputs using existing methods. While standards generally apply to all the methods and procedures for the collection, processing, storage, and presentation of statistical data, this becomes vitally important for users when trying to establish appropriate use of output. If a statistical agency has well established standards that are applied in the presentation of statistical output, these formats become understood and expected by users. The inclusion of relevant information on the sources, methods, and procedures of the statistics is also important, as are full explanations of standards applied to the data – for example, the structure of offence classifications utilized, and other statistical tools.

781. Confidentiality policies are a core element of standards application in the dissemination process. This can be vitally important in protecting the identity of respondents. The privacy of respondents is always a paramount concern to a statistical agency which hopes to obtain the future trust and cooperation of citizens in its processes. Confidentiality becomes even more vital when the crime victimization surveys tackle more sensitive subject matter, such as sexual assault, domestic violence, help-seeking behaviors and the effects of criminal victimization. Steps should be taken uniformly across statistical output to ensure that cells with small counts of persons are not revealed, and that personal information is not included. Where unit record files are to be released as a statistical output, this process can be lengthy and complex. If there are small numbers of persons with combinations of responses that are highly distinctive, it is often necessary to swap values between records, or group data in ways that protect the anonymity of the survey respondents. The strategies required for ensuring the data remain confidential varies as a product of

[2] Joint badging means that two agencies can have the logos of their organisations on the published product if it is a joint publication. For example Australian Bureau of Statistics and Australian Institute of Health and Welfare.

the formats in which the data is to be released, and the amount of control these formats give the statistical agency to control the further manipulation of data.

Dissemination formats

782. Dissemination formats are crucial in determining whether or not the statistical output can be easily accessed and utilized. Published reports are a traditional vehicle for statistical outputs, but increasingly other options have been made available as a result of emerging technologies and communication channels. Users may also specify their own formats that are required for certain analysis to be possible. Ultimately, decisions about the appropriate medium for the survey output should be geared towards convenience. Formats can be made available through a number of channels: Direct contact with clients, through library programs to disseminate statistical findings, contributions of articles and like materials to other agencies for publication and release, or through the internet and World Wide Web.

Box: Potential formats for statistical dissemination

- Printed Publication
- Loose Leaf
- Brochure
- Map
- Microfiche
- Floppy Disk
- Magnetic Tape
- Cartridge Tape
- CD-ROM/DVD
- Removable Memory (e.g. flash drive)
- Data-report Hard copy
- Service
- Electronic Delivery/internet-based distribution (email, PDF, HTML, interactive modules)
- Other Media

783. Some formats and services are very resource-intensive to produce, and this expense is often passed on to users by the statistical agency in the form of access fees or charges. When considering the output formats available, survey managers should consider the costs involved in producing an output via a particular format, and how well the resulting output can be utilized with the necessary charges applied. For instance, if unit record data is required for users to complete comprehensive analysis of survey results, but the cost of producing such data is high, the cost applied to the final output may prohibit any of the targeted users from accessing the data. In such instances, the resource expended is not recovered, and the data cannot be utilized.

Metadata and documentation

784. Explanatory material is provided as a principal means of achieving informed use of statistics. It provides guidance and interpretation to help users understand the statistics presented, and determines appropriate applications of the data. The final output from a crime victimization survey should contain, or be clearly associated with, sufficient explanatory materials to enable users to:

- be aware of the nature and limitations of statistical information
- assess the usefulness of the statistical data for specific purposes, and
- use the findings about crime victimization for decision making.

785. Explanatory material should inform users about the quality of the statistics and any process or events that might affect interpretation. This can include measures of accuracy, such as providing technical notes that detail the effects of relative standard error, for example. The variation that can arise as a result of a particular survey methodology needs to be communicated to users – particularly in statistical environments where there are multiple sources of crime victimization data that may utilize different methods, as users will require this information to determine which data is best for their purpose, and which data cannot be compared. Draft explanatory notes from the point of view of a user who is not an expert in the field. Avoid using statistical jargon and technical terms. Do not simply translate jargon into 'ordinary language', but explain it, and recast texts where it is used. Simply translating technical language into 'ordinary' language without also dealing with the logic in which it is framed can result in texts that are even more confusing to readers than the original technical prose. Sampling issues, such as coverage difficulties, sample loss, potential bias and any significant events that occurred immediately prior or during enumeration need to be included in the explanatory material for the survey findings. Information about significant events may be obtained through media, or news about major events, or may be communicated by stakeholders who are aware of local issues that may have a significant impact upon the data.

