



Department
for International
Development



Tools for measurement, monitoring and evaluation

IN-DEPTH FOCUS ON SURVEYS

Date: 14 November 2012

Practice Products for the CCVRI

Improving Measurement in DFID Crime, Conflict & Violence Programming

This document is one of a series of Practice Products developed under the **Conflict, Crime, and Violence Results Initiative (CCVRI)**. The full set of products is intended to support DFID country offices and their partners to develop better measures of programme results in difficult conflict and fragile environments.

DFID recognises the need to focus on the **results** of its work in developing countries. To this end, DFID strives to account better for our efforts on behalf of UK taxpayers, offering clarity regarding the value and impact of our work. The Results Initiative operates under the assumption that we will achieve our development objectives with our national partners more effectively if we generate—collectively—a clear picture of the progress being made.

Within DFID, the Conflict Humanitarian and Security Department has established a partnership with a consortium of leading organisations in the fields of conflict, security and justice to develop more effective approaches to the use of data in the design, implementation and evaluation of programmes that contribute to reducing conflict, crime and violence. In addition to producing these Practice Products, the consortium has established a Help Desk function to provide direct and customized support to country offices as they endeavour to improve measurement of results in local contexts.

The Help Desk can be accessed by contacting helpdesk@smallarmssurvey.org.

The views expressed in this Practice Product are the sole opinions of the authors and do not necessarily reflect the opinions of all consortia partners. This Practice Product does not reflect an official DFID position.

Members of the consortium

Document Summary

Title:

Tools for measurement, monitoring and evaluation. IN-DEPTH FOCUS ON SURVEYS

Purpose and intended use of this document:

This document has a focus on household and polling surveys. This tool will help to assess whether the proposal is value for money and a prudent investment for DFID. The document provides clear and simple guidance for making this assessment. It introduces some technical considerations to highlight building blocks of survey proposals.

Key questions this document addresses:

- What are the requirements necessary for collecting the desired information in a reliable and meaningful way?
- When is a survey useful?
- What are the indicators of quality in a survey research *proposal* and how do we make a distinction between a strong vs. weak survey research proposal?
- What are the indicators of quality in a survey research *report* and how do we make a distinction between a strong vs. weak survey research report?
- What is the usefulness of data collected within a certain framework of elements (e.g. crime, violence and victimization, firearm possession, disarmament)?
- What are the minimum requirements (e.g. sample size)?
- What are the costs? Overview of survey budget items and examples of relevant costs from recent surveys.

Key messages/essential “take aways”:

- Undertaking a survey is essential when there is a need for information relevant to the direct experience of local residents.
- The ethical framework within which the survey is designed and administered is a crucial component of the research project.
- Different types of research not only produce different results, but have different costs and require different profiles of researchers to carry them out.
- A strong methodological component in the proposal will ensure that the survey is carried out in a manner that is cost-effective, valid, and respectful of cultural customs and traditions.
- Any credible survey proposal shall include pilot testing.
- If cost is an issue then it might be possible to reduce the cost of a survey while still retaining the same quality by using a smaller sample size.

Intended audience of this document (including assumed skill level):

DFID Country Officers, who need to read, review and appraise a proposal for a survey research project. No prior knowledge or experience with surveys required.

Key topics/tags:

SURVEY, VALUE FOR MONEY, RESEARCH ETHICS, RESEARCH QUESTIONS

Authors and their organizations:

Principal Author: Anna Alvazzi del Frate and Ryan Murray, Small Arms Survey

Contributors: Irene Pavesi, Matthias Nowak, Natalie Jaynes

Cross-references to other documents in the series:

Tools for measurement, monitoring and evaluation. IN-DEPTH FOCUS ON DATA SOURCES

Tools for measurement, monitoring and evaluation. IN-DEPTH FOCUS ON DATA USES

Table of Contents

1. COMPONENTS OF A SURVEY	6
1.1. When is a survey useful?.....	6
1.2. The ingredients of a quality survey proposal	7
1.2.1. Ethics Component	8
1.2.2. Research Component.....	11
1.2.3. Methodological Component.....	14
1.2.5. Research Questions, survey questions and indicator.....	18
1.3. Theory to Practice: Research Question Checklist.....	21
1.3.1 Indicators of a quality survey report.....	22
2. BUDGET	25
2.1. Survey Costs	25
2.2. Quality of Surveys, Value for Money.....	25
3. GLOSSARY	29

Introduction

This document has a focus on household and polling surveys. It is aimed at assisting the DFID Country Officer to read, review and appraise a proposal for a survey research project. Survey research is generally expensive, so this tool will help to assess whether the proposal is in fact value for money and a prudent investment for DFID. The document provides clear and simple guidance for making this assessment.

It introduces some technical considerations to highlight building blocks of survey proposals which can provide solid results as baseline and monitoring. The reader will gain knowledge of the requirements necessary for collecting the desired information in a reliable and meaningful way, taking into account the following:

- particular characteristics of post-conflict countries and weak institutional settings;
- producing cost-effective surveys;
- ensuring transparency of methods and replicability of surveys.

The document includes checklists for quality verification of survey proposals and associated reports. Items to be checked may include, inter alia:

- When is a survey useful?
- What are the indicators of quality in a survey research *proposal* and how do we make a distinction between a strong vs. weak survey research proposal?
- What are the indicators of quality in a survey research *report* and how do we make a distinction between a strong vs. weak survey research report?
- What is the usefulness of data collected within a certain framework of elements (e.g. crime, violence and victimization, firearm possession, disarmament)?
- What are the minimum requirements (e.g. sample size)?
- What are the costs? Overview of survey budget items and examples of relevant costs from recent surveys.

1. COMPONENTS OF A SURVEY

1.1. When is a survey useful?

A survey offers a unique glimpse into the attitudes, perception, beliefs, and knowledge of local communities, the voices of whom would otherwise remain silent in desk-based research or research with a select number of individuals representing the entire community (e.g. focus-groups, key-informant interviews).

Undertaking a survey is essential when there is a need for information relevant to the direct experience of local residents. This may include experience with crime, number of household members victimized by a crime, prevalence of household small arms possession, number of small arms present in the community, etc. If however, the research is more interested in indirect experiences of local community members then focus group interviews would suffice. Indirect experiences equate to attitudes, perceptions and estimations of a group of respondents, sampled from the local community, who speak on behalf of their community. These individuals relay their personal attitudes and perceptions towards specific topics like number of people victimized in their area, number of small arms present in the community, or attitudes towards security. Notwithstanding respondents' direct input, numbers can still be skewed due to biases in responding and reporting. This may occur in both in the household survey as well as focus group setting.

What is a Survey?

The term SURVEY refers to a strategic design of both quantitative and qualitative research and it relates to both practical and tactical matters to do with the detailed design of an instrument to be used (almost always a questionnaire or a composite of semi-structured questions) (Robson, 2002).

What is a Household Survey?

A HOUSEHOLD SURVEY is a survey design that treats the HOUSEHOLD as a unit of measurement. The term household may be defined differently according to the relevant context. In most cases, you will find that the term *household* equates to individuals who eat from the same pot or share from the same kitchen at least five of the seven days of the week. It is important that the proposal articulates this. Most household surveys use a questionnaire as the instrument of data collection.

What is a Victimisation Survey?

A VICTIMIZATION SURVEY is a survey targeting the personal experiences in the past where the respondent, or possibly any household member, was a victim of a crime or violent act in the past. The length of time precedent to the date of the interview can range depending on the goal of the research endeavour. Research questions in a victimisation survey, therefore, focus on details, experiences, and perceptions related specifically to the victimization event and may include the type of crime committed, the number of people injured or killed, the weapons (if any) used to perpetrate the crime, the perceived type of perpetrator(s) who committed the crime, the number of perpetrators, the time of day, the location, financial and/or psychological consequences, etc.

What is a Perception Survey?

A PERCEPTION SURVEY is a research endeavour assessing the attitudes and perceptions of the respective target population. Generally, a perception survey does not focus on factual information per se, such as dates, events, or specific experiences. Rather it gauges respondents' perceptions regarding certain phenomena, such as

factors affecting the overall functioning of their community, the causes of violence, the likely perpetrator and victim of violence, amongst others. A perception survey may also address respondents' attitudes towards their current security situation, the efficacy of their security providers, the presence of firearms in their community, the benefit of any potential disarmament campaign, etc. Perception surveys alone may not be sufficient to collect information useful for establishing baselines and monitoring trends. Surveys capturing the direct experience of respondents generally include large sections on perceptions. A balanced mix of the two approaches is necessary to obtain robust datasets.

1.2. The ingredients of a quality survey proposal

A survey consists of several different components which are necessary but not sufficient alone to administer an entire research endeavour. A quality survey proposal should consider the following (see also Table 1.1):

1.2.1. Ethics Component

- Ethics Review Team/Locally Informed Advisors
- Contact with local community
- Voluntariness, Confidentiality, Anonymity
- Informed Consent
- Risk Assessment

1.2.2. Research Component

- Qualitative Analysis
 - Desk-based analytical background research and contextual analysis
 - Focus Group Discussions
 - Key-Informant Interviews
 - Validation workshops with local communities
- Quantitative Analysis
 - Household Survey

1.2.3. Methodological Component

- Sample Design & Target Population
- Training of Supervisors/Enumerators
- Validation of Instruments
 - Cultural validation
 - Translation & back-translation
- Data entry, cleaning, validation and analysis

1.2.4. Research questions, survey questions and indicators

- Research questions
- Survey questions
- Survey indicators

1.2.5. Survey report

Each of the above components is discussed in more detail below.

