

PART I
SETTING THE STAGE

COPYRIGHTED MATERIAL

1

Comparative Survey Methodology

*Janet A. Harkness, Michael Braun, Brad Edwards,
Timothy P. Johnson, Lars Lyberg, Peter Ph. Mohler,
Beth-Ellen Pennell, and Tom W. Smith*

1.1 INTRODUCTION

This volume discusses methodological considerations for surveys that are deliberately designed for comparative research such as multinational surveys. As explained below, such surveys set out to develop instruments and possibly a number of the other components of the study specifically in order to collect data and compare findings from two or more populations.

As a number of chapters in this volume demonstrate, multinational survey research is typically (though not always) more complex and more complicated to undertake successfully than are within-country cross-cultural surveys. Many chapters focus on this more complicated case, discussing multinational projects such as the annual International Social Survey Programme (ISSP), the epidemiologic World Mental Health Initiative survey (WMH), the 41-country World Fertility Survey (WFS), or the triennial and worldwide scholastic assessment Programme for International Student Assessment (PISA). Examples of challenges and solutions presented in the volume are often drawn from such large projects.

At the same time, we expect many of the methodological features discussed here also to apply for within-country comparative research as well. Thus we envisage chapters discussing question design, pretesting, translation, adaptation, data collection, documentation, harmonization, quality frameworks, and analysis to provide much of importance for within-country comparative researchers as well as for those involved in cross-national studies.

This introductory chapter is organized as follows. Section 1.2 briefly treats the growth and standing of comparative surveys. Section 1.3 indicates overlaps between multinational, multilingual, multicultural, and multiregional survey

¹ Survey Methods in Multinational, Multiregional, and Multicultural Contexts, edited by Harkness et al. Copyright © 2010 John Wiley & Sons, Inc.

research and distinguishes between comparative research and surveys deliberately designed for comparative purposes. Section 1.4 considers the special nature of comparative surveys, and Section 1.5 how comparability may drive design decisions. Section 1.6 considers recent changes in comparative survey research methods and practice. The final section, 1.7, considers ongoing challenges and the current outlook.

1.2 COMPARATIVE SURVEY RESEARCH: GROWTH AND STANDING

Almost without exception, those writing about comparative survey research—whether from the perspective of marketing, the social, economic and behavioral sciences, policy-making, educational testing, or health research—remark upon its “rapid,” “ongoing,” or “burgeoning” growth. And in each decade since World War II, a marked “wave” of interest in conducting cross-national and cross-cultural survey research can be noted in one discipline or another (see contributions in Bulmer, 1998; Bulmer & Warwick, 1983/1993; Gauthier, 2002; Hantrais, 2009; Hantrais & Mangen, 2007; Øyen, 1990; and Chapters 2 and 25, this volume).

Within the short span of some 50 years, multipopulation survey research has become accepted as not only useful and desirable but, indeed, as indispensable. In as much as international institutions and organizations—such as the European Commission, the Organization for Economic Co-operation and Development (OECD), the United Nations (UN), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the World Bank, the International Monetary Fund (IMF), and the World Health Organization (WHO)—depend on multinational data to inform numerous activities, it has become ubiquitous and, in some senses, also commonplace.

1.3 TERMINOLOGY AND TYPES OF RESEARCH

In this section we make a distinction which is useful for the special methodological focus of many chapters in this volume—between comparative research in general and deliberately designed comparative surveys.

1.3.1 Multipopulation Surveys: Multilingual, Multicultural, Multinational, and Multiregional

Multipopulation studies can be conducted in one language; but most multipopulation research is nonetheless also multilingual. At the same time, cultural differences exist between groups that share a first language both within a country (e.g., the Welsh, Scots, Northern Irish, and English in the United Kingdom) and across countries (e.g., French-speaking nations/populations). Language difference (Czech versus Slovakian, Russian versus Ukrainian) is, therefore, not a necessary prerequisite for cultural difference, but it is a likely indicator of cultural difference.