786. In some instances, the survey output to be disseminated may be linked to an ongoing survey collection, but be the result of a significantly different questionnaire, definition used, sample, method or other survey process that is considered to have an effect on the data such that it is no longer comparable with previous iterations of the survey findings. In such instances, this is termed a *break in series*. Similar issues arise when presenting revisions to data that has been previously released to the public. These events can cause confusion among users, and lead to misinterpretations of data, or potentially even perceptions of invalidity of the data if not handled correctly.

Publications and access to the public

787. Part of the initial specifications for the survey should detail the requirements for outputs and data availability to the public. It is vital that the survey manager clearly determines strategies for dissemination to the public ahead of time. This allows the survey manager to plan activities and schedule resources from an internal perspective, with a focus on the deliverables and their due dates. The aim should generally be to release statistics as soon as possible after their compilation. The

release of survey outputs should be announced in advance of the expected release date, to enable stakeholders and other users to prepare for the arrival of the results. In many instances, researchers and policy makers keenly anticipate the release of new crime victimization data, and if scheduled dates of release are known ahead, it is possible for them to be prepared with appropriate resources in order to access and begin to work with the survey findings at the earliest opportunity. In this way, the results of the victimization survey will be as timely and relevant for users as possible, and the policies and programs development that results from the availability of new data will occur more quickly for users.

Comparisons between survey results and police reports

788. Police data and victimization survey data are quite different and highlight different aspects of crime. The purposes of the two types of data differ, and are variably suited to different applications depending on the research or policy questions to be addressed. The two data sources can be viewed as complementary to each other. When an incident of crime victimization occurs, there are a number of ways in which this can be measured and a number of stages where a measurement can be taken. This includes from the time that the person perceives they have been a victim through to reporting to police and the formal charging of the offender. Police statistics reflect a measure of crime reported to and recorded by police, while victimization surveys represent direct reports from members of the public about their perceptions and experiences of crime as collected in household surveys. Neither of these sources provides a definitive measure of crime victimization, but together they provide a more comprehensive picture of victimization than either measure alone. Both sources have their limitations and need to be clearly understood.

789. In a number of jurisdictions, data based upon police administrative records and data gathered from crime victimization surveys are presented together in the same publications. There are advantages and disadvantages associated with this approach. The obvious advantage is that one publication contains a broader view of crime. In contrast, a disadvantage is that it can lead to attempts at comparisons between the two data sources that should not be made. Additionally, if users do not fully understand the differences between the two datasets, it can lead to users choosing the number that looks ‘best’ or most helpful to support a particular hypothesis or argument. This is always a risk when disseminating statistical information, but one that should be carefully managed when dealing with crime statistics. If there are differences in the ways that offenses are defined or categorized, then this needs to be very clearly stated, or the presentation altered in such a way that comparisons between related, but distinct concepts, are difficult to achieve.

790. The statistical agency needs to be very clear about the way that the survey can be complementary to police statistics or other sources, and instances where there are differences that may impact upon statistical inferences that can be made. These differences should be explained in any published statistical outputs, regardless of whether police and survey data is released in the same published outputs.

791. One of the most informative pieces of data that can come from a crime victimization survey is a measure of crimes that have not been reported to the police,

and therefore a reporting rate that shows the propensity of people to report incidents to authorities of the criminal justice system. Depending on the research or policy questions, this can be a more helpful piece of data than police records in some instances. However, it should also be noted that the scope of offenses or topics covered by the crime victimization survey may not be easily reconcilable with the legal definitions of offenses used by police. These legal definitions are bound by criminal law, operational policies and procedures, and thus are also prone to change over time. For example, changes to a law, computer system or recording rules can all have effects on the number of criminal incidents recorded in police statistics. Crime surveys generally use broader descriptions of behavior and harms, which can be free from these effects.

Box: Comparison of survey and police data

Survey data	**Police data**
• personal reports from individuals • sample-based • various methodologies can be used to obtain the data • can be periodic • generally only samples persons of certain age groups (i.e., often excludes children or those younger than a certain age) • includes crime not reported to, and recorded by, police • concerned with personal and household victims of crime as the counting unit • covers a limited selection of offenses, which may include events which are or are not strictly in breach of the criminal law • error can result from respondent error, non-response, coding errors, editing, imputation and estimation, sampling error, and non-sampling error related to coverage	• sources from administrative records • generally gathered via a census approach • harvested from administrative systems electronically, or provided in machine-readable or paper format • usually ongoing or annual collections • includes persons of all ages • count includes only incidents reported to, detected by, and recorded by, police • concerned with criminal incidents as the counting unit • covers all criminal offenses defined by law • error can result from inaccurate reporting by the police or inaccurate recording by the police, non-responding police department or operational unit, processing error, edit failure, police discretion and/or decision making, changes in policy and procedures, and legislative change.