1.2.1. Ethics Component

The most integral component of a research project is the ethical framework within which the project is designed and administered. An ethical framework must consistently bear in mind the risks posed to the respondents, their community, and the researchers conducting the fieldwork. Without clear evidence of and concrete steps toward ensuring this, the research project should not be considered valid until these are clearly integrated into the project and into the proposal.

Ethics Review Team/Locally Informed Advisors

The proposal should indicate that it will establish a group of social research experts, external and unaffiliated with the project, who may authenticate the ethical integrity of the research proposal. If this is not possible, it should still demonstrate an established partnership with locally informed advisors who can vouch for the validity, contextual appropriateness, and ethical integrity of the project. Ideally, an Ethics review team is a group of at least three professionals specialized in diverse areas of social research, including research ethics, methods and practice. These professionals will advise the survey team on the ethical integrity and methodological validity of their research project.

Contact with local community

Local communities may not be reachable by survey teams without the mediation of trusted community leaders or elders. The survey proposal should elaborate on how survey teams intend to gain the confidence of the various communities they wish to approach.

Incentives for survey participants?

Some survey proposals may include incentives for respondents, which may be in the form of telephone cards, gadgets or money. While incentives for community leaders are good practice and planning for them may indicate quality in a survey proposal (knowledge and appreciation of local customs), there is no evidence that providing incentives to individual respondents or households represents a good practice. The use of incentives is becoming more frequent in the North, where surveys are facing progressively declining response rates. In the South, response rates to face-to-face surveys are very high (generally above 80%). In post conflict settings in particular, provided they are adequately reassured about the independence of the survey, respondents are eager to freely talk about their experiences. In some cases, providing incentives may not be desirable or even possible, especially in very poor areas where even a small gift may expose the recipients to the envy of other community members.

A very good incentive is the prospect of providing feedback on the results of the survey. This may be planned in advance, for example in the form of briefings with community leaders to illustrate the main findings.

Voluntariness, Confidentiality, and Anonymity

Voluntariness

Decisions made by respondents and their respective community leaders must be autonomous and voluntary whereby no external influence mediated or coerced their decision in any way. Voluntariness signifies the participating individual's right to refuse

participation, refuse any question at any time, refuse to participate at any time, and right to have their interview dismissed and questionnaire destroyed if they so choose, at any point (Murray, 2012a). The respondent's voluntariness should be made explicit during the consent process (see Murray, 2012a for an overview of the consent process).

Confidentiality

Confidentiality refers to protecting the privacy of an individual and her/his family. Furthermore confidentiality signifies that "no information retrieved during the interview or elsewhere will be disclosed to a third party that is not directly affiliated with your research team" (Murray, 2012a).

Anonymity

Anonymity refers to "withholding any identifying information or any information that could be used to locate or identify the respondent and/or his or her household and/or family" (Murray, 2012a). Therefore, the proposal should clarify that there will be no possibility to retrace names, GPS coordinates of the residence or building of employment, or any other information that would allow the respondent or her/his household/employer to be identified.

Informed Consent

The proposal should clearly indicate a process to obtain consent through a specific form to be read out to participants to illustrate the scope and characteristics surrounding participation (voluntariness, anonymity, confidentiality, possibility to refuse or to check the credibility of the initiative, etc. (See Annex for an example of Consent Form)). Consent Forms shall inform about the survey goals, the nature and possible consequences of participation, the rights of participants during and after their participation, and the projected outputs of the research. Consent should not only be obtained from each respondent, but may also be necessary from village heads or community leaders. This should be specified in the proposal.

Risk Assessment

Any research conducted in post-conflict or high violence settings which addresses sensitive themes such as armed violence, victimization, arms possession, etc, engenders a level of risk to the participating individuals as well as the research team. The risk of harm is diverse and may vary across contexts, but may include retaliatory violence from family members or groups, imprisonment, psychological/emotional reliving of painful experiences, loss of job, ostracism by community or family members. These consequences may equally affect the respondents, as well as the interviewers and the research team.

It is, therefore, very important that the research proposal articulates clearly the risks of harm to the respondents in the project proposal, the participating community as well as the interviewers and the research team.

Table 1.1 provides an ethics component checklist to ensure that the essential constituents of the overall ethics component are provided.

Table 1.1 Checklist for Ethics Components of Research Proposal

N°	Question	Response	Sub-Question	Response	Notes
1.	Is the proposal indicating the establishment of a group of social research experts who may validate the ethical integrity of the research (i.e. an ethics review team)?	↑ YES			
		↑ NO			
		IF YES →	Is there a confirmation (written or otherwise) by this ethics review team on the ethical integrity of the research proposal?	↑ YES ↑ NO	At least one of these should be checked YES
		IF NO →	Have the researchers received written validation from their local informed advisors as regards the ethical integrity of their research proposal?	↑ YES ↑ NO	If both are NO, the ethical integrity of the research project cannot be confirmed
2.	Is there an explicit mode of establishing contact with community leaders?	↑ YES			
		↑ NO			
		IF YES →	Have the methods of establishing contact been validated by the research team's local informed advisors?	↑ YES ↑ NO	This should be explicit in the proposal
			Will there be reimbursement (monetary or otherwise) to local community elders?	↑ YES ↑ NO	Generally, offering something as a token of appreciation to the community leaders is encouraged
3.	Is there a consent process for both community leaders and concerned respondents/households?	↑ YES			
		↑ NO			
		IF YES →	Does the consent process discuss the research project's objectives?	↑ YES ↑ NO	These are all required in the consent process. If NO to any one item, the researcher needs to add the respective component accordingly.
			Does the consent process discuss the nature of participation for the respondent/household?	↑ YES ↑ NO	
			Does the consent process discuss the possible consequences of participation?	↑ YES ↑ NO	
			Does the consent process ensure confidentiality?	↑ YES ↑ NO	
			Are the projected outputs made explicit to the respondent?	↑ YES ↑ NO	
			Is there a mechanism established for the respondent to contact the research team at a later date if s/he wishes to revoke her/his questionnaire?	↑ YES ↑ NO	
4.	Does the survey proposal make it evident that the risks of harm to the respondents, the community, and the research team have been assessed?	↑ YES			
		↑ NO			
		IF NO →	<i>If NO, the proposal needs to provide a valid indication of the risks that may be posed to all concerned by the research project</i>		
		IF YES →	Has the survey team verified these risks with the local informed advisors?	↑ YES ↑ NO	It is generally recommended that the survey team receives input from the local informed advisors
			If the risks of harm as a result of participation in this survey are high, does the proposal articulate what measures the survey team will enact in order to mitigate such risks?	↑ YES ↑ NO	It is essential that protective measures are enacted to mitigate any possible increase of risk
			If the risks of harm as a result of participation in this survey are high, does the proposal convincingly argue why the benefits of such research outweigh the possible risks involved?	↑ YES ↑ NO	It is important that the reasons for such research are made clear, in light of the elevated risk of harm

1.2.2. Research Component

Table 1.2. Research Components of a Survey

Analysis	Data Collection Method	Target	Instrument	Number of People
Qualitative	Desk-Based Research	Contextual background research		
		Contextual data analysis collected from the field		
	Focus Group Interviews	Local community perceptions and attitudes about otherwise sensitive issues	Open-ended questionnaire using same questions used in the household survey	Max 15 per focus group, divided by gender and age
	Key-informant interviews	Experts and officials possessing privileged knowledge in respective field	Semi-structured questionnaire using same questions used in the household survey	1 per interview
	Validation Workshop	Local community perceptions and attitudes about findings, interpretations and recommendations of study	Presentation demonstrating key findings, interpretations, and recommendations of study. Selected questions for groups to discussed on their remarks, questions, and concerns following the presentation	To be decided by your locally informed advisors
Quantitative	Household Survey	Households of local community. Survey is to address attitudes, perception, beliefs, and knowledge of household members of concerned community with respect to various issues	Structured, closed-format questionnaire	Depends on resources, budget, and time. Generally a sample size ranges from 1000-3000 households (see infra)

As can be seen above, the majority of a survey research component is based on qualitative data and analysis both inside the office and on the ground. The household survey therefore complements the wealth of qualitative data by contributing quantitative data provided by the actual communities concerned. Some definitions are provided below.

Qualitative research

Desk research

This component includes literature review, taking stock of previous research and survey results, existing administrative data, current legislation and any other relevant information to be compiled with a view to contextualizing the study and its results. A good survey proposal should also include indications of the possible sources where to retrieve information from.

Focus Group Interviews

A focus group interview, or discussion (FGD), represents a collective interview engaging between 5-10 people who are qualified to speak on behalf of the group from which you wish to retrieve information. During a FGD, survey questions may be responded to at length and with more in-depth feedback by the respondents. The survey questions address the overarching questions of the research project and may allow for follow up

questions to increase clarity and precision in the responses. Frequently, FGD inform an instrument that is intended for a larger survey (e.g. household survey) (see Table 1.2).

Key Informant Interviews

Key informant interviews (KII) are qualitative in-depth interviews with community experts from the field of interest or with individuals who have privileged information regarding the local community thanks to their group affiliation or rank. Some examples are taxi drivers, pharmacists, doctors, military officials, police, or community elders. These individuals have first-hand knowledge of current issues facing the community and can offer recommendations for solutions (see Table 1.2).

Validation Workshop

Once the data have been collected and analysed and the final draft of the preliminary report has been composed, the research team should organize a validation workshop in order to authenticate the findings in its cultural context with the team of local informed advisors but also local stakeholders, such as local residents, civil society organizations, local non-governmental organizations, and relevant members of government.