Within-country research can be multilingual, as reflected in national research conducted in countries as different as the Philippines, the United States, Switzerland, Nigeria, or in French-speaking countries in Africa. Cross-national projects may thus often need to address within-country differences in language and culture in addition to across-country differences, both with respect to instrument versions and norms of communication.

Multiregional research may be either within- or across-country research and the term is used flexibly. Cross-national multiregional research may group countries considered to “belong together” in some respect, such as geographical location (the countries of Meso and Latin America), in demographical features (high or low birth or death rates, rural or urban populations), in terms of developmental theory (see Chapter 4, this volume) or in terms of income variability. Other multiregional research might be intent on covering a variety of specific populations in different locations or on ensuring application in a multitude of regions and countries. Within-country multiregional research might compare differences among populations in terms of north-south, east-west or urban-rural divisions.

1.3.2 Comparative by Design

This volume focuses on methodological considerations for surveys that are deliberately planned for comparative research. These are to be understood as projects that deliberately design their instruments and possibly other components of the survey in order to compare different populations and that collect data from two or more different populations. In 1969, Stein Rokkan commented on the rarity of “deliberately designed cross-national surveys” (p. 20). Comparative survey research has grown tremendously over the last four decades and is ubiquitous rather than rare. However, Rokkan’s warning that these surveys are not “surefire investments” still holds true; the success of any comparative survey requires to be demonstrated and cannot be assumed simply on the basis of protocols or specifications followed. Numerous chapters in this volume address how best to construct and assess different aspects of surveys designed for comparative research.

Comparative instruments are manifold in format and purpose: educational or psychological tests, diagnostic instruments for health, sports performance, needs or usability assessment tools; social science attitudinal, opinion and behavioral questionnaires; and market research instruments to investigate preferences in such things as size, shape, color, or texture. Several chapters also present comparative methodological studies.

Comparative surveys are conducted in a wide variety of modes, can be longitudinal, can compare different populations across countries or within countries, and can be any mix of these. Some of the studies referred to in the volume are longitudinal in terms of populations studied (panels) or in terms of the contents of the research project (programs of replication). Most of the methodological discussion here, however, focuses on synchronic, across-population research rather than on across-time perspectives (but see Lynn, 2009; Duncan, Kalton, Kasprzyk, & Singh, 1989; Smith, 2005).

Comparative surveys may differ considerably in the extent to which the deliberate design includes such aspects as sampling, the data collection process, documentation, or harmonization. In some cases, the instrument is the main component “designed” to result in comparable data, while many other aspects are decided at the local level (e.g., mode, sample design, interviewer assignment, and contact protocols). Even when much is decided at the local level, those involved in the project must implicitly consider these decisions compatible with the comparative goals of the study.

If we examine a range of large-scale cross-national studies conducted in the last few decades (see, for example, Chapters 25–31, this volume), marked differences can also be found in study design and implementation. Studies vary greatly in the level of coordination and standardization across the phases of the survey life cycle, for example, in their transparency and documentation of methods, and in their data collection requirements and approaches.

1.3.3 Comparative Uses of National Data

Comparative research (of populations and locations) need not be based on data derived from surveys deliberately designed for that purpose.

A large body of comparative research in official statistics, for instance, is carried out using data from national studies designed for domestic purposes which are then also used in analyses across samples/populations/countries. Early cross-national social science research often consisted of such comparisons (cf. Gauthier, 2002; Mohler & Johnson, this volume; Rokkan, 1969; Scheuch, 1973; Verba, 1969). Official statistics agencies working at national and international levels (UNESCO Statistics; the European statistical agency, Eurostat; and national statistical agencies such as the German Statistisches Bundesamt and Statistics Canada) often utilize such national data for comparative purposes, as do agencies producing international data on labor force statistics (International Labour Organization; ILO), on income, wealth, and poverty (Luxembourg Income Study; LIS), and on employment status (Luxembourg Employment Study; LES). Such agencies harmonize data from national studies and other sources because adequately rich and reliable data from surveys that were deliberately designed to produce cross-national datasets are not available for many topics. The harmonization strategies used to render outputs from national data comparable (ex-post output harmonization) are deliberately designed for that purpose (see, for example, Ehling, 2003); it is the national surveys themselves which are not comparative by design. A partnership between Eurostat and many national statistical offices has resulted in the European Statistical System, an initiative which aims to provide reliable and comparable statistics for all the European Union and the European Free Trade Association Member States on the basis of national data.

