



Circum Network inc.
management and research consulting

74 Val Perché Street
Hull, Québec J8Z 2A6
(819)770-2423, ☎ (819)770-5196
service@circum.com
<http://circum.com>

rigour – transparency – creativity – relevance

Assessing Survey Research A Principled Approach

Paper prepared by
Benoît Gauthier, **Circum Network Inc.**

March 2, 2003

Assessing Survey Research A Principled Approach

by Benoît Gauthier

*A deviation from the ideal may
still be leagues ahead of no effort
to achieve an ideal.*

Robert C. Sorensen

This document presents an assessment framework for survey research. It is based on four fundamental principles: rigour, neutrality, balance and transparency from which a series of assessment criteria are derived. It does not offer a cookbook approach to the attribution of quality score points to survey research projects, neither does it report on the state-of-the-art lessons learned in good survey practices¹; rather, it aims at providing a structure within which the best professional judgement of the researcher can be put to work.

Introduction

Traditionally, the assessment of survey research has been an art more than a science. Each researcher uses their own criteria and put more or less emphasis on various aspects of the research. As Cohen and Shirley (1998,

¹ See Kaase (1999) and Diamond (2000) for such prescriptions.

1) stated: "Applied research professionals in the social scientific field are continually challenged by the difficulty involved in defining appropriate criteria for establishing and measuring the quality of their research. This lack of clarity regarding the definition of quality in social scientific applied research — along with the specific criteria by which quality is determined and measured — has resulted in considerable confusion regarding the appropriate objectives, scope, methodological guidelines and external applications for the discipline".

This confusion occurs in the literature when one author comments on the weight of the evidence brought to bear on a competing or supporting point. It also takes place when clients assess the value to give to survey evidence they commissioned or that is used, among other sources of information, in support of a business decision. It happens during one's self-examination of a research report to ensure that all elements of the required proof are offered.

However, the depth of examination of research evidence is nowhere more profound than when survey research is used in the judicial system. In this instance, two or more parties have every reason to either accept or reject the evidence; all parties tend to emphasize only the areas of the proposed evidence which support their position. Meanwhile, the courts and the quasi-judicial bodies need to be able to rely on a robust approach to assessing survey research so as to ensure that no important component of the research is left unexamined. The author's experience with the judicial system as a survey expert has motivated the development of a framework to assess survey evidence. It has also been used as a self-assessment tool and as a checklist to confirm the quality and completeness of the research.

This paper and this framework do not address the issue of whether or not survey research is the right approach to study a particular topic. Such an assessment would require investigating the relevance of the unit of analysis and the consistency of the use of the unit of analysis in the research (Grover, 1). This is beyond the objectives of this paper.

Criteria for quality research

The first fundamental question is that of the criteria defining quality research. Kaase (1999, 240) suggests that

"A *quality evaluation* methodology for surveys [...] would require an overall approach. In practice, surveys comprise numerous steps integrated in a defined programme. The quality of results is an outcome of the overall process. [...] The first requirement is a concept for the overall description of a survey type. The term *procedural model* is propose for this concept. A procedural model consists of various components or dimensions of the overall process. With regard to quality standards, it seems useful to distinguish the following six components: the master sample (coverage), missing data (nonresponse), interview mode (mode effects), data collection and processing (error possibilities/error avoidance), time requirement (rapidity), documentation (transparency)."

In our view, this segmentation of the survey work confuses *stages* of the process with *fundamental quality values*. We propose to separate these two concepts. We propose that four criteria are essential to passing judgment on a particular piece of survey research. They are:

- **Rigour:** does the study systematically apply recognized best practices of survey research? Of course, rigour is relative to the situation at hand. Certain research situations call for some methodological choices while others require other approaches. By essence, rigour is the most technical of the criteria used here; yet, the specific rules of assessment are likely to evolve through time as we develop more knowledge on best survey practices.² This is particularly the case with innovative approaches such as Internet-based survey research.³

² For example, from Hirschi and Selvin (1975, 7): "Ainsi, la méthodologie, ce n'est pas la vérité révélée et éternelle. C'est un corps vivant d'idées, qui change avec le temps. Nombreuses sont les méthodes, et ce qui est correct et valable aujourd'hui peut être incorrect et inacceptable demain."

³ For example, Couper (2000, 473) wrote: "tried and tested motivating tools used in mail surveys (e.g., advance letters, personalized signatures, letterhead, incentives, etc.) cannot be implemented in the same way in Web surveys, and functional equivalents are yet to be developed and tested. There is at present little experimental literature on what works and what does not, in terms of increasing response rates to Web surveys. Many of the techniques developed and tested over time to increase response rates in mail surveys may not work the same way in fully electronic Web surveys. Finding electronic equivalents of response-stimulating efforts is work that remains to be done."

- **Neutrality:** does the study observe the phenomena of interest without biasing the observations? The object of research is to depict reality faithfully; therefore, sound research must take steps to ensure that its depiction is neutral in its representation, that reality has every chance of surging in a representative way from the observations made.
- **Balance:** does the study use sufficient resources to demonstrate its thesis adequately, but no more than are necessary for this purpose? Balance demands both sufficiency and parsimony. Balance is the position one adopts to achieve stated objectives with a minimum use of resources, efforts, burden, etc.. According to the Oxford dictionary, balance refers to, among other things, "harmony of design and proportion" and the verb "to balance" is defined as "to bring into or to keep in equilibrium". Balance also gets into play in the adaptation of the research to the audience, study objectives and the research circumstances.⁴
- **Transparency:** does the study provide all the relevant background needed for an informed assessment? Science is based on the idea of reproduction: if the same observation is made repeatedly, the conclusion reached from these observations acquires more strength than conclusions drawn from idiosyncratic observations. Reproduction requires shared knowledge or intersubjectivity⁵. The corresponding

⁴ "[...] because of the diverse range of constituencies for whom social scientific applied research is designed and implemented, researchers are often compelled to adjust their objectives, methodology and style of presentation to conform to the needs, expectations, and methodological and evaluative criteria of different audiences. [...] In designing and implementing a particular research program, the applied social scientific researcher must continually balance the client's self-defined needs and objectives with the actual allocation of resources (financial, human, archival, time, etc.) toward the research process. By definition, the quality of each research study is determined, in large part, by the researcher's ability to creatively manage this tension between objectives and available resources in a way that most effectively serves the needs of the client" (Cohen and Shirley, 1998, 1-2)

⁵ "In saying that science is intersubjective, we mean that two scientists with different subjective orientations would arrive at the same conclusion if each conducted the same experiment. [...] If the earlier researchers had reported the design and execution of their studies in precise and specific details, however, and you were to replicate the study exactly, you should arrive at the same finding. This is what is meant by the intersubjectivity of science." (Babbie, 1990, 16–17)
 [...]il est évident qu'on ne peut pas s'attendre à ce que la science produise des vérités ou des connaissances absolues et irréfutables. Les résultats dits « scientifiques » ne sont vrais que temporairement, jusqu'à ce qu'une autre recherche ne les démontre suffisamment faux pour que la communauté scientifique les rejette et les remplace par des résultats plus fiables. [...] On ne peut jamais démontrer la vérité absolue d'un résultat. En effet, le caractère empirique (plutôt que logique) de la preuve requiert la vérification de tous les objets ou événements passés, présents et futurs relatifs à une certaine classe d'objets ou d'événements; cette condition, qui ne peut évidemment jamais être remplie, ni logiquement, ni pratiquement, circonscrit cette impasse qu'on ne peut éviter que par la convention. [...] L'intersubjectivité peut prendre un autre sens qui est beaucoup plus proche de la signification de l'observabilité et de la reproductibilité. L'intersubjectivité implique que toutes les étapes de la

fundamental requirement is the full disclosure of all elements of methods which may have a bearing on the quality of the observations and on the strength of the conclusions.⁶ We label this "transparency" very much along the lines of the demand made by people and the media on politicians to be "transparent" in their decision making. Transparency is also a basic management principle according to the Ordre des administrateurs agréés du Québec.⁷ Transparency is probably the most fundamental precept: without it, the other criteria have no meaning in terms of public communication of research results.⁸

Other criteria have been used to assess survey research, but we contend that they are subsumed within the four proposed here. They include:

- the compliance with professional and ethical standards (Hebert, 1999);
- the representativeness of the sample (Stern, 1979, 32 & 77; Corbin et al., 2000, 16);
- the validity of the operational definitions of concepts (Stern, 1979, 77; Corbin et al., 2000, 16);
- the logic of the design (Babbie, 1990, 41);
- the sufficient size of the sample (Corbin et al., 2000, 16);

méthode puissent être traduites en termes publics de sorte que tous les tests puissent potentiellement être menés par au moins deux scientifiques." (Mellos, 1997, 500–501)

⁶ "Basic criteria for measuring the social scientific integrity of the research include the following: transparency of the research methodology (in design, implementation and presentation)" (Cohen and Shirley, 1998, 7)

⁷ 2.2-1: « Transparence » : Qualité de ce qui laisse paraître la réalité tout entière, sans qu'elle ne soit altérée ou biaisée. Il n'est d'autre principe plus vertueux que la transparence de l'acte administratif par l'administrateur qui exerce un pouvoir au nom de celui de qui origine le pouvoir. Celui qui est investi d'un pouvoir doit rendre compte de ses actes à son auteur.

2.2-2: Essentiellement, on peut déterminer que l'administrateur doit rendre compte de son administration; que ce soit au mandant ou à une personne ou un groupe désigné par celui-ci, par exemple : à un conseil d'administration, à un comité de surveillance ou à un vérificateur.

2.2-3: Dans la mesure où le mandant le permet et qu'il n'en subit aucun préjudice, l'administrateur doit également agir de façon transparente envers les tiers ou les préposés pouvant être affectés par ses actes.

2.2-4: Ainsi, la transparence implique de rendre l'information accessible aux tiers, incluant les membres de son organisation afin d'assurer la saine gestion.

2.2-5: L'administrateur doit divulguer à son mandant tous ses intérêts propres, financiers ou personnels, de même que ceux de sa famille immédiate (telle que définie au paragraphe 4.4.13-5), qui puisse affecter son travail ou ses fonctions. (Ordre des administrateurs agréés du Québec, 1)

⁸ "A common theme running through all my remarks is that, as professionals, survey researchers have the obligation to provide as much information as possible about the work they do-to their clients, to respondents and others whose interests may be affected by the research, and, especially but not only when the research is in the public domain, to the general public. From a practical point of view, private interests, including the researcher's, the client's, and even of the respondents, will always place limits on this principle of openness. Nonetheless, I have no doubt that over the long run, and often even in the short run, the quality of the work we do is enhanced by being as open as possible." (Crespi, 1998)

- controlled implementation of the stimulus that the survey represents for the respondent (Stern, 1979, 80);
- the availability of the sampled individuals (Blais and Durand, 1997, 362);
- the capacity for individuals to respond (comprehension and information) (Blais and Durand, 1997, 362);
- the lack of bias of the questions (Corbin et al., 2000, 16);
- the unbiased communication of information from the part of the respondent (Blais and Durand, 1997, 362; Corbin et al., 2000, 16);
- the faithful recording of the information offered by the respondents (Blais and Durand, 1997, 362; Corbin et al., 2000, 16);
- the disclosure of interviewer instructions (Corbin et al., 2000, 16);
- the use of accepted statistical principles (Corbin et al., 2000, 16);
- the objectivity of the process (Corbin et al., 2000, 16);
- the adherence to a procedural quality model (Kaase, 1999, 240).

Survey research building blocks

The assessment framework is structured around the six building blocks of survey research. They are listed below, along with a brief definition.