A validation workshop ensures that the data have been appropriately interpreted and that the dissemination of the findings respects local norms and values. This serves, thus, as an integral component of the survey project. It should be scheduled after the envisaged deadline for a final draft of the report and it should occur in the area(s) most concerned by the themes addressed by the survey (see Table 1.2).

Quantitative Research

Household Survey

As stated above, a household survey is a survey design that treats the household as a unit of measurement. The term household may be defined differently according to the relevant context. In most cases, you will find that the term *household* equates to individuals who eat from the same pot or share from the same kitchen at least five of the seven days of the week. The typical instrument is a questionnaire, administered face-to-face with the respondent or over the phone, where this is feasible (see Table 1.2).

It may not be necessary to include all research components in a study. Table 1.3 below is a checklist which can assist in better understanding the merits of conducting a household survey, a key informant interview, or a focus group discussion. One or more **YES** responses with a cogent argument for **Why** would provide a justification to consider a survey as a complementary mode of data collection.

The different types of research not only produce different results, but have different costs and require different profiles of researchers to carry them out.

1.2.3. Methodological Component

In order for the administration of the survey to be carried out in a manner that is cost-effective, valid, and respectful of cultural customs and traditions, the proposal should indicate a strong methodological component which will serve as the supporting structure throughout the project. Constituents of a methodological component include sampling design, training of supervisors and enumerators, validation of instruments which include cultural validation process as well as a translation and back-translation process. These are discussed in greater detail below.

Target Population and Sampling Design

The *target population* is the “complete set of units to be studied”. The target population will always be strictly related to the overarching questions and objectives of the survey project.

The *sampling design* is a construct that lays out the methodological framework for recruiting a representative sample of target population whilst mediating the environmental, logistical, safety, and practical constraints that face every research team.

Types of Survey Designs

Stratified Random Sampling

Stratified random sampling (SRS) divides the target population into various strata, in which the researcher assumes there is great homogeneity. This homogeneity is in reference mainly to ethnic affiliation, language, national/regional identity, socio-economic status, geo-demographic attributes (i.e. rural vs. urban), sex, age, and personal experiences (crime, violence, disease morbidity).

As the aim of SRS is to retrieve a representative sample, recruiting proportions of individuals commensurate with those of the actual target population, a specific sampling method is quite useful. It is called Probability Proportion to Size (PPS) whereby the larger groups within each stratum have a higher likelihood of being sampled. This is also referred to as proportionate sampling.

Cluster Randomized Control Trials (30 x 30 method)

A survey using cluster randomized control trials (CRCT) is a survey method employing a type of randomized control trial in which groups, or clusters, of respondents (as opposed to individual respondents) are randomized (Bland, 2004). Advantages of CRCT include the capacity to study interventions across a group of individuals rather than toward specific individuals, and it is flexible enough so that it may be conducted in a variety of conditions (Edwards et al., 1999).

A prominent method is the 30 x 30 cluster method whereby the researcher samples 30 clusters of 30 households. This would be in lieu of a larger household survey sample size of over 2'000 households. The clusters would be randomly selected and treated as a unit of measurement wherein the experiences and responses of the households should be homogeneous but different from those of other clusters.

Restricted Sampling: Convenience Sampling & Purposive Sampling

In the scenario where a population is inaccessible, the researcher may conduct *convenience sampling* which permits the researcher to recruit populations which are accessible to the researchers (e.g. close to a main road or the research facility, areas of

greater security). In the case of *convenience sampling*, however, it is important to note that this sample would not be representative, as an important component that would represent the target population, would be missing from the sample.

Purposive sampling, on the other hand is conducted when the survey wishes to address a specific group of individuals, usually based on their personal experience, their profession, or their expertise in a certain area. A *snowball sample* is a variation of the purposive sample and is generated by identifying a first respondent from the desired target population and asking him/her to suggest another respondent from the same target population. Snowball samples may prove useful when trying to reach populations particularly difficult to reach, such as criminals, illegal migrants, etc.

The myth of random sampling

Virtually any sampling technique may generate a random sample, provided that selection of respondents contains some elements of unpredictability at some stage. The uncertainty of the estimate will be expressed by the sampling error. The presence of the word “random” itself does not add quality to the survey proposal. At the same time, its absence does not reduce quality.

Apart from the main selection of survey locations, the actual selection of households to be interviewed should be indicated in the survey proposal. A frequently used method is the *random walk technique*, which allows reaching respondents by applying simple rules (for example, tossing a coin to decide the direction to take, interviewing every third household, etc.). The method for selection of the individual respondent in the household should also be specified. For example, the person aged 15 or above whose birthday is next, or an adult woman, etc.

Sample size: finding the balance

Surveys use probability samples with the objective of collecting a base of data large enough to allow inferences about the survey population. The size of the sample required is determined by the desired precision (margin of error) which is sought at the various levels of stratification. For example, a smaller sample may be sufficient if representative results are sought for male and female respondents at the national level instead of sub-national level. Another way to look at the size of the sample is to consider whether information is sought on direct experience of victimization, which is a relatively rare event, or perceptions of respondents, which elicits answers from the entire sample. Finally, the available budget may dictate what size of the sample will be feasible: if the sample is smaller than desirable, it is still possible to obtain accurate estimates, by applying statistical adjustments – if possible supplemented by data from other sources (such as administrative records) to improve precision (see UNODC-UNECE, 2010, p. 31 and section on Estimation).

Table 1.4. Sample Sizes for Populations of 1 Million and More

Margin of Error (Per cent)	Confidence Level		
	99 per cent	95 per cent	90 per cent
± 1	16,576	9,604	6,765
± 2	4,144	2,401	1,691
± 3	1,848	1,067	752
± 5	666	384	271

Table adapted from Jackson (2007) as cited by Morra-Imas & Rist (2009), p. 366

Small is beautiful?

Despite what intuition may tell us, a small sample can still be representative. However, small percentages of responses are subject to large fluctuations due to the margin of error. For example, see the following example with a level of confidence of 90%:

- 20 of 1,000 respondents said “yes” to question X, thus 2% of the sample. The margin of error in this case is 0.7, so there is 90% probability that the rate falls between 1.3 and 2.7%.
- In another survey, 40 of 2,000 respondents said “yes” to the same question X, thus again 2% of the sample. The margin of error in this case is only 0.5, so there is 90% probability that the rate falls between 1.5 and 2.5%.

Therefore, the larger the size of the sample, the higher the probability that the true percentage lies close to the observed value.

A larger sample implies higher costs: it will be necessary to balance the size of the sample with the desired level of accuracy. For example, if it is expected to measure rare events and to analyse them according to various strata, the size of the sample should allow for such a breakdown.

Training of Supervisors/Enumerators

The survey proposal should include a training module for the prospective supervisors and interviewers (also known as enumerators). The training should possibly occur over the course of 4 to 5 days and should include topics such as the following:

- Ethical survey administration
- The goals of the questionnaire and the definitions used
- The specific topics and questions of the questionnaire
- Review of translated version to ensure uniform understandings
- How to conduct an interview
- Mock interviewing
- Pilot interviewing in the field with real respondents
- Group exercises to build team spirit and comradery

Quality indicator

Carrying out surveys in post conflict or high violence areas requires well trained personnel and involves an important capacity building component. Training – although relatively expensive – represents an important predictor of the quality of the survey.

Training may also be provided as training of trainers. This may be the case when different languages / dialects are going to be used by different teams of enumerators.

A training course on interviewing and supervision should generally be accompanied by a thorough training manual that provides definitions, examples, and exercises for class work and homework. There should generally be an exam at the end. Participants in the training should be graded continuously throughout the course and at the end. Those who pass will become your supervisors/enumerators while those who do not pass should generally not be invited to serve as interviewers/enumerators.ⁱ

Validation of Instruments

Before a survey instrument may be administered it will need to pass through a series of validation steps. First, the instrument, generally a questionnaire, will need to be initially validated by a team of locally informed advisors. This team will vet each question and determine their contextual appropriateness, relevance, and clarity. Next, the questionnaire should be translated into the local language(s) by professional translators. This should then be back-translated by a different translator. Any discrepancies noted in the questionnaire, between the original version and the back-translated version should be pointed out to the locally informed advisors who can generally assist in the refinement of the translation. Although it would be desirable that translators do not have prior knowledge or connection with the survey and its objectives, in reality translations require some experience in the subject matter to avoid using wording which would result odd for the respondents.

Pilot testing

The last validation procedure of the questionnaire will take place in the field whereby enumerators pilot the instrument and test their skills by interviewing actual respondents in the area around the training facility. By administering at least two questionnaires each, supervisors and enumerators can provide feedback from their experiences and the responses of respondents which will inform the researcher and her/his locally informed advisors on how to improve the final version of the survey instrument.

Quality indicator

Any credible survey proposal shall include pilot testing. The pilot may represent a serious challenge for survey teams, so it is important to verify that it is indeed carried out. Since it comes very late into the process, pilot testing may require last minute changes which may represent additional costs (for example, reprinting of questionnaires if they need modifications). A provision for such costs is an indicator of quality.

Data entry and analysis

Below are useful steps to ensure that the survey proposal addresses how data will be effectively managed. In general, proposals also mention software packages which are likely to be used at each respective step.

ⁱ In certain situations, refusal of employment as a supervisor/interviewer may jeopardize the safety of the researchers, the research team, the trainers or the research facility. It is therefore important to assess the safety situation of training individuals who may be perceived as being at-risk for retaliatory violence in the case of being refused a position on the research team. If this assessment occurs post-hoc (i.e. after the fact), it is still important for the researcher and the locally informed advisors to make the appropriate decision on how to engage this specific individual.