Media relations

792. An important consideration in presenting data to the public is preparation for dealing with the media. Crime victimization is often a topic which gains significant attention in the press and generally the media tend to focus on the negative rather than positive aspects of the results. The disseminating agency needs to plan in advance for this attention. . It is highly recommended that media training is provided to people designated as media contacts, especially given the highly politically sensitive topic of crime. Developing a set of questions and answers that cover a range of sensitive issues that might be raised by the media or general public prior to the release of the results would also be a useful resource tool for those dealing with the media.

793. Statistical agencies generally have a policy of equity of access to statistical information to all, whereby the initial statistical outputs from a survey are released to the general public and media at the same time and same date. Generally notifications of expected releases of statistical outputs are lodged on an organizations website.

794. Some statistical agencies have media relations units who provide comment on the wording of the media release. These comments generally are aimed at making the language used in the media release media friendly, thus increasing the chances of the findings being reported in the media. These same media units will fax or email copies of the media release (which is also released at the same time and date of the statistical outputs) to media agencies.

795. It is important to present an appropriately short and succinct media story in a press release form. For jurisdictions where the survey data is likely to be compared with other data such as from previous surveys or other police or justice data, it is helpful if the agency can be prepared about what is and is not comparable, and be prepared to explain this in simple, non-statistical terms.

796. One of the responsibilities for an agency releasing statistical relating to crime victimization (or any topic for that matter) is to promote the accurate media coverage of statistics wherever possible. Any release of statistics should include, as previously outlined, sufficient explanatory material to enable users to understand the origin and complexity of the data, in addition to any additional information that may relate to that specific survey. For example, this may include information about any significant events that may have occurred during the reference period or enumeration. While this information may assist more savvy users in understanding the data in depth, it is also important to have an agency contact listed for media inquiries, statistical clarification, and to coordinate general requests for additional information. Such a contact not only provides a face for the agency in relation to this topic, but also can assist media in finding the data that is most relevant to them from the survey, and advocate for appropriate representations of the findings. General public or other users of the statistics can also utilise such a contact to obtain further guidance about the appropriate interpretation of the statistics, and information about other data that may not be made available in the initial products disseminated.

797. Obtaining the support of key figures in the field of crime and justice - or specifically the topic of the survey - can be invaluable in providing independent support for the findings. Such figures can become additional ‘survey champions’ or

spokespeople, who can speak on the topic and be briefed on the detail of the survey methodology and results in order to provide a public face for the findings. Depending on the limitations placed on the agency in terms of providing analytical comment on the results, these contacts may often be in a better position to provide more analytical and/or speculative comment on possible drivers for an increase/decrease in crime victimization.

Box: Guide to media releases

A media release is not intended to summarize all the details of the final publication. Instead, the media release is better described as a short, attention-getting and topical presentation of the most significant features of the results from the crime victimization survey. These features are likely to be of most interest to media editors and the public at large.

The main objectives of a media release are:

- To achieve wide reporting of the findings from the particular survey to help increase the awareness of the data among decision makers and the general community.
- To invite favorable publicity for a particular survey or publication.
- To encourage a positive media approach to the purpose or content of the publication.
- To advance the public perception that the statistical organization is effectively fulfilling its important and objective role in society.

Experience shows that these objectives are most readily achieved by a media release that:

Is journalistic in style. That is, the release is short, written in plain language with an eye-catching heading and opening paragraph and a straightforward, logical presentation of interesting facts.

Honestly sets out to satisfy journalistic (and therefore public) curiosity about the subject matter. The release should not try to obscure unpalatable findings by omission or through misleading language.

Is self-contained. That is, the release should not present questions which are unanswered in the text.

While setting out to meet these requirements, it is important that the release not resort to sensationalism or distort the facts through misleading headings or content. **The media release should not undermine other statistical organizational objectives** (e.g. objectivity and impartiality) nor lead to a story that trivializes a dataset, or leads it off on a tangent, and hence undermines community discussion.

Start writing the media release text well before the publication's release date, and even before final clearance of the publication. Unless you are working from first-hand knowledge of the publication contents, you probably will be drawing on the 'Main Features' or 'Summary of Findings' text which is, of course, written for a different purpose.

There is a good argument for drafting the media release before the Main Features document is finalized. This helps to clarify issues of public interest or concern which might be overlooked in preparing the introduction to the publication.