- **Questionnaire:** the development of the verbal or visual stimuli used to solicit answers from respondents; this is obviously the basic tool used in survey research.
- **Sampling:** the selection of a subset of the population of subjects which is targeted for the study.
- **Data collection:** the application of the questionnaire to the elements of the sample.
- **Data management:** the transfer of responses provided by participants to electronic media as well as the editing of the data and the creation of new data; this includes devising weights.
- **Data analysis:** the exploitation of data to answer research questions; this includes the use of statistics and could include the use of designs

which are more sophisticated than simple one-survey descriptive research plans.

- **Reporting:** the presentation of the research process and the research findings, usually in the form of a written or graphical report.

Combining principles and building blocks

This assessment framework posits that "sound" survey research exhibits evidence of the four quality principles, within each of the six building blocks. The particularity of this assessment framework is to combine the six building blocks with the four assessment principles to produce a matrix of 24 zones (see Exhibit 1) which require attention in the review of survey research.⁹ Each zone comprises specific assessment criteria. Let's analyse what these criteria would be within this 6x4 framework. The following sections describe and justify the criteria summarized in Exhibit 1. The selection of most criteria reflect a large level of agreement in the literature concerning their importance; additional criteria are suggested by the author as deserving careful attention in assessing survey research.

⁹ This cross-classification of functions or phases with principles mirrors that used by the Ordre des administrateurs agréés du Québec in the establishment of its standards of sound management: in this case, six principles ("transparence, continuité, efficience, équilibre, équité, abnégation") are applied to five management functions (planning, organizing, directing, controlling, coordinating) to form the basis for the management standards. The idea of the cross-classification of phases and principles was actually borrowed from this source. See Breault (1999).

EXHIBIT 1 • Survey Assessment Framework

Principles				
	[R]igour the systematic application of best practices in the matter of survey research	[N]eutrality nature of what represents reality faithfully	[B]alance the sufficient yet parsimonious utilisation of resources toward the development of proof	[T]ransparency nature of what depicts reality completely and without alteration
Building blocks				
[q]uestionnaire	<ul style="list-style-type: none"> Content validity Pretests Response scales 	<ul style="list-style-type: none"> Reliability Absence of bias 	<ul style="list-style-type: none"> Minimisation of response burden 	<ul style="list-style-type: none"> Research objective Sponsor Complete reproduction of the questionnaire
[s]ampling	<ul style="list-style-type: none"> Adhesion to rules of random sampling Justification of non random sampling 	<ul style="list-style-type: none"> Target population and population reached; filtering procedures Sampling frame Final dispositions; response rate; refusal rate; replacement procedures Sampling margin of error 	<ul style="list-style-type: none"> Appropriateness of the nature of the sample considering the research purposes Sample size 	<ul style="list-style-type: none"> Population definition Sampling method Sample size Response rate
data [c]ollection	<ul style="list-style-type: none"> Interviewer training Quality controls, supervision Call back schedule 	<ul style="list-style-type: none"> Criterion validity (corroboration) Non-contamination Double-blindedness 	<ul style="list-style-type: none"> Justification of the type of survey Informed consent vs. hidden client Avoidance of scarring effects Confidentiality of data 	<ul style="list-style-type: none"> Data collection method Identification of data collection agent Dates, locations and periods of data collection Notable social events during the data collection period
data [m]anagement	<ul style="list-style-type: none"> Calculations and adjustments done correctly Coding performed rigorously 	<ul style="list-style-type: none"> Weighting criteria used Sources of population data Assurance that adjustments do not unduly favour the researcher's hypotheses or interests 	<ul style="list-style-type: none"> Avoidance of excessive weighting Effect of the adjustments made 	<ul style="list-style-type: none"> Mode of calculation of weights Variance of the weighting scheme Calculations and adjustments made
data [a]nalysis	<ul style="list-style-type: none"> Technically correct use of statistical tools For issues of causation, appropriate specification of the effect model 	<ul style="list-style-type: none"> Risk of disputes on results and interpretations Coherence of conclusions with results 	<ul style="list-style-type: none"> Appropriateness of methods considering the research purposes Appropriate of the research design considering the research purposes 	<ul style="list-style-type: none"> Raw and weighted sample sizes in tables Description of analysis methods used
[r]eporting	<ul style="list-style-type: none"> Disclosure of study weaknesses and possible bias 	<ul style="list-style-type: none"> Distinct presentation of objective results and of their interpretation Selflessness of the researcher 	<ul style="list-style-type: none"> Level of support to conclusions offered by study results Possible reservations on the internal and external validity of results 	<ul style="list-style-type: none"> Understandability of the presentation Disclosure of all elements of information required by all 24 framework cells to assess the value of results

[q]uestionnaire

Quality of questionnaire design is a controversial concept¹⁰ although the requirement for quality questionnaires goes undisputed.¹¹ We apply our four principles to the design of questionnaires. While the combinations may not produce all of the criteria which have been identified by other authors, they generate a coherent set of criteria which encompass key practitioners' concerns.

- **[q]uestionnaire [R]igour** qR combination

qR1 **Content validity**¹²

Content validity "depends on the extent to which an empirical measurement reflects a specific domain of content." (Carmines and Zeller, 1979, 20). A questionnaire that is developed with rigour displays content validity; reporting on the research may include statements on content validity.

qR2 **Pretests**¹³

Pretests include various activities which follow the initial development of the questionnaire and precede the utilization of the questionnaire to collect final data. Careful pretests can identify problems with

¹⁰ The quality concept has been developed much further, but there is a large variation between how organizations conceptualise quality. By referring to their concept as "Quality Profile", some organizations like the U.S. Survey of Income and Program Participation and the National Center for Education indicate that no quality component is conclusive by itself. The Statistics Canada "Quality Guidelines" include documentation and dissemination in their concept. The U.S. Bureau of Labour Statistics "Quality Measurement document" has a chapter on "Conceptualization". Statistics Sweden includes "Content, Time, and Availability" together with the traditional "Reliability" aspect. Eurostat has found it useful to include "Comparability, Coherence and Completeness" as main components. (Akkerboom and Lindstrøm, 1998, 2-3)

¹¹ "As a rule, neither sampling methods, interviewer organization nor analytical methods differentiate poor quality surveys from high quality survey research; the determining factor, and thus the decisive criterion for distinguishing quality, is the study concept and questionnaire development." (Petersen, 1998, 3)

¹² "One must first ensure that the questionnaire domains and elements established for the survey or poll fully and adequately cover the topics of interest. Ideally, multiple rather than single indicators or questions should be included for all key constructs." (AAPOR, 2)

¹³ "Structured questionnaires should be pretested among eligible respondents prior to the start of the main field period to establish the clarity, flow, and appropriateness of questions.." (CASRO, 1)
 "All questions should be pretested to ensure that questions are understood by respondents, can be properly administered by interviewers, and do not adversely affect survey cooperation." (AAPOR, 2)
 "Texts on survey research generally recommend pretests as a way to increase the likelihood that questions are clear and unambiguous." (Diamond, 2000, 248)

understanding of questions, flow of the questionnaire, interview duration, instructions, etc.

qR3 **Response scales**

Questionnaire comprise questions, obviously, but also a series of response scales. The quality of these scales at depicting the nuances of opinion or attitude and their lack of bias are important characteristics of rigorous research.

- **[q]uestionnaire [N]eutrality** qN combination

qN1 **Reliability**¹⁴

"Fundamentally, reliability concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials. This tendency toward consistency found in repeated measurements of the same phenomenon is referred to as reliability. The more consistent the results given by repeated measurements, the higher the reliability of the measuring procedure; conversely the less consistent the results, the lower the reliability [...] (Carmines and Zeller, 1979, 11-12). A questionnaire acts neutrally if its measurements are stable reflections of the phenomenon under study.

¹⁴

"Realistically however, existing (and preferably validated) scales should be adopted (or adapted) wherever possible in order to cultivate a cumulative tradition of research." (Grover, 1)
 Kaye lists several key questions relative to ascertaining the quality of statistical analyses. One of them is "Is the measurement process reliable?" (Kaye, 2000, 341)

qN2 **Absence of bias**¹⁵

A questionnaire is neutral if the wording of the questions and answer categories as well as the order of questions and answers does not predetermine, in part or completely, the answers given by the study participants.

• **[q]uestionnaire [B]alance**

qB combination

qB1 **Minimisation of response burden**¹⁶

Questionnaire balance is achieved through the minimisation of response burden. The task demanded of participants should be commensurate to the study objectives and to the benefits derived, by participants and society, through the research.

- 15 "The questions and questioning procedures are unbiased. The wording of the interview questions does not predetermine the answers to the research questions. The questioning and analytical procedures allow responses over the entire range relevant to the research objectives." (CASRO, 1)
 "Concepts should be clearly defined and questions unambiguously phrased. Question wording should be carefully examined for special sensitivity or bias." (AAPOR, 2)
 "According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Were the questions framed in a clear, precise, and nonleading manner?" (Hebert, 1999)
 "You might also need a survey expert to help you analyze the sequence of the questions in a questionnaire and whether the sequence is in itself leading and suggestive." (Hebert, 1999)
 "Although it seems obvious that questions on a survey should be clear and precise, phrasing questions to reach that goal is often difficult. Even questions that appear clear can convey unexpected meanings and ambiguities to potential respondents." (Diamond, 2000, 248)
 "The mode of questioning can influence the form that an order effect takes. In mail surveys, respondents are more likely to select the first choice offered (a primacy effect); in telephone surveys, respondents are more likely to choose the last choice offered (a recency effect). [...] To control for order effects, the order of the questions and the order of the response choices in a survey should be rotated." (Diamond, 2000, 255)
- 16 "The purpose of interviewing must be limited to the finding out of information or observation of reactions relevant to the research problem at hand." (PMRS, 1, section II)
 "Overly long questionnaires should be avoided at all costs." (PMRS, 1, rule 2.13)
 "The types of information to be collected are driven by the information objectives of the research." (CASRO, 1)
 "The questions are relevant and appropriate to the research issues. Interview content focuses on the research objectives." (CASRO, 1)
 "The interview should recognize the value of the respondent's time and the respondent's right to privacy. Lengthy interviews can be a burden. Length of the interview should be weighed against the needs of the research objectives, with consideration to the burden on respondents, the quality of the responses obtained, and the availability of alternative methods of obtaining the data (split samples, other sources, etc.)." (CASRO, 1)
 "[...] statisticians should collect only the data needed for the purpose of their inquiry." (ASA, 1)

- **[q]uestionnaire [T]ransparency** qT combination

qT1 **Research objective**¹⁷

Transparent questionnaire design demands that the research objectives be spelled out so that the value of the indicators can be assessed against the original measurement intentions.

qT2 **Sponsor**¹⁸

Reporting on questionnaire design should indicate who the research sponsor was.

qT3 **Complete reproduction of the questionnaire**¹⁹

To be transparent, questionnaire design must include the reproduction of the questionnaire *in toto*. This allows for the assessment of question wording as well as possible order effects. In the case of omnibus surveys — where several clients commission portions of a questionnaire — reporting should include a list of topics addressed prior

¹⁷ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...]; iii) the specific objectives of the study [...]." (PMRS, 1, rule 4.6)
 "The purpose of *clearly stating the research objectives* is to have complete agreement between the professional resaerch firm and the client [...]." (CASRO, 1)
 "The objectives of a high quality survey or poll should be specific, clear-cut and unambiguous. Such surveys are carried out solely to develop statistical information about the subject, not to produce predetermined results [...]." (AAPOR, 2)
 "The first step in ensuring quality in applied qualitative research is the process through which the research team assists the client in clarifying and articulating the scope and objectives of the research." (Cohen and Shirley, 1998, 7)
 "The report describing the results of a survey should include a statement describing the purpose or purposes of the survey." (Diamond, 2000, 236)
 "The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 1. the purpose of the survey; [...]." (Diamond, 2000, 270-271)
 "Critical in demonstrating adherence to standards of excellence and validity in defining the universe, designing the sample, learning what people think and do, and reporting the findings are: (1) a clean and explicit statement of objectives; [...]." (Sorensen, 352)

¹⁸ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required:[...]; ii) the name of the organization for which the study was conducted [...]." (PMRS, 1, rule 4.6)
 "Standard for Minimal Disclosure: Who sponsored the survey" (AAPOR, 1, rule III.1)

¹⁹ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: i) copy of the questionnaire [...]." (PMRS, 1, rule 4.6)
 "Standard for Minimal Disclosure: The exact wording of questions asked, including the text of any preceding instruction or explanation to the interviewer or respondents that might reasonably be expected to affect the response" (AAPOR, 1, rule III.2)
 "The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 5. the exact wording of the questions used, including a copy of each version of the actual questionnaire, interviewer instructions, and visual exhibits; [...]." (Diamond, 2000, 270-271)

to the relevant section and a statement concerning the possibility that the prior questions may have affected following answers.