Data Entry

Data entry can be done locally or remotely. The disadvantage of paper questionnaire data entry is the cost of labour, printing and the time it takes to enter the data. For paper questionnaires, a software called *Census and Survey Processing System (CSPro)*. This can be downloaded free of charge at US Census bureau websiteⁱⁱ.

A potential solution to such these setbacks is the use of digital handheld tablets or smartphones. Data may be entered during the interview on such devices, equipped with an app (application) that is compatible with smartphone/handheld devices (e.g. android).

There are still some drawbacks to using such devices in the field. First is the longevity and maintenance of the electronic device, specifically battery life and potential application bugs that need updating. Next, possessing such devices in disadvantaged areas may put interviewers at risk of threat, theft, or assault. Finally, there is a risk of theft or loss of the device due to its appeal. These are issues that would need to be discussed with the research team and the locally informed advisors to ensure that safety and quality are maintained throughout the administration of the survey.

Data Cleaning & Validation

Once the data have been entered, cleaned and validated, they may be analysed by various data analysis software. For quantitative data analysis, the most frequently used software is the *Statistical Package for the Social Sciences (SPSS)*ⁱⁱⁱ. For qualitative data analysis, *NVivo*^{iv} or *ATLAS.ti*^v.

1.2.5. Research Questions, survey questions and indicatory

What is a Research Question?

A *research question* should be considered as one overarching area of interest, or theme, which the research team wishes to address in their survey research. While this question should be focused and purposive, this research question may be broad, as it will later inform the design of the *survey indicators*, which will require great specificity and clarity.

Examples of research questions are:

- What is the prevalence of firearms possession in this country?
- What is the rate of armed violence victimization?
- How do local residents in this region perceive and feel toward their security providers?
- How safe do residents feel in their community?
- What local factors should be considered before launching a disarmament campaign?

The above-mentioned research questions are, indeed, general and broad. However, they invite a wealth of more specific questions, called *survey indicators*, which will narrow the focus and increase the clarity of these questions. These indicators will require validation from the research team's *local informed advisors*^{vi} and from their *ethics review team*.^{vii}

ⁱⁱ <http://www.census.gov/population/international/software/cspro/>

ⁱⁱⁱ <http://www-01.ibm.com/software/analytics/spss/>

^{iv} http://www.qsrinternational.com/#tab_you

^v <http://www.atlasti.com/index.html>

^{vi} See Glossary for definition of "Local Informed Advisors"

What are Survey Indicators?

Survey indicators represent the concretization of the subject matter addressed in the larger research questions. A survey indicator may represent a specific figure, quantity, or specific perception or attitude.

Examples are: age, number of firearms owned by households, number of households victimized by a crime, attitude towards police or toward the military, or number of households willing to report a crime to their traditional elders.

In order to formulate questionnaire questions, survey indicators are then elaborated into specific *survey questions* which must be validated as relevant and culturally appropriate. This is generally undertaken by the local informed advisors. These survey questions are generated to retrieve information relevant to the respective indicator as the actual questions that are asked during the interview, whether it is a household, focus-group, or key-informant interview. Survey questions are simply the questions that will be asked, either in a semi-structured interview (e.g. focus group) or in a structured closed-format interview (e.g. household survey).

What are Survey Questions?

Survey questions are the actual questions that are asked to the target population. Survey questions represent the evolution of an overarching research question, which transitions into a focused survey indicator. Addressing one survey indicator may necessitate several survey questions. See Table 1.5 for examples of the transition from research questions to survey indicators to survey questions. Survey questions should be precise, simple, and clear. For examples of good versus poor survey questions, see Annex IV. Good vs. Poor Survey Questions.

^{vii} See Glossary for definition of “Ethics Review Team”

Table 1.5. Evolution from Research Question to Survey Question

N°	1 st Step Research Question	2 nd Step Survey Indicators	3 rd Step Survey Question
1.	What is the prevalence of firearms possession in this country?	Number of households owning a firearm	Does anyone in your household own a firearm?
		General perception of firearm ownership in community	In your opinion, how many households in your community own at least one firearm?
		Number of firearms owned by households admitting firearm ownership	If your household does own a firearm, how many firearms does you household own?
		Evolution in firearms availability over the past 1 year	Compared to one year ago, how has the availability of firearms changed?
2.	What is the rate of armed violence victimization?	Number of households victimized be a crime within the past one year	Have you or anyone in your household been victim of a crime within the past one year?
		Of the households victimized by a crime, number where crime involved a firearm	If so, did the crime involve a firearm?
		Number of victims resulting from armed violence crime	If so, how many people in your household were victims as a result?
3.	How do local residents in this region perceive and feel toward their security providers?	<i>Rating of trust toward, efficiency of, accessibility of security provider</i>	On a scale from 0-4 (0= not at all, 4= very much), how do you rate the police on trust; on efficiency; on accessibility
		Likelihood of seeking help from security provider when victimized or under threat	If you were victim of a crime, would you seek help from the police? If you were threatened, would you seek help from the military?
4.	How safe do residents feel in their community?	Perceptions of security in and outside of home, during day and at night	On a scale from 0-4 (0=not at all, 4=very much) how safe do you feel... <ul style="list-style-type: none"> • In your own home during the day • In your own home during the night • Walking outside your own home during the day • Walking outside your own home during the night
		Perception of the evolution of security	<ul style="list-style-type: none"> • Compared to one year ago, how is the security in your community today? <i>(possible response options)</i> <ul style="list-style-type: none"> • Very much less safe • Somewhat less safe • No change • Somewhat safer • Very much safer
5.	What local factors should be considered before launching a disarmament campaign?	Rate of firearms retention despite incentives	Could you be persuaded to relinquish your firearm?
		Types of incentives desired	If so (<i>referring to above</i>), what would persuade you to relinquish your firearm?
		Perceptions of safety of community after a disarmament	In your opinion, would disarmament in this area affect security? If so, how? <i>(possible response options)</i> <ul style="list-style-type: none"> • Very much less safe • Somewhat less safe • Somewhat safer • Very much safer

1.3. Theory to Practice: Research Question Checklist

As mentioned above, a variable is an indicator defined as targeting or highlighting the theme/phenomenon in question. A variable has more than one category or value. Some examples are age, sex, number of firearms owned, rating of perception towards security provider, etc.

Table 1.6. Theory to Practice: Research Question Checklist

N°	Question	Definition	Examples
1.	Time Frame	The period of time you wish to focus on. This must be specified in the interview instruments and in the survey proposal	<ul style="list-style-type: none"> • Present-day • Since 1 month • Since 6 months • Since 1 year • Since 5 years
2.	Geographical Location	Geographical region of interest. This must be specified in the survey proposal and the samplings strategy must prove how it will succeed in collecting a representative number of respondents for the specified region	<ul style="list-style-type: none"> • Entire Country • One City • Select number of Counties/Districts
3.	Comparing & Specifying Group Attributes	The survey proposal must specify the groups that the research is interested in. The sampling and instruments must take into account these group specifications	<ul style="list-style-type: none"> • Geographic Region • Urban/Rural • Regions of High vs. Low Volatility • Ethnicity/Tribal affiliation • Gender • Education
4.	Direct experience of respondents	Capturing responses illustrating experience of respondents.	<ul style="list-style-type: none"> • Victimization • Ownership of firearms • Contact with the police or other authorities
5.	Context, specifications	Particular nuanced topic or theme of your overarching subject	<ul style="list-style-type: none"> • Place, time of victimization experience • Characteristics of perpetrators • Household Perceptions of being victim of armed violence • Types of weapons perpetrating armed violence • Armed violence & gender-based violence
6.	Abstraction	Is the research team interested in raw facts or in what local residents may indicate at a more abstract level (e.g. attitudes, perceptions, indirect knowledge)?	<ul style="list-style-type: none"> • Perceptions of security • Attitudes towards police or military • Estimates of firearm possession in community • Prediction of outcomes from a potential disarmament

1.3.1 Indicators of a quality survey report

The survey proposal should provide information on the type of analysis and reporting which is expected to produce. This will include sections on each of the questionnaire headings and responses to the main research questions identified by the proposal. The type of statistical analysis shall also be specified (generally percentages) and whether a weight will be calculated (including how it will be calculated). The final report shall include essential information on the main methodological aspects, according to good statistical practice. This should be clearly marked and easy to find and include sample size, date of administration, and narrative explaining units of measurement and analysis. (see Table 1.7).