The main principles of news writing (and therefore media release writing) are straightforward:

- Keep It Short and Simple. It is recommended that releases are kept to an absolute 400 word limit, and to one page. This provides maximum economy for distribution purposes.
- Keep your primary readers/audiences in mind: Media and general public.
- Emphasis on active voice (subject-verb-object).
- Short sentences, generally one idea to a sentence.
- Who, what, when, where and (where feasible) why.

News is about people. It is important to remind ourselves that putting the people into a media release makes it much more interesting to the intended audience.

Sometimes there will be a clear idea for an opening paragraph such as a large increase in a particular crime type. But if the main point of the survey results is not a stand-out you will have to give some thought to choosing a topic for the 'lead' or 'intro' as journalists call it. In this situation you need to mentally step out of your professional frame of mind and ask yourself these questions:

- What do these findings say about people's lives? Do they indicate changes or new insights into the way people live, work, relax, plan, worry or relate to each other?
- Consider the ‘biggest, newest, latest, most’ in your release. The media uses this approach a great deal.
- Would newspaper readers or radio listeners be surprised by any of these findings?
- Would any of this information cause some people to change their attitudes or their plans for the future?

Change, or in some cases the lack of it, makes news. If your publication's main features summary does not make comparisons with previous time frames or previously published figures, you should check these yourself to see if you can use them to make your media release more effective. You may also wish to include information from another publication relating to subject of the media release. Such inclusions are permissible provided the title and catalogue number of the source publication is given.

As a rule, any material in the media release should be available in the Main Features, and the material in the Main Features should be sourced from the publication.

You might come up with three or four possible candidates for 'most interesting point'. You can use them all in the text. Make a choice for number one and then write your opening paragraph as a simple statement (you can decide on a heading later).

Journalists are taught to think of an inverted pyramid while writing their stories. The most important fact or facts come at the broad top (the inverted base), with supporting information immediately following, then additional interesting facts in descending order of importance, with the least important/interesting/significant material in the narrow apex at the bottom. The advantage of this image is obvious --- shortening the text is a relatively easy matter of cutting from the bottom. When space is at an absolute premium the 'base' --- the first paragraph --- will stand on its own.

While recognizing that most news organizations will use the media release only as a starting point for their own version of the story, one should aim to present the text in good style so that it could be published without further editing. This means that the journalist is working from a text that is valid in its own right and needs no basic correction or explanation --- it is ready for 'customizing' to their own requirements with creative editorial treatment.

Try to keep the heading comfortably to one line. A short, eye-catching heading will draw an editor to the lead paragraph, so look first to your lead for a heading. Try to vary or simplify the language to avoid repetition, although some repeats are to be expected.

A heading should make basic sense but should be seen as a form of shorthand: it need not have a verb and generally will not have room for qualifying phrases.

The style of the media release

- Spell out numbers one through nine and all numbers that occur at the start of a sentence. Percentages are an exception because of the use of the symbol (5%).
- Be sparing with bullet points: They can be useful in small numbers.
- Spell out acronyms when first used, followed by the acronym in brackets: Consumer Price Index (CPI). Afterwards just CPI is fine.
- Publication titles in italics.
- Use capitals sparingly. Media releases follow standard (media) style.
- One well-spaced page is the optimum length for any media release.
- Allowing for letterhead and 'master' items such as contact details, the actual content must be less than three-quarters of an A4 page in length. In some special cases extra material can be distributed to key media contacts, but this is costly and labor-intensive.
- Authors should keep in mind that the website 'Main Features' document will be available online to media, who will also have access to the publication itself. This is not a reason to leave key points out of the media release, but an incentive to avoid trying to cram too much into the available space.
- Journalistic writing and scripting is not a language in its own right --- it is simply a version of the way people speak to each other in normal situations. It could be called 'conversational style'.
- The best approach to writing a media release is to write as if you were speaking to

a busy, intelligent stranger not of your own profession. Such a listener would quickly grow impatient with language that was not simple, concise, to the point and free of jargon.

- Choose words that are used in everyday life: *Chose* or *preferred* instead of *expressed a preference for, travelled* for *were conveyed, most* in place of *the highest proportion of.* Of course this list could go on for pages.
- Avoid bureaucratic language. In general, journalists and their audiences think in terms of people, not persons --- they regard that word as a bureaucratic term which rarely appears in normal conversation. Such terminology can irritate sub-editors and other journalists using the text, while it might be acceptable in a specialist document.

Do ensure the figures in the media release text are absolutely correct. Ask one or more colleagues to check them for you and then re-check them yourself. The issue here is not rounding of figures, which often makes them more digestible to the public, but transposition of percentages or simple typing errors which cause embarrassment if not discovered before the text is released.