[s]ampling

- *[s]ampling [R]igour* sR combination

sR1 ***Adhesion to rules of random sampling***²⁰

Only random sampling permits the statistically-based enunciation of factual statements about the population. Since most research aim to provide a rigorous depiction of their study topic, random sampling should be the norm and adherence to random sampling principles should be expected of rigorous research.

20

"Survey reporting guidelines: [...] Sample design, including method of selecting sample elements, qualifying/ disqualifying criteria. Method of selection within household" (CASRO, 1)

"A survey's intent is not to describe the particular individuals who, by chance, are part of the sample, but rather to obtain a composite profile of the population. In a bona fide survey, the sample is not selected haphazardly or only from persons who volunteer to participate. It is scientifically chosen so that each person in the population will have a measurable chance of selection. This way, the results can be reliably projected from the sample to the larger population with known levels of certainty/precision." (AAPOR, 2)

"Virtually all surveys taken seriously by social scientists, policy makers, and the informed media use some form of random or probability sampling, the methods of which are well grounded in statistical theory and the theory of probability." (AAPOR, 2)

"According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Did the expert select a representative sample of the universe to interview?" (Hebert, 1999)

"Probability sampling methods, in contrast, ideally are suited to avoid selection bias. Once the conceptual population is reduced to a tangible sampling frame, the units to be measured are selected by some kind of lottery that gives each unit in the sampling frame a known, nonzero probability of being chosen. Selection according to a table of random digits or the like leaves no room for selection bias." (Kaye, 2000, 345)

"The use of probability sampling techniques maximizes both the representativeness of the survey results and the ability to assess the accuracy of estimates obtained from the survey." (Diamond, 2000, 242)

"All opinion polls should be based on scientific and representative measurements of public opinion. Far too often the term opinion poll is misused to describe unscientific and unrepresentative measurements of public opinion. Representativeness means the obtaining of measurements which can be generalised to apply without any statistical bias to the whole population under consideration." (ESOMAR, 1)

sR2 Justification of non random sampling²¹

There are situations where random sampling is unfeasible but where the study topic is sufficiently potent to warrant departing from the previous principle. In such circumstances, the use of non random sampling is acceptable but must be justified.

21 "[...] the suitability of a specific sampling method for the research purpose will always involve some judgment [...]" (CASRO, 1)
"If a non-representative sample is to be used, possible implications and effects of the sampling method should be discussed." (CASRP, 1)
"Survey reporting guidelines: [...] Implications and limitations of non-representative sampling methods, if used." (CASRO, 1)
"The irrelevance of performance standards to professional ethics is obvious with respect to surveys conducted in countries, or parts of countries, with limited or inadequate survey resources or with life style characteristics that, for example, may make for low completion rates. Under conditions that prevail in such cases, it may prove to be very costly (even perhaps impossible) to implement satisfactorily a probability sample in conformity with state-of-the-art standards that would apply in the USA or Western Europe. Insistence on conformity to performance standards in such cases could result in cancelling a study, whereas some flexibility in the application of such standards might make it possible to conduct a survey whose results might still be usable. For this reason, even though I unequivocally advocate the theoretical superiority of probability sampling, I also assume that its standards should be adjustable to meet the realities of specific times and places." (Crespi, 1998)

- **[s]ampling [N]eutrality**

sN combination

sN1 **Target population and population reached; filtering procedures**²²

"In litigation, the survey's universe must fit the facts of the case" (Hebert, 1999). Obviously, this applies beyond the courts. The theoretical target population must correspond to the topic studied.²³ Moreover, the population reached must fit the theoretical population sought. When applicable, the filtering procedures used during the fieldwork must provide a precise empirical match between the people selected for the study and the individuals whose behaviour or attitudes are important to the study topic.

22

"For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: v) the universe covered (intended and actual) [...]." (PMRS, 1, rule 4.6)

"A universe which is relevant to the problem being studied, [is a] vital requirement of high quality research." (CASRO, 1)

"Survey reporting guidelines: [...] Definition of the universe which the survey is intended to represent [...] Respondent qualification requirements." (CASRO, 1)

"Standard for Minimal Disclosure: A definition of the population under study." (AAPOR, 1, rule III.3)

"Standard for Minimal Disclosure: information on eligibility criteria and screening procedures." (AAPOR, 1, rule III.5)

"According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Was the survey universe properly determined?" (Hebert, 1999)

"The target population consists of all elements (i.e., objects, individuals, or other social units) whose characteristics or perceptions the survey is intended to represent. [...] The definition of the relevant population is crucial because there may be systematic differences in the responses of members of the population and nonmembers.[...] The universe must be defined carefully. [...] The survey report should contain a description of the target population, a description of the survey population actually sampled, a discussion of the difference between the two populations, and an evaluation of the likely consequences of that difference." (Diamond, 2000, 239-240)

"In a carefully executed survey, each potential respondent is questioned or measured on the attributes that determine his or her eligibility to participate in the survey. [...] The criteria for determining whether to include a potential respondent in the survey should be objective and clearly conveyed, preferably using written instructions addressed to those who administer the screening questions. These instructions and the completed screening questionnaire should be made available to the court and the opposing party along with the interview form for each respondent." (Diamond, 2000, 247-248)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail [...] 3. a description of the sample design, including the method of selecting respondents, the method of interview, the number of callbacks, respondent eligibility or screening criteria, and other pertinent information; [...]." (Diamond, 2000, 270-271)

"If a survey is to be relevant, care must be taken in defining the universe — the population about whose behavior generalization will be made from survey results, which is, therefore, the group from whom individuals are selected for interview." (Sorensen, 353).

23

"To use a common example, in litigation involving plaintiff's claims that defendant's deceptive advertising destroyed the market for plaintiff's product, the universe consists of those consumers who intend to buy from the plaintiff or defendant in the future. The consumer who belongs to this universe should be the (1) potential purchaser; (2) potential decision maker; and (3) person to whom the advertising is addressed." (Hebert, 1999)

sN2 **Sampling frame**²⁴

The sampling frame is the list²⁵ from which the sample is drawn. The quality of the sampling frame is an important determinant of the neutrality of the sampling process since an incomplete, ill-focussed or otherwise biased frame would reduce the ability to infer results to the target population.

²⁴ "[...] a sample which adequately represents that universe, [is a] vital requirement of high quality research." (CASRO, 1)
"The source of the sample (lists, on-line groups or services, randomly generated phone numbers, mall intercepts, etc.) should be revealed and the adequacy of the source given the study purpose should be discussed." (CASRO, 1)
"Survey reporting guidelines: [...] Definition of the sampling frame, i.e., sampling points actually used and the procedures used in selecting sampling points. If lists used, source and name of list." (CASRO, 1)
"Standard for Minimal Disclosure: a description of the sampling frame used to identify this population." (AAPOR, 1, rule III.3)
"Critical elements in an exemplary survey are: (a) to ensure that the right population is indeed being sampled (to address the questions of interest); and (b) to locate (or "cover") all members of the population being studied so they have a chance to be sampled. The quality of the list of such members (the "sampling frame") whether it is up-to-date and complete is probably the dominant feature for ensuring adequate coverage of the desired population to be surveyed." (AAPOR, 2)
"At the minimum, any IS survey research should describe and justify the sample frame . Estimation of possible frame error bias (or lack thereof) by a comparison estimation of the probability of the target population being included in or excluded from the sample frame is desirable." (Grover, 2)

²⁵ Sometimes, the frame is a list of clusters of ultimate sampling units.

sN3 ***Final dispositions; response rate; refusal rate; replacement procedures***²⁶

The neutrality of the sample can be assessed, in part, by the results obtained in attempting to contact sampling units. The response rate and the refusal rate are key indicators of neutrality. Sample replacement procedures should be assessed carefully to identify possible bias induced by discarding original sampling units and adding new units.

26

"For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: vii) the contact record based on the last attempt to obtain an interview with the exception of mall surveys and quota samples where it is not appropriate [...]." (PMRS, 1, rule 4.6)

"Callback and replacement procedures, if used, should be described." (CASRO, 1)

"The final disposition of the sample should be described in detail [...]." (CSARO, 1)

"A low cooperation or response rate does more damage in rendering a survey's results questionable than a small sample, because there may be no valid way scientifically of inferring the characteristics of the population represented by the nonrespondents." (AAPOR, 2)

"When objects like receipts (for an audit) or vegetation (for a study of the ecology of a region) are sampled, all can be examined. Human beings are more troublesome. Some may refuse to respond, and the survey should report the nonresponse rate. A large nonresponse rate warns of bias, but it does not necessarily demonstrate bias." (Kaye, 2000, 345)

One very stringent interpretation of response rate levels is reported by Diamond (2000, 245): "One suggested formula for quantifying a tolerable level of nonresponse in a probability sample is based on the guidelines for statistical surveys issued by the former U.S. Office of Statistical Standards. According to these guidelines, response rates of 90% or more are reliable and generally can be treated as random samples of the overall population. Response rates between 75% and 90% usually yield reliable results, but the researcher should conduct some check on the representativeness of the sample. Potential bias should receive greater scrutiny when the response rate drops below 75%. If the response rate drops below 50%, the survey should be regarded with significant caution as a basis for precise quantitative statements about the population from which the sample was drawn."

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 4. a description of the results of sample implementation, including (a) the number of potential respondents contacted, (b) the number not reached, (c) the number of refusals, (d) the number of incomplete interviews or terminations, (e) the number of noneligibles, and (f) the number of completed interviews; [...]." (Diamond, 2000, 270-271)

"Regardless of how the sample is drawn, it is good practice to disclose in easily comprehensible form how the sampling procedure has been implemented in the field. This includes the — albeit costly — best practice of providing a precise nonresponse account (number and description of contacts, reasons for nonresponse, etc.) as well as detailed information on the conduct of interviews (number of necessary contacts, date and time of day, difficulties in accessing target individuals, etc.)." (Kaase, 1999, 183)

"When people increasingly refuse access to interviewers, costs rise accordingly and a decreasing interview completion rate poses a fundamental question of validity: what would those people not interviewed have said about their attitudes and perceptions if they had been interviewed, and would their answers have differed significantly from those of people who were interviewed?" (Sorensen, 362)

sN4 **Sampling margin of error**²⁷

The neutrality of the sample is represented, in part, by the width of the sampling margin of error. A wide margin of error may hint at a non neutral sample.