Table 1.7. Indicators of a Quality Survey Report

N°	Indicator	Definition	Explanation & Examples
1.	Sample Size	The total number of respondents interviewed in the survey. This may also be supplemented by the final number of respondents after data cleaning and weighting	The survey report should clearly indicate the number of respondents initially recruited as well as those ultimately included after data cleaning and weighting. Total sample size is usually denoted as “N=”
2.	Date of Administration	The month(s) and year the survey was administered. This should also include the total duration of survey administration	The report should give the beginning date of the survey administration (by specific month and year) and indicate the duration (e.g. 3 months; 15 weeks, etc.)
3.	Percentage Explanation Narrative	Explanation on how what the percentages actually represent. A percentage cannot tell you the actual number of respondents, only the proportion of respondents relative to a greater whole. This whole, however, may be very small, for example 10 people. Thus if you say “50%”, it may sound like a lot, but it is only 5 people. The narrative should indicate the “n”, or number of respondents, along with the percentage	<u>EXAMPLE OF QUALITY NARRATION</u> “Of those who were victimized (n=10), 50% (n=5) said the event occurred at night” <i>Remark: We should see the actual number of respondents so that we can conclude that it may be too low to generalize to the target population. The author(s) should also highlight this limitation as well</i>
4.	Unit of Measurement	The entity which serves as one case. In a household survey, the unit of measurement is a household. Therefore each household represents one case. In a civil society organisation survey, the respondent is unit of measurement. Therefore, each respondent individually represents one case. This means that the respondent may work in the same civil society organization or domain (e.g. peacebuilding, humanitarian, development, etc.). However, in a household survey, this is not the case. As the household is the unit of measurement in a household survey, there should not be more than one resident from each household interviewed. Otherwise, this is likely to cause bias and potential redundancy in household experiences.	<u>EXAMPLES OF QUALITY SAMPLING</u> “In this survey, we sampled 2’400 households. Our total sample size is, therefore, 2’400 respondents.” “In this survey, we sampled 400 civil society organizations working in a total of 30 organizations within this region. Our total sample size is, therefore, 400 respondents.” <u>EXAMPLE OF POOR SAMPLING</u> “In this survey we sampled 500 households and 5 individuals from each household. Our final sample size was, therefore, 2’500 respondents”. <i>Remark: We should not see more than one respondent from each individual household. This would be an indication of a poorly sampled survey.</i>

Use of surveys for monitoring purposes

Repeating surveys represents serious methodological challenges, especially in the phase of interpretation of results. While it is possible to provide controlled characteristics for the wording of the questions and the overall methodology applied, a number of variables may intervene in generating results at different repetitions of the survey. It is important to take such challenges into account when comparing results from surveys carried out with similar methodologies either in different locations or at different points in time.

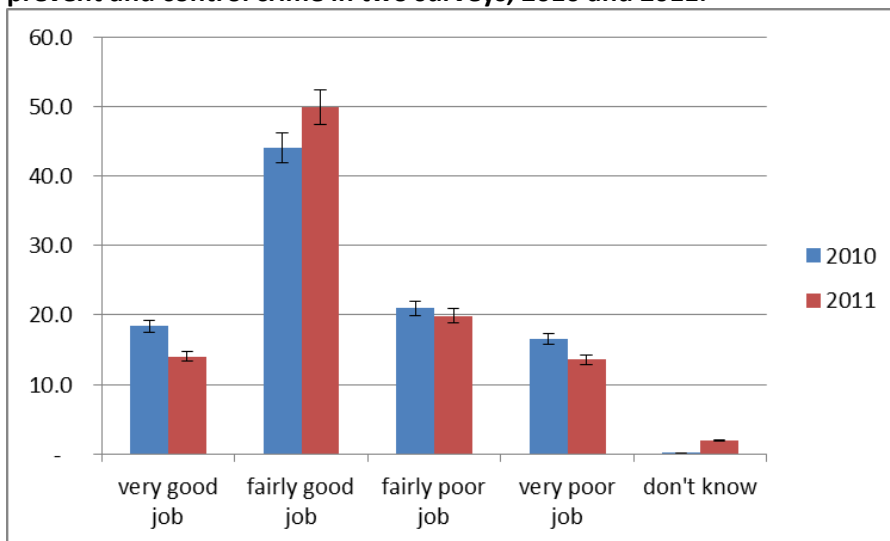
Some of the issues to be considered as possibly influencing results are the following:

- Introduction to the survey and training of interviewers
 - The way the survey is presented to respondents and the tone of the interviewers may determine their understanding of the questions
- Position of the questions in the questionnaire
 - If sequence of questions is different, results may change

Furthermore, it is prudent to interpret survey results with a margin of error, which may largely overlap in different surveys, thus creating some uncertainty in the reading of trends. However, large differences – not attributable to clear trends – are more likely to be found in rare events based on a small number of cases than in perception data based on the entire survey sample. For example, Figure 1 shows results of two different surveys carried out in Kenya in the years 2010 (UNODC, 2010) and 2011 (Wepundi et al., 2012). The two surveys carried out independently and with different objectives, were based on a similar methodology and questionnaire, including some questions with the same formulation and categories of response.

Responses to the question “What is your perception of the overall ability of the police to prevent and control crime: do you think they do a very good job, a fairly good job, a fairly poor job or a poor job?” were very similar in the two surveys. Although respondents may generally appear slightly more optimistic in 2011, the only marked differences are the decreased % of respondents who said the police do a very good job and the increased % of respondents who did not answer.

Figure 1: Percentage of respondents’ perceptions of the overall ability of the police to prevent and control crime in two surveys, 2010 and 2011.



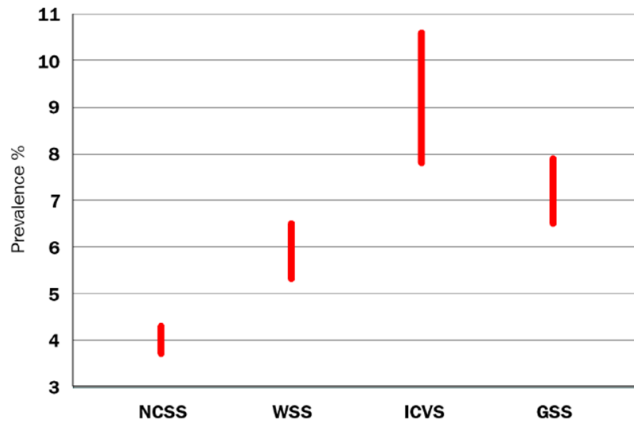
Sources: UNODC, 2010, p. 5 and Small Arms Survey Kenya database, 2011 (unpublished)

Conversely, despite efforts towards controlling data collection methods, different surveys may generate very different results. For example, Figure 2 shows results from 4 different surveys carried out by the Australian Bureau of Statistics within a period of approximately 2 years, referring to rates of assault victimization by females aged 18 and above. The surveys had different purposes and wording of introduction, while the wording of the question was the same, but with different positioning in the questionnaire.

Figure 2: Rates of assault victimization by females aged 18 and above within Australia, per four separate nationwide surveys



Survey Results (Assault - Females aged 18 and over)



Source: Australian Bureau of Statistics, 2004.

2. BUDGET

2.1. Survey Costs

Cost of surveys may vary depending on locations and availability of infrastructure. In some areas, it may be possible to work with local partners who have already survey experience or with statistical offices with survey capacity. The efficiency of the different solutions may vary. In some cases, even though local capacity is available, it still requires extensive training, which may be seen as a longer term investment.

Key elements to be taken into account in developing budget estimates include personnel costs (local researchers, supervisors, enumerators, data entry staff, and drivers) as regards their fees, transportation and per diems. Other costs to be considered are rental of vehicles (in most cases it will be necessary to use 4 x 4 vehicles to gain access to surveyed areas) and fuel, stationary and equipment. The cost of printing and reproduction of questionnaires may be significant, as well as the cost of transportation / shipment of filled questionnaires.

A general method to estimate survey costs is the *cost per completed interview*. This concept, largely used also in European surveys, refers to the global cost of personnel, transportation, and per diems, incorporating cross-cutting costs such as training, data entry and data analysis, divided by the number of completed interviews. Of course, such it should be clear in advance what items in the budget are included in the calculation of the cost per interview. In general, such costs are estimated on the basis of actual fieldwork and exclude, for example, the research fees attached to the design and development of the survey questions as well as any international travel of researchers. For example the *cost per completed interview* for surveys conducted by the Small Arms Survey in 2011 ranged from 12 USD in Nepal, 25 USD in Karamoja and 29 USD in Kenya.

Use of smart phones / tablets for interviewing: cost considerations

The use of mobile devices for conducting surveys presents clear advantages with respect to the traditional paper and pencil method. There is no need for internet or telephone connection, responses can be stored on the device offline, then shared with a central point of coordination when a wireless connection becomes available. Using such devices helps in training interviewers, who are facilitated in survey administration (the next question is automatically prompted), and eliminates data entry errors. Technology is becoming more and more accessible in terms of price. An estimate made by Small Arms Survey indicates that the cost of purchasing one device each for a team of 40 interviewers and 10 supervisors may be lower than the sum of the costs the use of the devices replaces: printing 2,500 questionnaires; transporting them back to the data entry site; recruiting 10 data entry staff for 10 days. Other considerations for the use of such devices may include the opportunity of equipping interviewers with flashy instruments which may make them too visible or even targets for theft and assault.

2.2. Quality of Surveys, Value for Money

As mentioned above, if survey proposals meet quality standards, they are likely to represent a good investment. Still, VFM may be difficult to measure. In the presence of good quality proposals, it may

be possible to assess whether the same quality could be achieved at lower costs. For example, it may be possible to adopt a smaller sample, without compromising on the quality of the results.

Example:

Survey proposal 1

Cost 100,000 - Sample 2,000

Will provide analysis disaggregated by geographical level (by county / province); urban/rural; sex (M/F); and age (5 age groups).

Survey proposal 2

Cost 50,000 - Sample 1,000

Will provide analysis disaggregated by broad geographical area (border/non border); sex (M/F).

Although proposal No 2 will result in a higher margin of error than No 1, it does not promise to deliver results where it would be dealing with too few cases. It may therefore be possible to consider whether lowering the costs would still maintain a “good enough” survey.

Applying here the example of margin of error of 90% presented above:

Survey proposal 1

- An observed frequency of 2% provides a margin of error of 0.5, so there is 90% probability that the rate falls between 1.5 and 2.5%.

Survey proposal 2

- An observed frequency of 2% provides a margin of error of 0.7, so there is 90% probability that the rate falls between 1.3 and 2.7%.