Do keep an open mind on what might interest the media and the public. If a survey has raised questions or issues considered peripheral to the main purpose of the exercise, but these are particularly interesting in themselves, there is no point trying to hide them. At the same time if material is of such interest as to be in a media release, it should also be in the Main Features.

Do keep in mind the ultimate use of the statistics being highlighted. If possible include a reference to use of the statistics in the text.

Box: An example of a media release

MEDIA RELEASE

June 27, 2008 Embargoed 1130 am (AEDT)

Nearly $1 billion dollars lost by Australians to personal fraud : ABS

Nearly one billion dollars ($980 million) was lost as a result of personal fraud according to the first National Personal Fraud Survey, by the Australian Bureau of Statistics (ABS) released today.

The survey, conducted in 2007, asked people aged 15 and over about their experiences relating to personal fraud incidents in the preceding 12 months. The survey found that 453,100 Australians lost on average $2,160 as a result of personal fraud.

Other results from the survey include:

- A total of 806,000 Australians reported they were victims of at least one incident of personal fraud in the previous 12 months. This represented 5% of the population aged 15 years and over.
- Half a million Australians experienced a form of identity fraud. The majority 383,300 (77%) were victims of credit or bank card fraud; identity theft accounted for the balance.
- Nearly 6 million Australians (36%) were exposed to a range of selected scams; that is they received, viewed and/or read an unsolicited invitation, request or notification designed to obtain personal information or money or obtain a financial benefit by deceptive means.
- 329,000 people fell victim to at least one type of scam by responding to or engaging with the unsolicited offer. The three main categories of selected scams were: lotteries (84,100 victims), pyramid schemes (70,900) and phishing and related scams (57,800).

Media Note: Personal fraud included: credit or bank card fraud, identity theft (includes the unauthorised use of a person's personal details), and the following selected scams; lotteries, pyramid schemes, phishing and related scams, financial advice, chain letters and advance fee fraud.

Further information is available in *Personal Fraud, Australia 2007*(cat. no. 4528.0), available free of charge from the ABS website <www.abs.gov.au>.

Less formal dissemination

798. It should also be acknowledged that dissemination does not solely include physical statistical outputs. This can also include more informal information sharing

through liaison and promotion activities. If a community engagement strategy has been utilized in order to address the needs of a specific cultural group or other sub-population, dissemination can be a crucial phase of the survey process in providing meaningful feedback to respondents and their community. This may require the presentation of data in a different manner in order to make the results more relevant, and perhaps consideration of dissemination in a number of languages. Other forums, conferences and meetings can be utilized in order to return the findings of the survey to their source. If a longer-term strategy to engage communities is in train, these are methods of maintaining the trust and support of groups.

Summary of points: Publication and dissemination

- Preparation of data to be released from a statistical activity's source file usually involves many steps. Verify and ensure that released data, after all the processing steps, are consistent with the source data obtained. In the case of regrouped data or derived variables this means that one should be able to reproduce the same results from the source data.
- Verify the quality of the publications by ensuring that the information presented is relevant, the analysis is accurate, the release is timely considering the data collection period, the data are interpretable and coherent.
- Ensure that all rules concerning confidentiality of the data are followed by suppressing data that may identify an individual respondent.
- Test an electronic product before release to ensure that it performs as planned.
- Provide data quality measures or, where possible, tools for their calculation (e.g., CV look-up tables, sampling variance programs) along with the disseminated product.
- Provide documentation and metadata along with the disseminated material that contains, as appropriate, descriptions of its quality and the methodology used to ensure that users do not draw conclusions from the data that are inaccurate or inappropriate.
- Develop a dissemination product consistent in style and formatting to other previously released data from the survey to assist in its use and utility.
- Where possible, provide a contact person who is knowledgeable about the survey, a telephone number and/or an e-mail address for each release of information. This will help ensure the proper use of the data when users have questions concerning concepts, definitions, approaches and methodologies.

Data disclosure controls

Importance of preserving confidentiality

799. Cooperation from respondents in completing survey returns is dependent on the trust they have with the agency requesting them to provide personal information. Confidentiality is a key element of that trust and agencies need to protect the information supplied by respondents and ensure that information about individual people, businesses or other entities is not disclosed. Some national statistical agencies

have a legal obligation to protect information supplied by individuals and businesses. The sixth United Nations Fundamental Principle of Official statistics states:

800. "Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."

801. Without trust, national statistical agencies would not get the high level of response and accuracy to deliver good information to the community. If respondents believe or perceive that an agency will not protect the confidentiality of their data, they are less likely to cooperate or provide accurate data. Any assurances provided to respondents about privacy or how the data will be used should also be upheld. A lack of trust or misuse of data could result in negative publicity for that organization or in some cases legal action.