• **[s]sampling [B]alance**

sB combination

sB1 **Appropriateness of the nature of the sample considering the research purposes**²⁸

A sample features balance if the nature of the sampling procedure is in sync with the objective of the research. For example, descriptive studies should emphasize the representativity of the sample while comparative or correlational studies should focus on the ability to compare groups in a valid fashion (see later concepts of internal and external validity).

sB2 **Sample size**²⁹

Ceateris paribus, the size of the sample determines the ability of the study to uncover statistically significant differences between subgroups of the sample and it is one of the determinants of the statistical accuracy of the estimates produced. Therefore, the sample should be large enough to support the claims of the study but not so large as to provoke type I errors (or the erroneous rejection of the null hypothesis). Every survey also contributes to the overall social burden represented by survey research; therefore any one survey should not acquire data from substantially more respondents than necessary.

²⁷ "Survey reporting guidelines: [...] If statements are made regarding the overall sampling error of the survey, it should be stated that total survey error includes both sampling error and response error." (CASRO, 1)

"Standard for Minimal Disclosure: A discussion of the precision of the findings, including, if appropriate, estimates of sampling error" (AAPOR, 1, rule III.6)

"Sampling errors should be included for all statistics presented, rather than only the statistics themselves." (AAPOR, 2)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 7. estimates of the sampling error, where appropriate (i.e., in probability samples);[...]" (Diamond, 2000, 270-271)

²⁸ "Basic criteria for measuring the social scientific integrity of the research include the following: [...] the scale, balance and consistency of the research sample" (Cohen and Shirley, 1998, 7)

²⁹ "The sample size should also be specified, and there should be some discussion of its appropriateness, considering the purpose of the study." (CASRO, 1)

"The sample must be large enough to allow the expert to extrapolate his or her findings to the universe. Experts will testify that, in general, the larger the sample size, the more reliable the survey results. On the other hand, a survey with a sample size of 5000 could have useless results if it uses improper questions or picks the wrong universe." (Hebert, 1999)

- **[s]sampling [T]ransparency** sT combination

sT1 **Population definition**³⁰

Sample transparency demands that the population be clearly defined in reporting. This is a transparency requirement as well as a necessity to be able to assess the population targeted and reached (see sN — sample neutrality).

sT2 **Sampling method**³¹

The sampling method must be spelled out in reporting for sample transparency to be achieved. It is also required to assess the appropriateness of the sampling procedures (see sB — sampling balance).

³⁰ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: v) the universe covered (intended and actual) [...]." (PMRS, 1, rule 4.6)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail [...] 2. a definition of the target population and a description of the population that was actually sampled; [...]." (Diamond, 2000, 270-271)

"The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - a description of the intended and actual universe covered [...]." (ESOMAR,1)

³¹ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: v) [...] details of the sampling method and selection procedures; [...] viii) the method of recruitment when prior recruitment of respondents is undertaken [...]." (PMRS, 1, rule 4.6)

"The following guidelines are [...] aimed at insuring that sample design and management are disclosed in sufficient detail to allow clear judgments of a sample's adequacy for the stated research purpose." (CASRO, 1)

"The description of the sampling plan should include the criteria by which a given sample element (i.e., an individual consumer, household, business, etc.) is selected to be in the sample." (CASRO, 1)

"Standard for Minimal Disclosure: A description of the sample selection procedure, giving a clear indication of the method by which the respondents were selected by the researcher, or whether the respondents were entirely self-selected." (AAPOR, 1, rule III.4)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail [...] 3. a description of the sample design, including the method of selecting respondents, the method of interview, the number of callbacks, respondent eligibility or screening criteria, and other pertinent information; [...]." (Diamond, 2000, 270-271)

"The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - [...] the size, nature and geographical distribution of the sample (both planned and achieved); and where relevant, the extent to which any of the data collected were obtained from only part of the sample - [...]." (ESOMAR,1)

"Regardless of how the sample is drawn, it is good practice to disclose in easily comprehensible form how the sampling procedure has been implemented in the field." (Kaase, 1999, 183)

sT3 Sample size³²

The size of the final, unweighted sample must be indicated in reporting. The size of important sub-samples used in analysis and reporting should also be indicated clearly.

sT4 Response rate³³

The response rate itself serves as an indicator of sample neutrality. The transparency concern is that the response rate value be clearly presented in reports on the study and that it be calculated according to agreed upon formulas.³⁴

³² "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: vi) the size and nature of the sample [...]" (PMRS, 1, rule 4.6)

"The sample size should also be specified [...]" (CASRO, 1)

"Survey reporting guidelines: [...] Sample size." (CASRO, 1)

"Standard for Minimal Disclosure: Size of samples" (AAPOR, 1, rule D.5)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 4. a description of the results of sample implementation, including (a) the number of potential respondents contacted, (b) the number not reached, (c) the number of refusals, (d) the number of incomplete interviews or terminations, (e) the number of noneligibles, and (f) the number of completed interviews; [...]" (Diamond, 2000, 270-271)

"The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - [...] - the size, nature and geographical distribution of the sample (both planned and achieved); and where relevant, the extent to which any of the data collected were obtained from only part of the sample - [...]" (ESOMAR,1)

³³ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] xi) a statement of response rates, how they were calculated, and a discussion of possible bias due to non-response [...]" (PMRS, 1, rule 4.6)

"The final disposition of the sample should be described in detail, as should any completion rate or incidence rate calculations." (CSARO, 1)

"Standard for Minimal Disclosure: if applicable, completion rates" (AAPOR, 1, rule D.5)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 4. a description of the results of sample implementation, including (a) the number of potential respondents contacted, (b) the number not reached, (c) the number of refusals, (d) the number of incomplete interviews or terminations, (e) the number of noneligibles, and (f) the number of completed interviews; [...]" (Diamond, 2000, 270-271)

"The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - [...] where technically relevant, a statement of response rates and a discussion of any possible bias due to non-response." (ESOMAR,1)

"Regardless of how the sample is drawn, it is good practice to disclose in easily comprehensible form how the sampling procedure has been implemented in the field. This includes the — albeit costly — best practice of providing a precise nonresponse account (number and description of contacts, reasons for nonresponse, etc.) as well as detailed information on the conduct of interviews (number of necessary contacts, date and time of day, difficulties in accessing target individuals, etc.)." (Kaase, 1999, 183)

³⁴ There is more than one way to measure response rates. Professional organizations propose various standard formulas. The key is to report which formula was used and to abide by the rules of that formula.

data [c]ollection

- **data [c]ollection [R]igour**

cR combination

cR1 **Interviewer training**³⁵

While recognizing the importance of interviewers in the survey process, Kaase (1999, 195-196) minimises the importance of standardization in this area: "[...] the recruiting and training of interviewers deserve particular attention from the quality point of view. [...] Although the literature fully recognises the key role of the interviewer in the survey process [...], no scientifically grounded quality criteria can be identified that go beyond those mentioned in point 7³⁶ of the AAPOR code." We suggest that rigorous data collection require that interviewer training be commensurate to the complexity of the task and that the demonstration of the extent and nature of interviewer training be an indicator of quality.

35

"For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: ix) the method of field briefing sessions [...]" (PMRS, 1, rule 4.6)

"A properly trained interviewer is one who has been instructed in general interviewing techniques and who has been briefed on the particular project." (CASRO, 1)

"Train interviewers carefully on interviewing techniques and the subject matter of the survey." (AAPOR, 2)

"According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Did the interviewers use sound interview procedures and did they lack knowledge of the purpose of the survey?" (Hebert, 1999)

"Interviewers should be trained in delivering probes to maintain a professional and neutral relationship with the respondent (as they should during the rest of the interview), which minimizes any sense of passing judgment on the content of the answers offered. Moreover, interviewers should be given explicit instructions on when to probe, so that probes are administered consistently." (Diamond, 2000, 254)

"Properly trained interviewers receive detailed written instructions on everything they are to say to respondents, any stimulus materials they are to use in the survey, and how they are to complete the interview form. These instructions should be made available to the opposing party and to the trier of fact. [...] Interviewers require training to ensure that they are able to follow directions in administering the survey questions. [...] The more complicated the survey instrument is, the more training and experience the interviewers require." (Diamond, 2000, 264-265)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 5. the exact wording of the questions used, including a copy of each version of the actual questionnaire, interviewer instructions, and visual exhibits; [...]" (Diamond, 2000, 270-271)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 9. copies of interviewer instructions, validation results, and code books [...]" (Diamond, 2000, 270-271)

"Interviewers, being human beings, require careful, programmed instructions for respondent selection and interview procedures." (Sorensen, 359)

36

"Good interviewer techniques should be stressed, such as how to make initial contacts, how to deal with reluctant respondents, how to conduct interviews in a professional manner, and how to avoid influencing or biasing responses."

cR2 Quality controls, supervision³⁷

Rigorous data collection requires controlled procedures. The extent and quality of these controls determine the procedural robustness of the exercise — very much in an ISO-9000 way. While all data collection approaches demand quality controls, telephone interviewing has been seen as a way to make "data available under better controlled conditions" (Kaase, 1999, 157). This assertion may be put to the test soon with the joint use of computer-assisted interviewing, Internet-based shared data bases and low-cost large-bandwidth communications — which have fuelled a new trend toward centralised data bases accessed for telephone interviewing from home.

cR3 Call back schedule³⁸

Part and parcel of the achievement of high response rates, call back procedures must open communications with the potential respondents at a variety of times when they may be available.

- ³⁷ "The practitioner must automatically verify or monitor a minimum of 10% of each interviewer's completed interviews unless it is specifically made clear that this practice will not be followed." (PMRS, 1, rule 4.5)
 "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] xv) an adequate description of verification or monitoring procedures and results of the same [...]." (PMRS, 1, rule 4.6)
 "Standard industry practice requires that the data collection agency validate 15% of the interviews, using validation questions supplied by the researcher." (CASRO, 1)
 "Controlling the quality of fieldwork is done by observing/monitoring, verifying and/or redoing a small sample of the interviews." (AAPOR, 2)
 "Judging the adequacy of data collection may involve examining the process by which measurements are recorded and preserved." (Kaye, 2000, 342)
 "Three methods are used to ensure that the survey instrument was implemented in an unbiased fashion and according to instructions. The first, monitoring the interviews as they occur, is done most easily when telephone surveys are used. [...] Second, validation of interviews occurs when respondents in a sample are recontacted to ask whether the initial interviews took place and to determine whether the respondents were qualified to participate in the survey. [...] A third way to verify that the interviews were conducted properly is to compare the work done by each individual interviewer." (Diamond, 2000, 267)
- ³⁸ "Callback and replacement procedures, if used, should be described." (CASRO, 1)
 "Survey reporting guidelines: [...] number of callbacks" (CASRO, 1)
 "The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail [...] 3. a description of the sample design, including the method of selecting respondents, the method of interview, the number of callbacks, respondent eligibility or screening criteria, and other pertinent information; [...]" (Diamond, 2000, 270-271)
 "The probability sample defeats itself if, in each household, one or more of the following things happens: [...] (3) insufficient effort is made to call back and identify — or otherwise statistically compensate for — a respondent in an assigned household". (Sorensen, 356).

- **data [c]ollection [N]eutrality** cN combination

cN1 **Criterion validity (corroboration)**³⁹

"Criterion-related validity is defined as the correlation between a measure and some criterion variable of interest" (Bohrnstedt, 1983, 97). A demonstration via a single study is not as robust as a demonstration which builds upon past research. In particular, where data collected in one study can be corroborated by other data which are considered representative of the reality under study, the demonstration of data collection neutrality is reinforced.

cN2 **Non-contamination**

Durand and Blais (1997, 172) define non-contamination as follows: "The interviewer and the questionnaire, the observer and the observation grid are not part of daily life. Their simple presence may provoke specific effects, quite distinct from those we want to measure. One refers to this as instrument contamination effect. What is observed then is different from what would have happened without the presence of the instrument — the instrument influencing study subjects into modifying their behaviour." (translation by the author) A contaminating research instrument or protocol alters reality and, therefore, does not pass the data collection neutrality test.