Proposal 2 is likely to be “good enough” to measure phenomena at the heading levels, especially if the expected frequency of responses is relatively high (indicatively above 2%, as very rare phenomena may not be accurately captured among small samples).

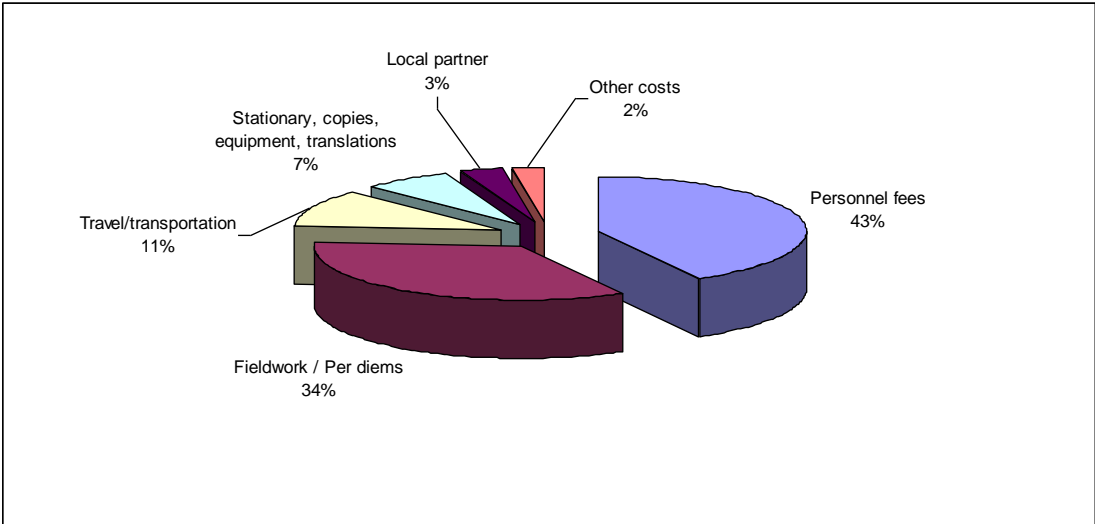
Example of Survey Budget Components

This section will provide an example of a survey budget, highlighting the various components to be considered when calculating costs. The survey budget can be disaggregated into the following main sections: (i) Transportation/Travel, (ii) Project Personnel, (iii) Fieldwork, (iv) Equipment, (v) Other costs (see table 2.1 and Figure 2.1)

Table 2.1 – Budget categories for a survey

	Approximate % of total
Personnel fees (Includes researchers, statisticians, supervisors, enumerators, data entry staff, drivers)	43
Fieldwork / Per diems (Includes per diems for the duration of training, fieldwork / data collection)	34
Travel/transportation (Includes reimbursement of fuel, local travel costs, rental of vehicles)	11
Stationary, copies, equipment, translations (Includes photocopies of questionnaire)	7
Local partner (Includes administrative fees and secretarial support from local partner)	3
Other costs Miscellaneous costs, research authorization, incentives for elders	2

Figure 2.1. Example of proportion of estimated costs for a nationwide survey in Kenya



As illustrated above, the costs of surveys can vary depending on the context in which the survey is carried out. For example, costs of transportation may be influenced in areas in which the fieldworkers cannot access other than with the assistance of a third-party (typically the UN). What this means in practise is that a) the third party assistance should be explicit and secured and b) the transport line item may often be covered by an ‘in-kind’ contribution in the event that the third party provides transport. This then may bring down the overall direct costs of the survey.

The cost of training

As discussed above, training of survey teams over several days represents a quality indicator. Depending on the contexts, training may also be provided as direct training of entire teams or “training of trainers”. The latter case may be necessary when different languages / dialects are going to be used by different teams of enumerators. These different settings may have an impact on the cost of the survey and should be highlighted to give the advisor the possibility to assess the value for money of the type of training suggested. For example, a relatively cheap training provided in English only may not reach all members of the team and result in poor performance during the fieldwork.



3. GLOSSARY

Ethics review team (aka Ethical Review Board): A group of at least three professionals specialized in different areas of social research. Areas should include research ethics and research methods and practice. These professionals will review your project proposal and advise you on its ethical rigor and methodological validity to ensure a research project that respects the rights and safety of those implicated in the research endeavour.

Local informed advisors: Local partners, local residents, and/or professionals who have worked extensively in the targeted area and who may provide otherwise inaccessible information about and insight into the various contextual components of the region and its population.

Research team: A concerted team of actors directly contributing to the administration of your survey research, from its conception and administration to reporting. These individuals rely on access to the data or data-related materials in order to satisfy their role in the research team. Therefore, these members, and only these members, should have access to data and data-related materials.

Survey Indicator: An operationally defined variable that adequately measures what the research team intends to measure. Survey indicators are the concretization of the subject matter addressed in the larger research questions. Specific questions which have been validated as relevant and culturally appropriate are then generated to retrieve information relevant to the respective indicator as the actual questions that are asked during the interview, whether it is a household, focus-group, or key-informant interview. See below for examples of the transition from research questions to survey indicators to survey questions.

(Source: Murray, 2012a)

REFERENCES:

- Australian Bureau of Statistics (2004). *Measuring Crime Victimization, Australia: The Impact of Different Collection Methodologies*. 4522.0.55.001 Information Paper. Canberra: ABS.
<http://www.abs.gov.au/ausstats/abs@.nsf/mf/4522.0.55.001#COMPARISON%20OF%20CRIME%20VICTIMISATION>
- Barnett, Chris, Barr, Julian, Christie, Angela, Duff, Belinda and Shaun Hext (2011). *Governance Indicators Report*. ITAD. http://www.dfid.gov.uk/r4d/PDF/Outputs/Mis_SPC/60797_ITAD-Gov-Conflict-Indicator-Rpt-Jan11.pdf
- Bland, JM (2004). Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Medical Research Methodology*, 4:21, doi: 10.1186/1471-2288-4-21
- DFID, UK Aid (n.d.). *Interim Guidance Note: Measuring and managing for results in fragile and conflict-affected states and situations*. (unpublished)
- Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ (1999). "Ethical issues in the design and conduct of cluster randomised controlled trials". *BMJ* 318 (7195): 1407–9. PMID 10334756
- Hext, Shaun (2012). *Programmatic Approaches to Conflict, Crime and Violence Data*. A report commissioned by DFID CHASE. Hext Consulting Ltd (unpublished)
- Jackson, Gregg B (2007). *Sampling for IEG Managers*. Presentation as George Washington University, Washington DC, December 18
- Kisielewski, M., Rosa, J.C. and Asher, J. 2010. *Statistical Approaches to Developing Indicators of Armed Violence*. Geneva: Technical paper prepared by StatAid for Small Arms Survey
[http://www.genevadeclaration.org/fileadmin/docs/Indicators/Statistical Approach to Armed Violence Indicators.pdf](http://www.genevadeclaration.org/fileadmin/docs/Indicators/Statistical_Approach_to_Armed_Violence_Indicators.pdf)
- Morra-Imas, Linda G. & Rist, Ray C. (2009) *Road to Results*, World Bank Publications, Washington D.C.
- McLean Hilker, Lyndsay and Ann Kangas, with Ellen Vanboegaerde (2011). *DFID's use of surveys and polls on conflict, security and justice*. Social Development Direct. (unpublished)
- Murray, Ryan (Forthcoming 2012a). *Small Arms Survey Manual to Ethical Field Research in Post-Conflict Settings*. Geneva: Small Arms Survey.
- Murray, Ryan (Forthcoming 2012b). *Small Arms Survey Guide to Ethical Field Research in Post-Conflict Settings*. Geneva: Small Arms Survey.
- Robinson, Louis N. (1933). "History of Criminal Statistics (1908-1933)", in *Journal of Criminal Law and Criminology (1931-1951)*, Vol. 24, No. 1. May – Jun. pp. 125-139. <http://www.jstor.org/stable/1135156>
- Robson, Colin (2002). *Real World Research. A Resource for Social Scientists and Practitioner-Researchers*. Second Edition. Blackwell.
- UNODC (United Nations Office on Drugs and Crime) (2010). *Victimization Survey in Kenya. Executive Summary*. Vienna: UNODC. http://www.unodc.org/documents/data-and-analysis/dfa/Executive_Summary_KENYA_final.pdf
- UNODC (United Nations Office on Drugs and Crime) and UNECE (United Nations Economic Commission for Europe) (2010), *Manual on Victimization Surveys*. Geneva-Vienna: UN.
http://www.unodc.org/documents/data-and-analysis/Crime-statistics/Manual_on_Victimization_surveys_2009_web.pdf

Wepundi, Manasseh, Eliud Nthiga, Eliud Kabuu, Ryan Murray and Anna Alvazzi del Frate (2012). *Availability of Small Arms and Perceptions of Security in Kenya: An Assessment*. Geneva: Small Arms Survey.

Annex I. Example of a Consent Form

Kenya

PERCEPTIONS & ATTITUDES OF SECURITY AND SMALL ARMS QUESTIONNAIRE CONSENT FORM

Keep this for your records:

1001

↑
Tear and give to participant

Check
When Read

- I would like to ask you a series of questions regarding your experiences here in your community.
- Some questions are personal and some people may find them difficult to answer.
- Questions pertain to attitudes towards the current security situation in your area and about experience with violence in the past.
- Other questions ask about any history of crime victimization in your household. By household we mean people who live under the same roof and eat from the same kitchen.
- We also would inquire about weapons possession and type of weapons owned by your community neighbors and household members.
- Finally, we would discuss with your thoughts on disarmament and how you believe it would affect your security and that of your household.

~~~~~

- The interview should last no longer than one hour.