802. Personal information may also be subject to privacy rules. Information must be stored securely and only used for its intended purpose. The agency collecting the information should ensure that no one outside the organization has access to information that could identify an individual. This could include the questionnaire, database records or the outputs produced from the survey. Data must be protected at all stages of the of the collection process.

803. Technological developments provide considerable challenges in ensuring that an individual's personal information is protected, therefore it is critical that a survey agency develops a set of rules or policies, methods and techniques to sustain the trust of its providers. Strict procedures in collecting, processing, storing and releasing information should be in place.

Methods of accessing data

804. Users of data have become more sophisticated and are demanding fine level data for research. The form that the data is released in will affect confidentiality requirements. The most common form of presenting data is via multidimensional tables or data cubes.

805. Another form of dissemination is through micro data access. Microdatasets containing individual records provide a rich source of data, however the risk of disclosure of personal information can be high. Microdata can be confidentialized by removing identifying information, recoding values or changing variables. This process can be complex and involves a subjective process therefore the confidentiality should be performed by someone who understands the data and the risks. The rapid expansion of databases in the private sector containing information about identifiable persons can increase the risk of identification through data matching even though names and addresses are not included on the microdata file.

Confidentializing data

806. Some offense types in crime victimization surveys, such as sexual assault, may have low prevalence, resulting in some cells in tables having few contributors. Cells with few contributors in the unweighted data will need to be altered by collapsing rows or columns or both, suppressing cell values or perturbing cell values prior to data being released. Care should be taken to ensure that cells that are suppressed cannot be calculated from the totals or other values. Cells that are suppressed because they fall below a set number of contributors are called primary suppression cells. It may be necessary to suppress other cells within a table to ensure that the value in the primary suppression cell cannot be revealed; this is referred as secondary or consequential suppression. If the confidentializing action severely restricts the amount of information available from a table (i.e. too many cells are suppressed), then the table design should be reviewed; collapsing some rows and/or column categories may improve the table. The table below is an illustration of data that have been suppressed; the total person counts are still publishable, but the male and female counts cannot be supplied. There are various methods in confidentializing data and this will depend on policy or the situation. Some software packages are also available to protect microdata against disclosure.

Table: Victims of robbery: Type of weapon used in most recent incident

	Males		Females		Persons	
	'000	%	'000	%	'000	%
Total number of incidents of robbery	67.6	..	23.8	..	91.4	..
Knife	np	np	np	np	11.1	18.8
Other weapon	np	np	np	np	8.5	14.5

np = not published .. = not applicable

References

Guidelines and core principles for managing statistical confidentiality and microdata access - Taskforce on Managing Confidentiality and Microdata Access - Conference of European Statisticians, June 2006.

Ch. III.U. Closing the circle evaluating existing surveys

807. A survey system can be viewed as a cyclical process. It begins in the design stage, with developing and testing the survey. Stages two and three are the implementation of the survey and the production of survey results. Stage four is the review and evaluation of the process as it has been carried out. Finally, in the last stage, the cycle is begun again, as the next iteration of the survey is developed and implemented.

808. Previous chapters have discussed various procedures that should be incorporated into the first three stages of the survey cycle to ensure that the survey's goals are clear and its protocols and processes are well developed, and the quality of the work is high and

accomplished in a cost effective manner. For example, systems should be created to monitor data collection costs and the quality of interviewers' work and their progress in completing their assignments. During data processing, the data must be examined closely to determine whether the questionnaire is measuring what it was designed to measure; and the extent to which non-sampling error; that is error associated with respondent recall, coverage, etc. are present in the data.

809. This chapter covers the fourth stage of the survey process cycle; the post-enumeration evaluation of the survey system. Once the data are collected and the results produced, the various processes, protocols and procedures employed in the survey should be thoroughly reviewed to assess the reliability, accuracy and utility of the data and to improve and enhance its ability to collect timely, accurate data in as cost efficient manner possible in future enumerations.

810. Most, if not all, of the discussion that follows can apply to any survey regardless of subject matter. Evaluations are of particular importance to victimization surveys because such surveys are generally expensive undertakings measuring relatively rare, often difficult to measure and sensitive events. Problems that arise can have substantial impact on the survey's results and therefore must be identified and corrected.