³⁹

"Criterion-related validity or predictive validity check must also be performed, and refers to the ability of the scale to predict (or at least relate to) one or more external variables." (Grover, 1)

Kaye lists several key questions relative to ascertaining the quality of statistical analyses. One of them is "Is the measurement process valid?" (Kaye, 2000, 342). Kaye continues "Reliability is necessary, but not sufficient to ensure accuracy. In addition, to reliability, validity is needed. [...] When an independent and highly accurate way of measuring the variable of interest is available, it may be used to validate the measuring system in question."

cN3 **Double-blindedness**⁴⁰

Double-blindedness means that neither interviewers, nor respondents are aware of the interest (and, sometimes, the identity) of the study sponsor — and, therefore, cannot adjust, consciously or not, their behaviour to their perceptions of expected answers. This is a controversial criterion of data collection neutrality. While several sources identify double-blindedness as a requirement, consideration must be given to its impacts on the ability of interviewers to conduct an intelligent interview and on the likelihood of respondent refusal if faced with an unknown sponsor. Also, ensuring double-blindedness is sometimes difficult as sponsor identity may be revealed by some questions or probes. This author belongs to the sceptical group of researchers on this criterion but inserts it in the assessment framework nonetheless because of the frequency at which it is mentioned in the literature.

- **data [c]ollection [B]alance** cB combination

cB1 **Justification of the type of survey**

The type of survey conducted must be in sync with the study objectives. Surveys can be administered in various ways (face-to-face, telephone, mail, Web, etc.) and can take a number of forms (one-time, time-series, correlational, comparison groups, etc.). It is important that the type of survey conducted provides appropriate support to the analysis.

40

"The survey should be double-blind. The people who implement the survey should have no involvement in its design and the interviewers and their supervisors should not know the purpose of the survey. If they know who is sponsoring the survey, or what the survey is supposed to test, it is possible that they will unconsciously or consciously skew the results by the way they ask and record the questions. For the same reasons, the survey expert should not implement the survey, other than to ensure the accuracy of the data collection and to analyze the data." (Hebert, 1999)

"[...] any potential bias is minimized by having interviewers and respondents blind to the purpose and sponsorship of the survey and by excluding attorneys from any part in conducting interviews and tabulating results." (Diamond, 2000, 238)

"To ensure objectivity in the administration of the survey, it is standard interview practice to conduct double-blind research whenever possible: both the interviewer and the respondent are blind to the sponsor of the survey and its purpose. [...] Nonetheless, in some surveys (e.g., some government surveys), disclosure of the survey's sponsor to respondents (and thus to interviewers) is required. Such surveys call for an evaluation of the likely biases introduced by interviewer or respondent awareness of the survey's sponsorship. In evaluating the consequences of sponsorship awareness, it is important to consider (1) whether the sponsor has views and expectations that are apparent and (2) whether awareness is confined to the interviewers or involves the respondents." (Diamond, 2000, 266)

"The law and my professional judgment both come down in favor of the interviewers in the field being told absolutely nothing about the purpose of the interview, the identity of the client, or the fact that their interviews may be used in litigation." (Sorensen, 361)

cB2 Informed consent vs. hidden client⁴¹

Balanced data collection is based on the prerequisite that individuals accept to participate knowing what the task is and being reasonably reassured about the use of the information they are about to offer. There is a tension between this ethical requirement and the possible impact of divulging information on the sponsor or the topic of the study. Balanced data collection displays an adequate equilibrium in this regard.

cB3 Avoidance of scarring effects⁴²

Balanced data collection errs on the side of prudence with regard to the possible consequences that participation in a study may have on participants. Of particular concern are instances where participants are brought to declare illegal or socially unacceptable behaviour.

- 41 "You will always be told the name of the person contacting you, the research company's name and the nature of the survey" (Council for Marketing and Opinion Research, 1). Most important here is the last requirement that the participant must be told about the survey exercise they are about to embark.
- "Upon request interviewers will provide the Respondent with the name, address, and phone number of the head office of the organization on whose behalf the interviewer is conducting the survey research as well as the CSRC registration number and toll free telephone number." (Canadian Survey Research Council, 1)
- "IRBs should consider, in other words, the impact of informed consent procedures on the data collection objectives of the research. In the normal survey which presents minimal risk, detailed information about the objectives of the survey and the questions to be asked is apt to bias respondent participation without safeguarding respondent rights. In these surveys, the usual practice of a short introduction about the purpose of the study, the sponsor and the topics to be covered is sufficient. This statement should include the instruction that responses will be held in confidence and that questions the respondent does not want to answer can be skipped. More detailed methods of informing respondents may be considered when survey participation does pose substantial risk. The key here is to provide necessary information for informed choice without dramatic increases in nonresponse or response error, which can render survey efforts useless." (AAPOR, 4)
- [...] statisticians should [...] inform each potential respondent about the general nature and sponsorship of the inquiry and the intended uses of the data." (ASA, 1)
- "We at the American Statistical Association recommend strongly that everyone cooperate with surveys, if those who sponsor them can: [...] provide you with information so that you can make an informed decision about whether or not to participate." (ASA, 2)
- "The screening segmentation questions must not reveal to the respondent the purpose of the survey or seem to convey some impression about the interview's concern" (Sorensen, 357).
- 42 "No procedure or technique shall be used in which the respondent is put in such a position that he or she cannot exercise the right to withdraw or refuse to answer at any stage during or after the interview." (PMRS, 1, rule 2.3)
- "Questions or procedures that might put respondents "at risk," by asking confidential, disturbing, or threatening information should only be included where directly necessary to the research issues, and techniques should be used to minimize discomfort, concerns about security, apprehension, and/or misreporting." (CASRO, 1)
- "We shall strive to avoid the use of practices or methods that may harm, humiliate, or seriously mislead survey respondents." (AAPOR, 1, rule D.1)
- "Survey participation can, however, put respondents at significant risk when, for example, the inquiry concerns stigmatizing or illegal activity and inadequate attention is paid to ensuring respondent anonymity and the confidentiality of responses." (AAPOR, 4)
- The document provides an in-depth discussion of the possible scarring effects.
- "To some degree, there is an element of self-interest in our obligations to respondents. For example, if we do not respect the privacy and confidentiality of what respondents tell us, eventually we will lose the public's cooperation. Beyond self-interest, this obligation also rests on the proposition that we do not have the right to trick anyone into acting against his or her self-interest." (Crespi, 1998)

cB4 Confidentiality of data⁴³

Unless otherwise stated clearly, there is an implicit contract between the researcher and the participant that the participant's responses are to be used only for the purpose of the research and without connection to the identity of the respondent. Balanced data collection must respect this contract.

- **data [c]ollection [T]ransparency** cT combination

cT1 Data collection method⁴⁴

Transparent reporting of the data collection phase of the study requires that the mode of data collection be clearly established.

cT2 Identification of the data collection agent⁴⁵

The assessment of the research may be based in part on the track record of the data collection agent. Therefore, the data collection agent identity must be disclosed as part of data collection reporting.

- ⁴³ "Confidentiality of respondent data will be maintained. Individual respondents' data will not be provided to clients or other third parties without the permission of the respondent." (Canadian Survey Research Council, 1)
 "The identity of individual respondents must not be revealed by the practitioner to the client or anyone other than persons belonging to the organization of the practitioner concerned [...]." (PMRS, 1, rule 2.5) Exceptions are listed.
 "At every stage in the design, development, execution, and reporting of the research, procedures should insure the confidentiality and security of data provided by respondents, or by clients." (CASRO, 1)
 "Unless the respondent waives confidentiality for specified uses, we shall hold as privileged and confidential all information that might identify a respondent with his or her responses. We shall also not disclose or use the names of respondents for non-research purposes unless the respondents grant us permission to do so." (AAPOR, 1, rule D.2)
 "Exemplary survey research practice requires that one literally do "whatever is possible" to protect the privacy of research participants and to keep collected information they provide confidential or anonymous." (AAPOR, 2)
 [...] statisticians should [...] establish their intentions, where pertinent, to protect the confidentiality of information collected from respondents, strive to ensure that these intentions realistically reflect their ability to do so, and clearly state pledges of confidentiality and their limitations to the respondents; ensure that the means are adequate to protect confidentiality to the extent pledged or intended, that processing and use of data conform with the pledges made, that appropriate care is taken with directly identifying information (using such steps as destroying this type of information or removing it from the file when it is no longer needed for the inquiry), that appropriate techniques are applied to control statistical disclosure; ensure that, whenever data are transferred to other persons or organizations, this transfer conforms with the established confidentiality pledges, and require written assurance from the recipients of the data that the measures employed to protect confidentiality will be at least equal to those originally pledged." (ASA, 1)
- ⁴⁴ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] xii) the method by which the information was collected (e.g., mail, intercept, telephone) [...]." (PMRS, 1, rule 4.6)
- ⁴⁵ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] ii) [...] name of the organization conducting it, including sub-contractors [...]." (PMRS, 1, rule 4.6)
 "Standard for Minimal Disclosure: Who sponsored the survey, and who conducted it." (AAPOR, 1, rule III.1)

cT3 *Dates, locations and periods of data collection*⁴⁶

Minimal reporting requirements include the identification of the timing and location of data collection as well as, eventually, the stages of data collection which occurred. This contextual information is important to the assessment of the representativity and generalisability of the data.

cT4 *Notable social events during the data collection period*

Survey research studies are generally meant to be generalisable beyond the period when the data collection took place. One prerequisite for this to occur is that no significant social⁴⁷ event took place during the data collection which could severely influence the responses collected. Such events (or their absence) should be reported for more transparency.

data [m]anagement

- **data [m]anagement [R]igour** mR combination

mR1 *Calculations and adjustments done correctly*⁴⁸

The management of survey data often requires that calculations be performed and that data adjustments⁴⁹ be conducted. Such calculations and adjustments must be performed according to best practices at the time they were performed.

⁴⁶ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] iv) the dates on or between which the fieldwork was done and the time periods of interviewing [...]." (PMRS, 1, rule 4.6)

"Survey reporting guidelines: [...] Where interviewing was conducted (localities, or national if national). Dates interviewing conducted." (CASRO, 1)

"Standard for Minimal Disclosure: Method, location, and dates of data collection." (AAPOR, 1, rule III.8)

⁴⁷ We focus here on social events, but other types of events (political, economic, technological, etc.) may also have a bearing on data collection results.

⁴⁸ "Cleaning specifications should be written for all information collected, including information collected by computer assisted interviewing systems and on-line or interactive interviews." (CASRO, 1)

"Special codes should be provided for missing items, indicating why the data are not included. And, ideally, the "filling in" or imputation of these missing data items (based on rigorous and well validated statistical methods) should be undertaken to reduce any biases arising from their absence." (AAPOR, 2)

⁴⁹ The literature often use the term "data editing". See, among others, Kovar (1997), Nordbotten (2000), Blum (1997), Whitridge and Bernier (1999), K-BASE (1).

mR2 Coding performed rigorously⁵⁰

Coding is the delicate qualitative operation of collapsing answers to open-end questions into meaningful categories. Rigorous coding requires that the categories depict the original data with accuracy (if not completeness, by definition) and that the codification protocols be clear. Coding of the same material by two independent coders is one way to demonstrate rigour in coding.