~~~~~

- Answers will be treated with utmost privacy and confidentiality.
- You would choose the location of the interview that is private and comfortable for you.
- We will not share your responses with anyone not directly related to the survey (No one outside KNFP and Small Arms Survey).
- We will not be recording your names anywhere. No one will be able to link your responses to your name .
- If you feel uncomfortable at any time, you can let us know how we may help you feel more comfortable while you participate in our survey

~~~~~

- Your participation is at all times voluntary.
- There is no obligation to answer these questions,.
- You are free to refuse any question you do not wish to answer.
- You are free to stop the interview at any time.

~~~~~

- If, at any time, during or after the completion of this questionnaire, you would like your responses destroyed, please inform your administrative/village leader who will inform our contact person at KNFP. They will then destroy the questionnaire.
- In order to destroy your data, you must have Questionnaire Identification Number that is written on the questionnaire. I would provide you with that number.
- We would greatly appreciate your help in responding to this survey. Would you like to participate?

If respondent says "YES", tear off the number at the top right part of this page and give it to the respondent

Annex II. Summary of minimum survey criteria

Summary for establishing minimum criteria of quality of surveys

- Involvement of a local partner
- Transparent sample design and methodology
- Size of the sample adequate to expected level of analysis
- Questionnaire contains a mix of questions on direct experience and perceptions of respondents
- Clear indication of qualitative research (Focus Groups Discussions, Key Informant Interviews, etc.) to be conducted in support of a) development of the questionnaire; b) contextualisation of results
- Clear criteria for the survey design, implementation and reporting
- Overview of existing data sources, as well as other surveys done or ongoing
- Training of sufficient length (also if the interviewers are already trained), based on transparent training materials
- Pre-testing and piloting of the questionnaire
- Planning for recontacting of target respondents who are not available
- Planning for quality control and data checking

Summary for establishing minimum criteria of quality in looking at survey results:

- The final output should clearly indicate:
 - o nature and limitations of data;
 - o possibility to use data for specific purposes, including decision making;
 - o comparability with previous survey data or with other relevant data – if any.
- Standard statistical rules impose that any survey results should be accompanied by the following information:
 - o Size of the sample and sampling methodology
 - o Date of data collection
 - o Method of data collection
 - o Level of confidence of the results
 - o Weighting procedures.
- Survey data should not be presented alone, but triangulated with other data sources; especially qualitative research (Focus Group Discussions – FGDs or Key Informant Interviews – KIIs) conducted within the framework of the same study and based on similar survey questions.
- Presentation of results should be clear and balanced, facilitating the identification of levels and trends.

Annex III. Sampling in Depth

In order for the administration of the survey to be carried out in a manner that is cost-effective, valid, and respectful of cultural customs and traditions, the proposal should indicate a strong methodological component which will serve as the supporting structure throughout the project. Constituents of a methodological component include sampling design, training of supervisors and enumerators, validation of instruments which include cultural validation process as well as a translation and back-translation process. These are discussed in greater detail below.

Sampling Design & Target Population

The *target population* is the “complete set of units to be studied”. According to the UN Manual on Victimization Surveys (2010), a target population should be defined according to the survey research questions and project objectives. The target population will always be strictly related to the overarching questions and objectives of the survey project. While the respondents will be members of this specific target population, their total number in the survey will be constrained to a limited representative sample. The size of the survey sample and the method by which the research team recruits their sample will depend on several different logistical, empirical, and practical factors, all of which will directly impact the sampling design. All such elements, including justification of the reasons why one method is preferred, should appear in the survey proposal with ample details.

The sampling design is a construct that lays out the methodological framework for recruiting a representative sample of target population whilst mediating the environmental, logistical, safety, and practical constraints that face every research team. Examples of such constraints are flooded roads due to a rainy season, inaccessible villages due to lack of traversable roads, regions which are deemed as “no-go” because of security reasons, posing elevated risk on the researchers and their respondents, and budget limitations to the area that can be covered in the time allotted. Such factors will need to be taken into consideration and will inform the sampling design.

Although the constraining factors mentioned above are oftentimes unpredictable or unique in each case, the sampling design should be constructed from valid and standardized sampling algorithms and practices. The following types of design may appear in survey proposals:

Stratified Random Sampling

Stratified random sampling (SRS) divides the target population into various strata, in which the researcher assumes there is great homogeneity. This homogeneity is in reference mainly to ethnic affiliation, language, national/regional identity, geo-demographic attributes (i.e. rural vs. urban), sex, age, and personal experiences (crime, violence, disease morbidity). Generally, stratifying the survey sample is conducted hierarchically and can be effectuated over many stages. In terms of a hierarchical stratification, a larger stratum, like regional or district affiliation, is identified depending on the objectives of the specific project. For example, if the project wishes to inform programs or policies targeting geopolitically defined regions (as opposed to ethnically defined regions), using official geopolitical boundary delineations (such as region, district, or county) would be most appropriate. If, however, the purpose of the project is to target programs/policies mediated by ethnic or tribal affiliation, the first appropriate stratum would be ethnicity or tribal affiliation (Robson, 2002).

A second common stratum is geo-demographic location. This simply means whether the residents live in a rural, peri-rural, peri-urban, or urban area. It would be natural to assume that experiences within each location would share similarity within the specific location yet differ between geo-demographic locations.

As the aim of SRS is to retrieve a representative sample, recruiting proportions of individuals commensurate with those of the actual target population, a specific sampling method is quite useful. It is called Probability Proportion to Size (PPS) whereby the larger groups within each stratum have a higher likelihood of being sampled. This is also referred to as proportionate sampling.

Disproportionate sampling is also quite useful in the case where there are small but important groups to be taken into consideration, such as ethnic minorities. In this case, it may be useful to oversample certain small groups and to weight the mean of these groups during the analysis. Equally, the survey proposal may indicate

the possibility to oversample specific stratum groups if great variability of responses is observed during pilot-testing (e.g. within a certain district, within rural areas, females as opposed to males). Again, during analysis results from this group can be weighted to achieve the appropriate proportions commensurate with the target population.

The myth of random sampling

Virtually any sampling technique may generate a random sample, provided that selection of respondents contains some elements of unpredictability at some stage. The uncertainty of the estimate will be expressed by the sampling error. The presence of the word “random” itself does not add quality to the survey proposal. At the same time, its absence does not reduce quality.

Apart from the main selection of survey locations, the actual selection of households to be interviewed should be indicated in the survey proposal. A frequently used method is the *random walk technique*, which allows reaching respondents by applying simple rules (for example, tossing a coin to decide the direction to take, interviewing every third household, etc.). The method for selection of the individual respondent in the household should also be specified. For example, the person aged 15 or above whose birthday is next, or an adult woman, etc.

Cluster Randomized Control Trials (30 x 30 method)

It is not uncommon that groups of inhabitants experience similar events, similar disease outbreaks, and similar conflicts and thus share very comparable perspectives. In the case of disease or conflict, it is very common that these phenomena affect a specific cluster of people, perhaps limited to a village, an ethnicity, or a specific region or district. In this case, where residents vary very little within their group (e.g. ethnicity, village, geodemographic location) but vary greatly between groups, conducting a survey using cluster randomized control trials (CRCT) is very useful. The CRCT method is a type of randomized control trial in which groups, or clusters, of respondents (as opposed to individual respondents) are randomized (Bland, 2004).

Advantages of CRCT include the capacity to study interventions across a group of individuals rather than toward specific individuals, and it is flexible enough so that it may be conducted in a variety of conditions (Edwards et al., 1999). A typical case where CRCT would be useful is in the case of conflict or disease outbreak contained to specific area or group of people. In this case, it would be useful to cluster those known to have experienced such an outbreak and compare to a cluster of those who have not. A prominent method is the 30 x 30 cluster method whereby the researcher samples 30 clusters of 30 households. This would be in lieu of a larger household survey sample size of over 2'000 households. The clusters would be randomly selected and treated as a unit of measurement wherein the experiences and responses of the households should be homogeneous but different from those of other clusters.

Restricted Sampling: Convenience Sampling & Purposive Sampling

It is not uncommon that environmental, safety, or situational variables may prohibit the research team from undertaking a vast survey with a large sample size. Equally, it may not be feasible for the research team to access remote or exclusive populations due to logistical (e.g. remote village with no road access) or safety complications (e.g. a no-go area due to ongoing conflict). In such cases, the researchers may wish to go ahead and retrieve as much information as s/he can from the available population. In the scenario where a population is inaccessible, the researcher may conduct convenience sampling which permits the researcher to recruit populations which are accessible to the researchers (e.g. close to a main road or the research facility, areas of greater security). This would still provide researchers with informative data that could help move them further in understanding the local context and their issues. In the case of convenience sampling, however, it is important to note that this sample would not be representative, as an important component that would represent the target population, would be missing from the sample.

Purposive sampling, on the other hand is conducted when the survey wishes to address a specific group of individuals, usually based on their personal experience, their profession, or their expertise in a certain area. This is purposive because it is targeted towards specific individuals in target populations which are not large enough to allow for representative sampling with adequate proportions. A *snowball sample* is a variation of the purposive sample and is generated by identifying a first respondent from the desired target population and asking him/her to suggest another respondent from the same target population. Snowball samples may prove useful when trying to reach populations particularly difficult to reach, such as criminals, illegal migrants, etc.