The importance of the evaluation process

811. This evaluation is a vital component of the survey process. It is not solely about the quality of the survey estimates and whether the procedures and methodologies have been carried out well, but whether the procedures, questionnaires, and methodologies meet the survey's goals, whether the survey data have the intended utility, whether the costs for fielding the survey are appropriate and within reasonable constraints and whether any of these factors require refinement or revision. It is a systematic review of the entire survey process, both a qualitative and quantitative review of the program.

812. Evaluation of the survey takes many forms. It involves reviewing the data to assess the presence and extent of sampling and non-sampling error, determination of whether the survey's questionnaire is obtaining the information intended, whether the survey's other protocols are operating as designed, and whether the survey's overall goals are being achieved.

813. The nature of review and evaluation necessarily depends on the nature of the survey being evaluated. The evaluation of a one time survey focuses on the past to determine the quality of the data that has been collected. Ongoing or periodic surveys must also be reviewed towards incorporating improvements and additions into the methodology of future iterations. If problems have been identified in the current enumeration, these must be identified and methods developed and tested to correct the problems. It also possible that the survey results may identify additional questions or issues that could be addressed in future enumerations. Again, these must be researched and protocols developed for incorporation into the survey. The following sections describe some of the areas that should be reviewed in the post-enumeration evaluation. Among the topics are data quality, survey goals and objectives, the utility and benefits of the data collection, and the protocols and questionnaires used. Additionally, a robust program will also incorporate a program of methodological research.

Data quality

814. Earlier chapters of this manual have discussed data quality and how to identify and measure and address sampling and non-sampling error. Some of this review take place during data collection, but much of the review can only be accomplished upon completion of enumeration when all the data are available for analysis. Post enumeration review should encompass a comprehensive examination of the errors associated with conducting a survey; including coverage errors, response errors, and data processing errors. The assessment of the existence and magnitude of these errors is important regardless of whether the survey is a one time effort or an ongoing program. It creates a level of transparency in presenting the survey results. Publishing information about the quality and limitations of the survey estimates helps to promote acceptance of the estimates and preempt criticism about them. Such analyses also inform efforts to eliminate problems in subsequent iterations of periodic or ongoing surveys.

Survey goals

815. As discussed in chapter II.A, it is vitally important to establish clear goals for a victimization survey. The post-enumeration evaluation should include a review of the goals, and the extent to which they have been achieved. For example, a major goal of the survey might be to measure national victimization rates at a certain level of precision. The review should include whether or not this goal has been met. If not, the cause for the failure should be identified. It may be that the sample was too small, or that the participation rate was too low, or that the victimization rate measured by the survey fell below that anticipated during the planning stages. Identifying the cause can enable implementing a correction.

816. It may be necessary to modify the survey's goals to meet the constraints imposed by budgetary or resource limitations. For example, it may not possible to increase sample or introduce other modifications to the survey to increase precision of estimates. Such a situation may require modifying the goal to accept a lower level of precision, or focus on different sets of victims or offenses or create a substantially different objective for the survey.

817. Survey goals may also change in response to changes in information needs. These needs may be identified during the post enumeration survey review, but also may also come from outside. A governmental agency may impose a new role for the survey or a new problem may arise that can be addressed through a victimization survey.

Utility of the survey

818. Victimization survey results have a variety of potential uses. They can provide a social indicator function measuring the magnitude of the crime problem and its impact on society. The data can inform policy and legislation to address societal or crime related issues. They can be used in research related to crime and criminal justice, and they can also be used for planning, administration, and evaluation of programs and policy. The uses are dependent on the survey's goals and the questions that it asks. In addition, the utility of the survey is also a function of other factors, including the quality of the data, and how it is disseminated.

819. The post-enumeration evaluation should examine the current and potential utility of the survey. The review should examine who has used the data and the ways the data have been used. Various methods can be applied to this effort, such as conducting a literature review, convening focus groups or conducting surveys of potential user communities. The

review should examine how well the dissemination process facilitates or hinders the various potential uses and benefits of the data.

820. The development and nurturing of a broad user community is vital for the growth and continuance of any survey program. The survey must be seen to be serving an important function, be it informing the public, informing public policy, providing a platform for research into societal problems or providing other important information. Evaluating the current and potential uses of the survey enables the sponsor to develop and expand upon those uses. This analysis can identify the impediments that exist towards any of the potential uses for the survey so that the means can be developed to eliminate them.

821. The evaluation should also examine how well the survey's goals coincide with the actual uses of the data. A goal of the survey may be to inform policy, but if the information collected does not promote or enable such use, than that goal cannot be met.