- **data [m]anagement [N]eutrality** mN combination

mN1 Weighting criteria used⁵¹

Survey research often require application of *ex post facto* weights to compensate *a priori* stratification or unequal participation rates. Which variables are taken into consideration by these weights is one aspect of data management neutrality since weights are meant to bring sample data closer to the reality of the population.

mN2 Sources of population data⁵²

Similarly, quality research uses quality source of population data to compute weights. It is recognized, however, that recency and quality

- ⁵⁰ "Coders should not individually and independently set codes. A senior person should review codes." (CASRO, 1)
 "Because the interviews might result in ambiguous responses, the expert should be able to show that he or she had a written system for categorizing responses. The expert should be able to show that he or she can rationally account for the ambiguous responses in the survey results, and how; or that he or she threw out these responses, and why; and how these responses affected the survey results." (Hebert, 1999)
 "Coding of answers to open-ended questions requires a detailed set of instructions so that decision standards are clear and responses can be scored consistently and accurately. Two trained coders should independently score the same responses to check for the level of consistency in classifying responses." (Diamond, 2000, 268)
 "The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 6. a description of any special scoring (e.g., grouping of verbatim responses into broader categories);[...]" (Diamond, 2000, 270-271)
- ⁵¹ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] vi [...] details of any weighting methods used [...]" (PMRS, 1, rule 4.6)
 "If the sampling procedure requires that the resulting sample be weighted, the objectives of the weighting [...] should be specified." (CASRO, 1)
 "Standard for Minimal Disclosure: description of any weighting or estimating procedures used." (AAPOR, 1, rule III.6)
- ⁵² "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] xiv) list of sources of secondary research [...]" (PMRS, 1, rule 4.6)
 "The practitioner should provide to the client in the report, or in a supporting document, in addition to the items listed in 4.6, the following information: [...] ii) an assessment of the reliability of the sources used in secondary research [...]" (PMRS, 1, rule 4.15)
 "If the sampling procedure requires that the resulting sample be weighted, [...] the sources of the weights should be specified." (CASRO, 1)

don't always co-exist; sometimes, the researcher must choose between robust old data (e.g., census data) and up-to-date more feeble data (e.g., projections from census data).

mN3 ***Assurance that adjustments do not unduly favour the researcher's hypotheses or interests***⁵³

The various adjustments made via editing, weighting and otherwise should not unduly favour the position preferred by the researcher or by the study sponsor; excessive adjustments which would support the researcher's or the sponsor's thesis, even if they appear reasonable, would not respect the principle of data management neutrality.

- ***data [m]anagement [B]alance*** mB combination

mB1 ***Avoidance of excessive weighting***

Data weighting is normal practice in survey research. However, excessive⁵⁴ weighting where the answers of some individuals are given a value far beyond that given to some other individuals are to be avoided to achieve balance in data management. Sometimes, it is better to exclude some variables from a weighting scheme in order to reduce weighting variance, even at the cost of somewhat weaker representation of the population.

mB2 ***Effect of the adjustments made***

The overall effect of adjustments, editing and weighting should not be so large as to denature the data collected.

⁵³ "If data are weighted, weights should be fairly and consistently applied." (CASRO, 1)

"Editing should be done before data are to be entered and should remove all illegible, incomplete and inconsistent interviewer errors. If correct responses are not obvious from the questionnaire, then responses should be coded as no answer or the questionnaires should either not be used or returned to data collection." (CASRO, 1)

⁵⁴ We are not aware of an objective threshold beyond which weighting becomes excessive. This has to be determined on a case by case basis.

- **data [m]anagement [T]ransparency** mT combination

mT1 **Mode of calculation of weights**⁵⁵

The protocol used to calculate the weights should be clearly reported.

mT2 **Variance of the weighting scheme**⁵⁶

The variance of the weighting scheme affects the quality of the sample. A perfect simple random sample does not require weighting and it constitutes the basis for calculation of inferential statistics. The variance of the weighting scheme provides an indication of the level of design effect that should be accounted for in study statistics (Spencer, 2000).

mT3 **Calculations and adjustments made**⁵⁷

Transparent data management demands that calculations and adjustments be described in enough detail that the assessor can judge whether they were warranted and well executed.

data [a]nalysis

- **data [a]nalysis [R]igour** aR combination

aR1 **Technically correct use of statistical tools**⁵⁸

An obvious requirement of rigorous data analysis is that statistical tools be used correctly and that their underlying assumptions be not too severely broken.

⁵⁵ "The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - [...] and any weighting methods used - where technically relevant, a statement of response rates and a discussion of any possible bias due to non-response." (ESOMAR,1)

⁵⁶ "The effect of the weighting on sampling error should be disclosed." (CASRO, 1)

⁵⁷ "Force cleaning, which is done by having the computer change all answers meeting certain criteria, should only be used when changes are logical. It is advisable to make the client aware that force cleaning is being used and the specifications under which data are being changed." (CASRO, 1)

"Survey reporting guidelines: [...] If an index or some other constructed variable is used, show the method of calculation." (CASRO, 1)

⁵⁸ "According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Was the data analyzed in accordance with accepted statistical principles?" (Hebert, 1999)

"Basic criteria for measuring the social scientific integrity of the research include the following: [...] conformity of quantitative data to standard disciplinary measurements for statistical margin of error." (Cohen and Shirley, 1998, 7)

Kaye (2002) and Rubinfeld (2000) provide some guidelines as to which statistical tools are appropriate for different purposes.

aR2 ***For issues of causation, appropriate specification of the effect model***⁵⁹

Where the study intends to demonstrate the existence of a cause and effect relationship, the research should develop an effect model which identifies the main influences on the dependent variable. Data analysis should be designed to net out competing hypotheses, to all extent possible.

• ***data [a]nalysis [N]eutrality*** aN combination

aN1 ***Risk of disputes on results and interpretations***⁶⁰

Neutral data analysis derives conclusions which most informed analysts would derive from the same data. Excessive excursions beyond the information provided by the data shall be avoided.

aN2 ***Coherence of conclusions with results***⁶¹

Neutral data analysis provides conclusions and interpretations which are clearly in line with the results obtained from the study.

⁵⁹ Kaye lists several key questions relative to ascertaining the quality of statistical analyses, among them, in the context of the substantiation of causal relationships, "What are the independent and dependent variables?" and "What are the confounding variables?" (Kaye, 2000, 348)

"Ideally, a multiple regression analysis builds on a theory that describes the variables to be included in the study. [...] Failure to develop the proper theory, failure to choose the appropriate variables, and failure to choose the correct form of the model can bias substantially the statistical results, that is, create a systematic tendency for an estimate of a model parameter to be too high or too low." (Rubinfeld, 2000, 423)

⁶⁰ "Members must not provide or allow interpretations of the research which are inconsistent with the data without protest" (PMRS, 1, rule 1.1ii)

⁶¹ "We shall not knowingly make interpretations of research results, nor shall we be tacitly permit interpretations that are inconsistent with the data available." (AAPOR, 1, rule I.3)
 "[...] great care should be taken to be sure that the conclusions and the findings presented are consistent." (AAPOR, 2)

- **data [a]nalysis [B]alance** aB combination

aB1 ***Appropriateness of methods considering the research purposes***⁶²

Methods of data analysis should be in sync with the objectives of the research. Generally speaking, balanced descriptive studies are based on descriptive statistics and balanced cause and effect studies are based on bivariate and, more likely, multivariate methods.

aB2 ***Appropriate of the research design considering the research purposes***⁶³

There should be a balance between the research objective and the research design. Generally speaking, descriptive purposes require descriptive studies (e.g., no need for comparison groups) whereas studies with a comparative objective necessitate comparative designs (e.g., time-series, comparison groups, correlational approaches)

⁶² "Members must recommend those techniques and methodologies which are appropriate to the objectives of the research, avoiding those which they believe may give misleading results" (PMRS, 1, rule 1.1i)

"A wide variety of methods are available to collect primary research data. The professional research organization selects the method that provides the most effective means of reliably and validly achieving the study's information objectives." (CASRO, 1)

"We shall recommend and employ only those tools and methods of analysis which, in our professional judgement, are well suited to the research problem at hand." (AAPOR, 1, rule I.1)

"The following questions should be considered in evaluating the admissibility of statistical evidence. These considerations are motivated by two concerns: [...] (2) Are the methodological choices that the expert made reasonable, or are they arbitrary and unjustified?" (Rubinfeld, 2000, 441)

⁶³ "The structure of the questions is suitable for the statistical techniques that will be used in the analysis of the data." (CASRO, 1)

"Controlled experiments are, far and away, the best vehicle for establishing a causal relationship." (Kaye, 2000, 347)

"[...] we must be guided by a commitment to serving the client's interests as if we were the client her- or himself. Thus, we are obligated to develop and implement survey designs that will meet client objectives at a reasonable cost. [...] In other words, I maintain that professional ethics require us to make clear to clients the implications of adopting a particular course of action but not necessarily to insist upon a rigid conformity to performance standards." (Crespi, 1998)

- **data [a]nalysis [T]ransparency** aT combination

aT1 **Raw and weighted sample sizes in tables**⁶⁴

There is a consensus in the literature and, to a large extent, in the practice of survey research, that raw (unweighted) sample sizes must be supplied for all subgroups which are subjected to reporting. Several sources also require that weighted sample sizes be reported although this author does not consider this an absolute requirement as long as inferential statistics are computed using raw frequencies.

aT2 **Description of analysis methods used**⁶⁵

Transparent data analysis translates into a clear description of the procedures used so that replication can take place.

⁶⁴ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] x) weighted and unweighted bases for all conventional tables, clearly distinguished between the two [...]." (PMRS, 1, rule 4.6)

"All tables should clearly indicate the response base, whether it is total respondents, respondents meeting certain criteria, or any other base, such as total units or households." (CASRO, 1)

"Weighted data should carry a notation to that effect." (CASRO, 1)

"Usually, the small-base problem is obvious if the presentation is reasonably complete. An expert who says that 50% of the people interviewed had a certain opinion also should reveal how many individuals were contacted and how many expressed an opinion. Then we know whether the 50% is 2 out of 4 or 500 out of 1,000." (Kaye, 2000, 355)

"The completeness of the survey report is one indicator of the trustworthiness of the survey and the professionalism of the expert who is presenting the results of the survey. A survey report generally should provide in detail 8. statistical tables clearly labeled and identified as to source of data, including the number of raw cases forming the base for each table, row, or column; [...]." (Diamond, 2000, 270-271)

⁶⁵ "For each survey, a practitioner must provide to the client [...] information sufficient to replicate the study [...]. If applicable, the following information is required: [...] xvi) the detail of any special statistical methods used in the analysis of the results [...]." (PMRS, 1, rule 4.6)

"Survey reporting guidelines: [...] If special scoring, data adjustment or indexing methods are used, these should be described." (CASRO, 1)

"Survey reporting guidelines: [...] If statistical significance is noted, there should be an indication of the test used, significance level, and number of tails." (CASRO, 1)

"The following questions should be considered in evaluating the admissibility of statistical evidence. These considerations are motivated by two concerns: (1) Has the expert provided sufficient information to replicate the multiple regression analysis?" (Rubinfeld, 2000, 441)

"Critical in demonstrating adherence to standards of excellence and validity in defining the universe, designing the sample, learning what people think and do, and reporting the findings are: [...] (2) full documentation of raw data (particularly the completed questionnaires) which embody methods and results; and (3) step-by-step demonstration of how methods were used and how conclusions were reached." (Sorensen, 352).

[r]eporting

- **[r]eporting [R]igour** rR combination

rR1 **Disclosure of study weaknesses and possible bias**⁶⁶

Rigorous reporting includes the upfront disclosure of weaknesses and possible biases. Such disclosure gives the researcher an opportunity to weight the importance of these limitations and to provide context for them.