Sample size: finding the balance

Surveys use probability samples with the objective of collecting a base of data large enough to allow inferences about the survey population. The size of the sample required is determined by the desired precision (margin of error) which is sought at the various levels of stratification. For example, a smaller sample may be sufficient if representative results are sought for male and female respondents at the national level instead of sub-national level. Another way to look at the size of the sample is to consider whether information is sought on direct experience of victimization, which is a relatively rare event, or perceptions of respondents, which elicits answers from the entire sample. Finally, the available budget may dictate what size of the sample will be feasible: if the sample is smaller than desirable, it is still possible to obtain accurate estimates, by applying statistical adjustments – if possible supplemented by data from other sources (such as administrative records) to improve precision (see UNODC-UNECE, 2010, p. 31 and section on Estimation).

Annex IV. Good versus Poor Survey Questions

The table below may assist in ensuring valid, reliable, and ethical questions in your survey questionnaire

N°	Question to ask	Definition	Incorrect	Correct
1.	Is the language simple?	Make sure the language used is colloquial and free of jargon that may be unfamiliar or foreign to your population	To a local rural Sub-Saharan village respondent: <i>"Your responses will be treated with the utmost confidentiality"</i>	To a local rural Sub-Saharan village respondent: <i>"We will not tell anyone about what we discuss today or the answers you give us"</i>
2.	Can the question be shortened?	Make sure your questions are short and clear.		
3.	Is the question double-barrelled?	Question which asks more than one question In such cases, questions should be broken down into one simple question each	<i>"Is the police in your area effective and trustworthy?"</i>	(a) <i>"Is the police in your area effective?"</i> (b) <i>"Is the police in your area trustworthy?"</i>
4.	Is the question leading?	Question where either the question structure or wording pushes people to provide a response that they would not have given had the question been asked in a more neutral way. Leading questions give respondents the impression that here is a 'correct' response. Thus, avoid linking an attitude, position, policy or whatever with a prestigious person	<i>"Do you oppose or favour cutting defence spending even if cuts turn the country over to communists?"</i> Question types to be avoided: <i>'Do you agree that...?'</i> <i>'...Does this seem like a good idea to you?'</i> Terminology is very important: 'Abortion' vs. 'killing unborn babies'	<i>"Would you favour cutting defence spending cuts?"</i>
5.	Is the question negative?	Question that contains a negative. This can often be confusing for the respondent.	<i>"Marijuana should not be decriminalized."</i>	<i>"Marijuana should remain illegal"</i>
6.	Is the respondent likely to have the necessary knowledge?	For issues where there is a doubt, ask filter questions to see if people are aware of the situation of concern The questionnaire should always offer the respondent the opportunity to claim ignorance to the situation if he/she does not know about it (e.g. do not know)	<i>"How effective has the UN Firearms Protocol been in your village?"</i> This question assumes that the respondent knows about this particular protocol. If the respondent has never heard of such protocol, their response may be biased towards what s/he believes the interviewer wishes to hear.	<i>"Are you aware of the UN Firearms Protocol?"</i> Here, you can immediately control if the respondent has the appropriate knowledge to answer the subsequent questions related to this subject matter. If so, you can ask the following question: <i>"How effective has this protocol been in your village?"</i>
7.	Will the words have the same meaning for everyone?	Definitions of certain terms may vary according to the cultural context. Examples are firearms, physical violence, armed violence, domestic violence, and crime. For instance, In certain contexts, domestic violence may be excluded from residents' definition of crime, which may lead to a bias in reporting; more specifically, an underreporting		
8.	Is there a prestige bias in the question?	When an opinion is attached to the name of a prestigious person and the respondent is then asked to express their own view on the same matter, the question can suffer from a prestige bias	<i>"What is your view about your village elders' choice to hire militia?"</i> Question would also be double-barrelled because could reflect attitude about village elders or about the militia	<i>"In your opinion, has your village been affected by the employment of militia"</i> If the respondent says yes: <i>If so, "In what way? Very positively, somewhat positively, somewhat negatively, very negatively."</i>
9.	Is the question ambiguous?	Questions that have ambiguous terms or verbs or refer to events or phenomena which are vague	<i>"Do you feel your village is affected by political instability?"</i> Without asking how the village is affected, we cannot ascertain the type of impact (positive vs. negative) political instability has	<i>"Do you feel your village is affected by political instability?"</i> If the respondent says yes: <i>If so, "In what way? Very positively, somewhat positively, somewhat</i>

			<p>on the village</p> <p><i>“During the war, was anyone in your household victimized by a crime?”</i></p> <p>This is vague in two ways. First, “the war” may not refer to a specific conflict. Second, “the war” is not a specific time period that may be perceived differently between respondents</p>	<p><i>negatively, very negatively?”</i></p>
10.	<p>Is the question too precise?</p>	<p>Question that asks for very precise information about information that may be difficult to recall. This can produce unreliable responses that will not be useful for your survey.</p> <p>Although this may, at times, yield precise figures, they are likely to be both inaccurate and unreliable</p>	<p><i>“How many times in the last year did any member of your household travel outside this village?”</i></p>	<p><i>“Approximately how many times per month do you or anyone in the household travel outside this village? Daily, once a week, biweekly, once a month, a few times a year, rarely, never?”</i></p>
11.	<p>Is the frame of reference for the question sufficiently clear?</p>		<p><i>“How often do you see police?”</i></p> <p>Without a frame of reference, the response can be subjective and difficult to classify with the other respondent’s responses</p>	<p>If you are interested in frequency, you may wish to articulate that the permissible response types are: <i>“daily, once a week, biweekly, one a month, a few times in the year, rarely, never”</i></p>
12.	<p>Does the question artificially create opinions?</p>	<p>It is important to always allow the respondent to offer a “non-response”. A non-response would be either <i>“don’t know”</i> or <i>“refuse to answer”</i>.</p> <p>Not including these will produce highly unreliable data as it would force the respondent to offer an actual response.</p> <p>To prevent this, always include the options “Do not know” and “Refuse to Answer” as responses for every question</p>	<p><i>“Have you or any member of your household been a victim of a crime or violent encounter in the last one year?”</i></p> <p>YES 1 () NO 2 ()</p> <p>This question is missing the “non-response” options “Do not know” and “Refuse to Answer”</p>	<p><i>“Have you or any member of your household been a victim of a crime or violent encounter in the last one year?”</i></p> <p>YES 1 () NO 2 () <u>DO NOT KNOW 88 ()</u> <u>REFUSE TO ANSWER 99 ()</u></p>
13.	<p>Is personal or impersonal wording preferable?</p>	<p>It is important to know when to ask the respondent directly and when to ask about experiences at the household level (i.e. indirectly).</p> <p>For instance, When discussing very sensitive issues, such as crime victimization, domestic violence, firearms possession, it is important to ask the respondent indirectly, in terms of household experiences</p>	<p><i>“Have you been victim of a crime?”</i></p> <p><i>“Do you own a firearm?”</i></p> <p>Not only can this place the respondent in an uncomfortable position, it may lead to an underreporting.</p>	<p><i>“Have you or anyone in your household been victim of a crime?”</i></p> <p><i>“Do you or anyone in your household own a firearm?”</i></p> <p>The respondent is now free to respond without attributing his/her answers directly to him/herself</p>
14.	<p>Is the question wording unnecessarily detailed or objectionable</p>	<p>There is information that we may wish to be precise, such as age. Placing people in Putting people in categories (age, income) prevents question from being too precise</p>	<p>WHAT IS YOUR AGE? 16-20 YEARS OLD 1 () 21-35 YEARS OLD 2 () 36-45 YEARS OLD 3 () 46+ YEARS OLD 4 () DO NOT KNOW 88 () REFUSE TO ANSWER 99 ()</p>	<p>WHAT IS YOUR AGE? AGE 1 ___ ___ DO NOT KNOW 88 () REFUSE TO ANSWER 99 ()</p>
15.	<p>Does the question have dangling alternate answers?</p>	<p>Alternate answers that are dangling are left “hanging” in the sentence because it does relate to the theme or subject it should.</p> <p>As a rule of thumb, the subject matter should come <i>before</i> the alternate answers are listed</p>	<p><i>“Would you say that it is frequently, sometimes, rarely, or never that you see the police?”</i></p> <p>Here, the alternate answers come before the subject matter and are dangling.</p>	<p><i>“How often do you see police? Frequently, sometimes, rarely, or never?”</i></p> <p>Here the subject matter comes before the alternate responses and is the correct way of phrasing the question.</p>
16.	<p>Does the question contain gratuitous qualifiers</p>	<p>Qualifiers are words or phrases, especially adjectives, used to attribute a quality to another word.</p> <p>When a qualifier is gratuitous it is unwarranted or uncalled for and may bias the responses of your respondents</p>	<p><i>“Do you oppose or favour cutting defence expenditure <u>even if it endangers our national security?</u>”</i></p> <p>The underlined phrase serves as a gratuitous qualifier that is unnecessary</p>	<p><i>“Do you oppose or favour cutting defence expenditure?”</i></p>

<p>17. Is the question a 'dead giveaway'?</p>	<p>Questions that are a dead giveaway are often absolute and all-inclusive or exclusive.</p> <p>It is best to avoid absolute, all-inclusive or exclusive words. See "Incorrect" column for examples of all-inclusive or exclusive words.</p> <p>Since words allow no exceptions few people will agree with the statement that includes them and this in turn will result in low variance and poor question discrimination</p>	<p>Examples of all-inclusive or exclusive words: <i>all, always, each, every, everybody, never, nobody, none, nothing</i></p> <p><i>"Everybody should be disarmed"</i></p> <p><i>"Nobody is safe in this village"</i></p>	
--	---	---	--