Review of survey protocols

822. While monitoring the survey's protocols during data collection should identify those that are problematic, a post-enumeration review of the survey's protocols and procedures is also important to the success of the survey. Often, only when the data are being analyzed is it discovered that a question series or survey procedure is not as successful as had been believed. The analysis may determine, for instance, that there is a great deal of missing data, or too many responses to a particular question fall into a "catch-all" category. A review may identify mode effects on the estimates, or establish that additional attempts to reach respondents do not substantially increase the likelihood of obtaining an interview. These analyses may not correct the problems in the past enumeration, but it does inform improvements in subsequent iterations of the survey.

823. Such an evaluation can include experiments to determine the viability of alternative survey protocols. For example, when the U.S. National Crime Survey was first implemented, a proxy respondent was used to obtain information for household members 12 and 13 years of age unless a parent explicitly stated that the youth could respond for him or herself. This procedure was implemented because of a fear that parents would find the questions too sensitive or difficult for such young respondents. Because information from proxy respondents was deemed to be less reliable than that obtained from respondents themselves, a study was conducted to determine whether parents would consent to their children being interviewed for themselves. After the study found that few parents actually object to such interviews, the protocol was changed to accept personal interviews from 12 and 13 year olds unless a parent objected.

Methodological research

824. A program of systematic review of survey methods and protocols is necessary to maintain and improve the quality of the survey's data. However, this aspect of survey implementation is often under-funded and may be lacking at all because it is not seen as directly related to the mission of producing and publishing the survey's information. It is often viewed as an extra function to carry out if funds and resources are available. A well designed program of methodological research, however, can create economies in data collection and improve the quality of the survey data.

825. The research program should examine the quality of the data being collected and explore ways to improve the quality and utility of the data, as well as the extent to which economies can be introduced into the program without jeopardizing the accuracy or reliability of the results.

The U.S. experience

826. The U.S. National Crime Victimization Survey was introduced in 1972 (as the National Crime Survey). At the time it was implemented, the idea of measuring crime using a household survey was not widely accepted by the public or by policy makers. Many questions remained about the utility and reliability of the survey. The initial results of the new survey raised a number of new questions about the viability of a national victimization survey. The sponsor of the survey contracted for two evaluations of the survey's methodology, utility and benefits (see McMullan, et al. 1978 and Penick and Owens, 1979). The recommendations of these studies guided a decade long process of research and study that culminated in a redesign of the survey in 1992. The research carried out during this period was intended to examine virtually every aspect of the victimization survey, including sample selection, survey mode, survey error properties, subject matter coverage, respondent recall issues, questionnaire design, and analytic capabilities. It was carried out by a consortium of research organizations that possessed a broad range of expertise and skills selected through a competitive process.

827. The survey redesign incorporated a number of changes to the survey including the introduction of computer assisted telephone interviewing, a new crime screening protocol that enhanced respondents' ability to recall hard to enumerate crimes such as violence by non-strangers and rape, the addition of sexual assault as a measured offense, inclusion of life-style questions to enhance analyses of risks of criminal victimization, and a revision of the survey's protocol to measure repeat victimization.

828. Over the history of the NCS/NCVS, other research projects have been undertaken to evaluate and improve the survey's methodology. One such study examined the survey's reference period, and compared three month, six-month and 12-month reference periods in terms of accuracy of respondent recall and the costs of data collection.

829. In 2008, in response to rapidly rising costs of enumeration in a time of fiscal austerity, the survey's sponsor instituted a new program of methodological research, similar to that conducted twenty years earlier, but with a focus on decreasing the costs of producing reliable data on crime victimization. This effort will examine the feasibility of using less expensive modes of collection, changes in the survey's reference periods, identifying more efficient sampling frames, and evaluating the survey's goals and objectives.

References

Groves, Robert M. and Cork, Daniel L., eds. *Surveying Victims. Options for Conducting the National Crime Victimization Survey*. National Research Council. Panel to Review the Programs of the Bureau of Justice Statistics. Washington D.C. 2008.

McMullan Jr., Philip S., Collins, James J., Gandossy, Robert, and Lenski, Joan G. *Analysis of the Utility and Benefits of the National Crime Survey (NCS).* Research Triangle Institute. Research Triangle Park. 1978.

Penick, Bettye E. and Owens III, Maurice E.B., eds. *Surveying Crime*. National Research Council. Panel for Evaluation of Crime Surveys. Washington, D.C. 1976.

United Nations. *Handbook of Household Surveys*. (Revised Edition). Department of International Economic and Social Affairs. Statistical Office. Studies in Methods Series F. No. 31. New York. 1984.

Taylor, Bruce. *Redesign of the National Crime Survey*. U.S. Bureau of Justice Statistics. Washington D.C. 1989

Table of content