- **[r]eporting [N]eutrality** rN combination

rN1 **Distinct presentation of objective results and of their interpretation**⁶⁷

Reporting should clearly distinguish the statements of facts which are drawn from the observed data and the interpretation of these facts against the hypotheses or theories used by the researcher. It does not follow that research reports must have separate sections for the description of findings and the analysis of their implication, but the reporting must make the level of the presentation clear.

⁶⁶ "Survey reporting guidelines: [...] If the survey is a non-probability survey, it should be clearly stated that the results are not projectable to the entire universe. [...] In reports of qualitative research, a statement should be included to the effect that the research is based on a small, geographically limited sample and that, accordingly, the findings should be regarded as hypotheses subject to confirmation. The "tone" of the text reporting results should reflect this limitation. [...] In reports where a non-representative sample is used for the study, a statement should be included, either in the results or the technical appendix, discussing the implications and limitations of using a non-representative sample and the "tone" of the report should reflect these limitations." (CASRO, 1)

"[...] statisticians should be prepared to document data sources used in an inquiry; known inaccuracies in the data; and steps taken to correct or to refine the data, statistical procedures applied to the data, and the assumptions required for their application." (ASA, 1)

⁶⁷ "When reporting findings of a study in either written or oral form, the practitioner should make a clear distinction between the objective results and his or her own opinions and recommendations." (PMRS, 1, rule 4.14)

"The practitioner should provide to the client in the report, or in a supporting document, in addition to the items listed in 4.6, the following information: i) a discussion of any aspects of the research which may bias the results; ii) an assessment of the reliability of the sources used in secondary research [...]." (PMRS, 1, rule 4.15)

"Findings and interpretations should be presented honestly and objectively, with full reporting of all relevant findings, including any that may seem contradictory or unfavorable." (AAPOR, 2)

"Conclusions should be carefully distinguished from the factual findings [...]." (AAPOR, 2)

"According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Was the data that was gathered accurately reported?" (Hebert, 1999)

"When reporting on the results of a marketing research project the Researcher must make a clear distinction between the findings as such, the Researcher's interpretation of these and any recommendations based on them." (ESOMAR, 1)

rN2 Selflessness of the researcher⁶⁸

Neutral reporting should indicate the level of selflessness enjoyed by the researcher. In particular, the absence of a link between study results and remuneration is worth noting.

- **[r]eporting [B]alance**

rB combination

rB1 Level of support to conclusions offered by study results⁶⁹

Balanced reporting spells out the level of support that the study provides for the conclusions reached. In particular, inconclusive demonstration should not be offered as proof on one's hypothesis.

⁶⁸ "The practitioner should make known any current involvement in the same general subject area before accepting a project." (PMTS, 1, rule 4.12)

"For these reasons, statisticians should [...] disclose any financial or other interests that may affect, or appear to affect, their professional statements." (ASA, 1)

"[...] statisticians should inform a client or employer of all factors that may affect or conflict with their impartiality and accept no contingency fee arrangements". (ASA, 1)

"According to Manual on Complex Litigation §21.493 at 102 (3d ed Federal Judicial Center 1995), there are seven topics you need to address to assess whether a survey was properly done, i.e.: Was the objectivity of the entire process assured?" (Hebert, 1999)

"Ideally, expert who conduct research for litigants should proceed with the same objectivity as they would apply in other contexts. Thus, if experts testify or if their results are used in testimony by others, they should be free to do whatever analysis and have access to whatever data are required to address the problems the litigation poses in a professionally responsible fashion" (Kaye, 2000, 337-338)

"The validity and value of public opinion polls depend on three main considerations: i. the nature of the research techniques used and the efficiency with which they are applied, ii. the honesty and objectivity of the research organisation carrying out the study, iii. the way in which the findings are presented and the uses to which they are put." (ESOMAR, 1)

⁶⁹ "Members must not present research results with greater confidence than the data warrant." (PMRS, 1, rule 1.1iii)

"We shall not knowingly imply that interpretations should be accorded greater confidence than the data actually warrant." (AAPOR, 1, rule I.4)

"Researchers must not knowingly allow the dissemination of conclusions from a marketing research project which are not adequately supported by the data. They must always be prepared to make available the technical information necessary to assess the validity of any published findings." (ESOMAR, 1)

rB2 **Possible reservations on the internal and external validity of results**⁷⁰

The researcher must candidly express reservations that should be known about the strength of the demonstration provided by the research and about the generalisability of study findings.

- **[r]eporting [T]ransparency** rT combination

rT1 **Understandability of the presentation**⁷¹

Transparent reporting should be understandable to the audience targeted. Unnecessary use of jargon is a barrier to transparent reporting among the uninitiated. Conversely, an overly simplified presentation may elude important details for the specialist.

⁷⁰ "Instead, as responsible professionals, members must point out the relevant limitations of the research." (PMRS, 1, rule 1.1.iii)
 "The Client is entitled to the following information about any marketing research project to which he has subscribed: [...] (2) Sample - [...] and a discussion of any possible bias due to non-response." (ESOMAR,1)
 "All experiments are conducted with a sample of a certain population, at a certain place, at a certain time, and with a limited number of treatments. With respect to the sample studied, the experiment may be persuasive. It may have succeeded in controlling all confounding variables and in finding an unequivocally large difference between the treatment and control groups. If so, its 'internal validity' will not be disputed; in the sample studied, the treatment has an effect. But an issue of 'external validity' remains. To extrapolate from the limiting conditions of an experiment always raises questions. If juries react differently to competing instructions on the law of insanity in cases of housebreaking and of incest, would the difference persist if the charge were rape or murder? Would the failure of ex-convicts to react to transitory payments after release hold if conditions in the employment market were to change radically?" (Kaye, 2000, 349)
 "However, association is not causation, and the causal inferences that can be drawn from such analyses [observational studies] rest on a less secure foundation than that provided by a controlled randomized experiment." (Kaye, 2000, 351)

⁷¹ "In order to have quality, research findings must be prepared and presented to the client in a manner that is both understandable and compelling." (Cohen and Shirley, 1998, 8)

rT2 **Disclosure of all elements of information required by all 24 framework cells to assess the value of results**⁷²

Clearly, there is a tension between efficiency and cost, on the one hand, and transparent reporting, on the other hand, and some sponsors may not accept to defray the cost of transparency. This is a small problem if the dissemination of study results is limited to individuals and organizations who accept this tradeoff. When study results are disseminated more widely, however, there is a joint researcher-sponsor responsibility to tip the balance on the side of transparency. This framework gives researchers and sponsors a structure to determine where transparency investment will pay off.

Codes of ethics

Referring to the traditional literature in the field, one might ask: what about reliability, validity and ethics? We contend that traditional concerns related to research quality and, to a certain extent, to social acceptability or ethics, are contained within the framework proposed here; meanwhile, the framework goes much beyond that to offer more prescriptive guidance to researchers, auditors and assessors.

While this assessment model does not duplicate codes of ethics or rules of good practices such as the ones offered by AAPOR and PMRS, it offers the analyst a strong, defensible framework which ensures the thoroughness and balance of the review. The same approach can be used as a checklist for one's own research work.

72

"Excellence in survey practice requires that survey methods be fully disclosed reported in sufficient detail to permit replication by another researcher and that all data (subject to appropriate safeguards to maintain privacy and confidentiality) be fully documented and made available for independent examination.[...] Exemplary practice in survey research goes beyond such standards for "minimal disclosure," promulgated by AAPOR and several other professional associations ." (AAPOR, 2)

"At a minimum, both academic and commercial survey researchers are obliged to describe their research designs, and to present their results, as fully and openly as possible so that competent others can evaluate their projects. Unfortunately, there will be times when this obligation will conflict with the confidentiality of client interests, placing the researcher in a very difficult position." (Crespi, 1998)

"To make it possible for the court and the opposing party to closely scrutinize the survey so that its relevance, objectivity, and representativeness can be evaluated, the party proposing to offer the survey as evidence should describe in detail the design and execution of the survey." (Diamond, 2000, 232)

"Whatever information may be given in the published report of the survey, the publisher and/or the research organisation involved must be prepared on request to supply the other information about the survey methods described in the Notes on the application of Rule 25 of the International Code." (ESOMAR, 1)

This framework does not constitute a code of ethics. Research which would provide solid evidence regarding the issue at hand according to the criteria used here could still be considered unethical on other grounds. For example, the AAPOR code of ethics indicates that "If we become aware of the appearance in public of serious distortions of our research, we shall publicly disclose what is required to correct these distortions, including, as appropriate, a statement to the public media, legislative body, regulatory agency, or other appropriate group, in or before which the distorted findings were presented"; a contravention to this rule would not transform good research into bad research.

However, relative to qB (questionnaire–Balance) and cB (data collection–Balance), where the research endeavour interacts with the public, sound research criteria do include aspects which would usually be comprised in codes of ethics, such as the level of response burden, informed consent and confidentiality of the information collected.

This framework does not attempt to integrate all detailed prescriptions of codes of ethics. Rather, it reiterates the main themes which form the basis for responsible research behaviour. The debate is still open, however: can research be "sound" (i.e., pass the test of an assessment framework) without embodying every aspect of established codes of ethics?

An application to Internet surveys

It is our position that this framework provides an exhaustive depiction of the assessment territory for survey research. In this section, we put the model to a test: based on these criteria for sound survey research, which areas of the 24 assessment zones are likely to be impacted by a switch to Internet-based surveying and to other technological advancements?

Mick P. Couper (2000) identified the methodological concerns which are specific to Internet-based research — or more prominent in this type of research. They are:

- ***Sampling frame coverage issues:*** "Coverage error is presently the biggest threat to inference from Web surveys, at least to groups beyond those defined by access and use of the Web. The problems of sampling

in many Web surveys also present a formidable barrier to probability-based sample surveys on the Web. Coverage error is a function of the mismatch between the target population and the frame population." (Couper, 2000, 467).

- **Non-response issues:** "Nonresponse error arises through the fact that not all people included in the sample are willing or able to complete the survey. As with coverage error, nonresponse error is a function of both the rate of nonresponse and of the differences between respondents and nonrespondents on the variables of interest. [...] Several recent studies have compared response rates from e-mail studies to those from mail surveys of the same populations. [...] For all but one study, the e-mail surveys failed to reach the response rate levels of the mail surveys." (Couper, 2000, 473)
- **Questionnaire design uniformity.** "While the importance of question wording in influencing respondent answers is well-recognized, there is a growing literature that suggests that the design of the survey instrument (such as the placement of questions, flow of instrument, typographical features, etc.) also plays an important role, in both self-administered and interviewer-administered surveys. [...] unintended design or layout changes can affect the responses obtained both in interviewer-administered and in self-administered surveys. On the Web, unlike on paper, the appearance of a survey can vary from respondent to respondent because of different browser settings, user preferences, variations in hardware, and so on. Design may thus be much more important for Web surveys, both because there are more tools available to the designer (color, sound, images, animation, etc.) and because of variation in how these may be seen by respondents." (Couper, 2000, 476–477)
- **Panel conditioning.** "Another source of measurement error that is unique to panel or longitudinal surveys (often employed in Web surveys) is that of panel conditioning (or time-in-sample bias). Panel conditioning occurs through the ongoing participation of members in a panel. Given their experience with the survey over time, their responses may increasingly begin to differ from the responses given by people answering the same survey for the first time. Given the current lack of a suitable sampling frame, many survey organizations are creating panels of Web

survey respondents, and panel effects remain a concern for such surveys. Even though the surveys may vary over time, the mere act of participating in an ongoing panel may change respondent behavior and attitudes." (Couper, 2000, 476)

Where, in our proposed assessment framework, do these four particular concerns of Internet-based research fall?

Sampling frame coverage issues are a sample neutrality (sN) issue, as is the non-response problem: studies using truncated sampling frames and suffering from low participation have a hard time demonstrating neutrality in the representation of the target universe.

Browser-related variations in rendering of questionnaire on-line and the ensuing lack of uniformity in the stimulus offered to respondents is a questionnaire neutrality (qN) issue: it may affect in a non-random fashion the way respondents answer questions, thereby affecting the reliability of the estimates produced.

Panel conditioning is a data collection neutrality (cN) issue since it produces a contamination effect — the data collection itself risking to affect respondents attitudes and behaviour.

A few conclusions are in order:

- First, based on Couper's analysis as applied to our assessment framework, Internet-based survey research issues focus on neutrality concerns and on the ability of this type of research to properly depict reality.
- Secondly, only three of the 24 assessment areas (sN, qN and cN) raise particular concerns in the context of Internet-based research. This is not to say that the other 21 assessment areas are of no interest, but rather that they garner the same importance in this type of research that they normally possess in any type of survey research endeavour.
- Finally, the necessary conclusion of this analysis is that Internet-based research does not represent a fundamental paradigm shift in social research. The same criteria apply to it as they do to other types of

research, to the extent that the knowledge objectives are the same. If Internet-based research wants to compete with other types of research to quench managers' thirst for knowledge and competitive edge, it must be subjected to the same rules and abide by the same standards.

Can research project be scored?

All preliminary presentations of this work raised the issue of the ability of the framework to provide a marking system for survey research. Did research no. 1 get it correct under 95% of the criteria and therefore get an A? Did research no. 2 fail on 20% of the criteria and therefore get a D?

No. The assessment framework should not be used this mathematically. It is conceivable that a research project concerned with language assimilation meet all of the criteria laid out in the framework except one, say, the representativity of the sample, which would disqualify it as a credible description of this social phenomenon. A 95% mark would not do good to that research. Conversely, one can imagine a study on the impact of shape, form and colour on product branding which may fail on the same sample representativity criterion, but still be considered adequate for its purpose.

Assessing survey research requires the application of professional judgment to individual research situations; meeting a certain criterion may be of small importance in one setting while failing it may be crucial in another situation. The framework does not provide a quick rating scale to do so, but rather a systematic protocol to assist the process.

Conclusion

Kaase (1999) suggests that survey quality is best represented by accurate results and appropriate methods.⁷³ Accurate results, as an indicator of survey quality, are rarely available to the researcher, to the client, to the

73

"Survey research is a method for making generalisable statements about social facts with limited economic means. The core of this social science invention is constituted by two methods, *representative sampling* in conjunction with *standardised interviewing*. On this basis, the criteria for the quality of a survey are clear: the undistorted representation of a defined population by the sample, [and] the valid, reliable measurement of the parameters at issue by the interview. These are theoretical criteria that apply absolutely and always. The problem is to determine in any given case whether the criteria have been met, and if not why not. Two types of indicator estimate survey quality, the instrumental criteria, at it were: accurate results, [and] appropriate methods." (Kaase, 1999, 238).

assessor and to society. Therefore, our assessment framework has focussed on the appropriateness of the methods used by the research. It aims at offering an organized set of criteria for evaluating survey research exhaustively, fairly and constructively. While it is not construed as a simple checklist which could provide an overall assessment without the use of experience and judgment, it is meant as a guide for the experienced researcher and as a heuristic tool for the beginner.

References

- Akkerboom, Hans and Håkan L. Lindström (1998), "The Development of Quality Concepts for Questionnaire Testing and Evaluation", WAPOR regional seminar on Quality Criteria in Survey Research, June 25-27, 1998, Cadenabbia, Italy.
- American Association for Public Opinion Research (1), *Code of Professional Ethics and Practices*, <http://www.aapor.org/ethics/code.html>, as of 2002.09.18
- American Association for Public Opinion Research (2), *Best Practices for Survey and Public Opinion Research*, <http://www.aapor.org/ethics/best.html>, as of 2002.09.18
- American Association for Public Opinion Research (3), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, <http://www.aapor.org/ethics/stddef.html>, as of 2002.09.18
- American Association for Public Opinion Research (4), *Statement to Institutional Review Boards*, <http://www.aapor.org/ethics/stateirb.html>, as of 2002.09.18
- American Statistical Association (1), *Ethical Guidelines for Statistical Practice*, <http://www.amstat.org/profession/ethicalstatistics.html>, as of 2002.09.18
- American Statistical Association (2), *Surveys and Privacy*.
- Babbie, Earl (1990), *Survey Research Methods*, Belmont, California, Wadsworth Publishing Company, second edition, 1990.
- Blais, André and Claire Durand (1997), "Le sondage" in Benoît Gauthier (ed.), *Recherche sociale, de la problématique à la collecte des données*, Québec, Presses de l'Université du Québec, 1997, third edition, pp. 357–399.
- Blum, Olivia (1997), "Evaluation of data-editing using administrative records", Work Session on Statistical Data Editing, Conference of European Statisticians, Statistical Commission and Economic Commission for Europe, Prague, October 14-17, 1997.
- Bohrnstedt, George W. (1983), "Measurement" in Peter H. Rossi, James D. Wright and Andy B. Anderson, *Handbook of survey research*, New York, Academic Press, 1983, pp. 69-121.

- Breault, Bernard, *Exercer la saine gestion, Théorie appliquée à l'audit de saine gestion*, Farnham, Quebec, Publications CCH, second edition, 1999.
- Bureau du Vérificateur général du Canada (1998), *Document d'orientation fonctionnelle, Guide pour la tenue d'enquêtes*, août 1998.
- Canadian Survey Research Council (1), *Declaration of principles*, <http://www.csrc.ca/CSRC/whoweare/principles.php>, as of 2002.09.21
- Carmines, Edward G. and Richard A. Zeller. *Reliability and validity assessment*, Sage University Papers on Quantitative Applications in the Social Sciences, Beverly Hills, Sage Publications, 1979, 71 pages.
- Cohen, Michael and David Shirley (1998), "Intersubjectivity in Applied Social Scientific Research: Toward a Working Definition of Quality", WAPOR regional seminar on Quality Criteria in Survey Research, June 25-27, 1998, Cadenabbia, Italy.
- Corbin, Ruth M., A. Kelly Gill and R. Scott Joliffe (2000), *Trial by survey: survey evidence and the law*, Toronto, Carswell, 2000, 196 pages.
- Council for Marketing and Opinion Research (1), *Respondent Bill of Rights*, http://www.cmor.org/what_is_research_rights.htm, as of 2002.09.18
- Council of American Survey Research Organizations (1), *Guidelines for Survey Research Quality*, <http://www.casro.org/guidelines.cfm>, as of 2002.09.18
- Couper, Mick P. (2000), "Web Surveys, A Review of Issues and Approaches", *Public Opinion Quarterly*, vol. 64, Winter 2000, pp. 464-494.
- Crespi, Irving (1998), "Ethical considerations when establishing survey standards" in *International Journal of Public Opinion Research*, spring 1998, vol. 10, p. 75 and ss.
- Diamond, Shari Seidman (2000), "Reference Guide on Survey Research" in Federal Judicial Center, *Reference Manual on Scientific Evidence, Second Edition*, 2000, <http://www.daubertexpert.com/link1.htm> as of 2002.09.19
- Durand, Claire et André Blais (1997), "La mesure" in Benoît Gauthier (ed.), *Recherche sociale, de la problématique à la collecte des données*, Québec, Presses de l'Université du Québec, 1997, third edition, pp. 159-184.

- ESOMAR (1), *Guidelines to opinion polls*, <http://www.esomar.nl/guidelines/opolls.htm>, as of 2002.09.18
- Grover, Varun (1), *A Tutorial on Survey Research: From Constructs to Theory*, <http://dmsweb.badm.sc.edu/grover/survey/MIS-SUVY.html>, as of 2002.09.18
- Groves, Robert M. (1989), *Survey Errors and Survey Costs*, New York, John Wiley and Sons, 1989.
- Groves, Robert M. and Mick P. Couper (1998), *Nonresponse in Household Interview Surveys*, New York, John Wiley and Sons, 1998.
- Hartung, Uwe (2001), "Third Cadenabbia Seminar on Quality Criteria in Survey Research", *International Journal of Public Opinion Research*, Spring 2001, pp. 108–113.
- Hebert, William N. (1999), "Cross-Examining Survey Experts" in *Civil Litigation Reporter*, November 1999.
- Hirschi, Travis and Hanan C. Selvin (1975), *Recherches en délinquances, principes de l'analyse quantitative*, Paris, Mouton, 1975.
- Kaase, Max (1999), *Quality Criteria for Survey Research*, Berlin, Akademie Verlag, 1999, 288 pages.
- Kaye, David H. (2000), "Reference Guide on Statistics" in Federal Judicial Center, *Reference Manual on Scientific Evidence*, 2000, <http://www.daubertexpert.com/link1.htm> as of 2002.09.19
- K-BASE (1), "The knowledge base on statistical data editing", <http://www3.sympatico.ca/guylaine.theoret/k-base/> as of 2002.09.20
- Kovar, John G. (1997), "Canada: National Report", Work Session on Statistical Data Editing, Conference of European Statisticians, Statistical Commission and Economic Commission for Europe, Prague, October 14-17, 1997.
- Mellos, Koula (1997), "Une science objective" in Benoît Gauthier (ed.), *Recherche sociale, de la problématique à la collecte des données*, Québec, Presses de l'Université du Québec, 1997, third edition, pp. 497–514
- Noelle-Neumann, Elisabeth, "Quality criteria in survey research: Keynote speech presented at the WAPOR Thematic Seminar on June 27, 1996,

- Cadenabbia, Italy", *International Journal of Public Opinion Research*, vol. 9, no. 1, Spring 1997, pp. 29–32.
- Nordbotten, Svein (2000), *Evaluating Efficiency of Statistical Data Editing: General Framework*, Geneva, United Nations Statistical Commission, 2000.
- Ordre des administrateurs agréés du Québec (1), *Guide de l'administrateur agréé, Principes et normes de saine gestion*, Publications CCH/FM (regular updates)
- Petersen, Thomas (1998), "Three aspects of quality in survey research: indirect questions, field experiments and the need for researchers to test questionnaires themselves", WAPOR regional seminar on Quality Criteria in Survey Research, June 25-27, 1998, Cadenabbia, Italy.
- Professional Marketing Research Society (1), *Rules of Conduct and Good Practice 1999/2000*, <http://www.pmr-s-aprm.com/What/Conduct.html> as of 2002.09.18
- Rubinfeld, Daniel L.,(2000), "Reference Guide on Multiple Regression" in Federal Judicial Center, *Reference Manual on Scientific Evidence*, 2000, <http://www.daubertexpert.com/link1.htm> as of 2002.09.19
- Sorensen, Robert C., "Survey Research Execution in Trademark Litigation: Does Practice Make Perfection", *The Trademark Reporter*, vol. 73, pages 349-365.
- Spencer, Bruce D. (2000), "An Approximate Design Effect for Unequal Weighting When Measurements May Correlate With Selection Probabilities", *Survey Methodology*, December 2000, vol. 26, no. 2, pp. 137-138.
- Stern, Paul C. (1979), *Evaluating Social Science Research*, New York, Oxford University Press, 1979.
- Whitridge, Patricia and Julie Bernier (1999), "The impact of editing on data quality", UC/ECE Work Session on Statistical Data Editing, Conference of European Statisticians, Statistical Commission and Economic Commission for Europe, Rome, June 2-4, 1999.
- Zimmerman, Richard E., Linda Steinmann and Vince Schueler (1), *Designing Customer Surveys That Work*, <http://www.qualitydigest.com/oct96/surveys.html> as of 2002.09.18