Instruments designed for a given population are also frequently translated and fielded with other populations. Such translated versions can be tested for suitability with the populations requiring the translations (see Chapters 5, 6, and 7, this volume) and may produce data that permit comparison. Nonetheless, the

original (source) instrument was not comparative by design. Publications arguing the validity and reliability of “translated/adapted” instruments abound, particularly in health, opinion, and psychological research. While the suitability of these procedures and instruments is sometimes contested (e.g., Greenfield, 1997), such instruments may be translated into many languages and used extensively worldwide. In some cases, feedback from implementations in other languages can lead to adjustments to the original instrument. One prominent example is the development of the SF-36 Health Survey, a short (36-question) survey that has been translated and adapted in over 50 languages. The development of translated versions led to related modifications in the original English questionnaire (cf. Ware, undated, at <http://www.sf-36.org/tools/SF36.shtml/>).

Finally, we note that the distinction between comparative research and research that is comparative by design used here is not one always made. Lynn, Japac, and Lyberg (2006), for example, use the term “cross-national surveys” to refer to “all types of surveys where efforts are made to achieve comparability across countries. Efforts to achieve comparability vary on a wide spectrum from opportunistic adjustment of data after they have been collected to deliberate design of each step in the survey process to achieve functional equivalence” (p. 7). The latter of these would fall under our definition of “surveys comparative by design;” those based on “opportunistic adjustment of data after they have been collected” would not.

1.4 WHAT IS (SO) SPECIAL ABOUT COMPARATIVE SURVEY RESEARCH?

Many discussions of comparative survey research note at some point that all social science research is comparative (cf. Armer, 1973; Jowell, 1998; Lipset, 1986; Smith, forthcoming).

Some also suggest that there is nothing really special about comparative (survey) research. Verba (1971 and 1969) and Armer (1973) seem to take this position—but simultaneously also document difference. Verba (1969) states, for example: “The problems of design for within-nation studies apply for across-nation studies. If the above sentence seems to say there is nothing unique about cross-cultural studies, it is intended. The difference is that the problems are more severe and more easily recognizable” (p. 313). Armer (1973) goes a step further: “My argument is that while the problems involved are no different in kind from those involved in domestic research, they are of such great magnitude as to constitute an almost qualitative difference for comparative as compared to noncomparative research” (p. 4).

Later researchers, focusing more on the design and organization of comparative surveys, point to what they consider to be unique aspects. Lynn, Japac, and Lyberg (2006) suggest “Cross-national surveys can be considered to have an extra layer of survey design, in addition to the aspects that must be considered for any survey carried out in a single country” (p. 17). Harkness, Mohler, and van de Vijver (2003) suggest that different kinds of surveys call for different tools and strategies. Certain design strategies, such as decentering,

certainly have their origin and purpose in the context of developing comparative instruments (cf. Werner & Campbell, 1970). The distinction between comparative research and surveys that are comparative by design accommodates the view that all social science research is comparative and that national data can be used in comparative research, while also allowing for the need for special strategies and procedures in designing and implementing surveys directly intended for comparative research.

There is considerable consensus that multinational research is valuable and also more complex than single-country research (Kohn, 1987; Jowell, 1998; Kuechler, 1998; Lynn, Japac, & Lyberg, 2006; Rokkan, 1969). The special difficulties often emphasized include challenges to “equivalence,” multiple language and meaning difficulties, conceptual and indicator issues, obtaining good sample frames, practical problems in data collection, as well as the sheer expense and effort involved. A number of authors in the present volume also point to the organizational demands as well as challenges faced in dealing with the varying levels of expertise and the different *modi operandi*, standards, and perceptions likely to be encountered in different locations.

1.5 HOW COMPARABILITY MAY DRIVE DESIGN

The comparative goals of a study may call for special design, process, and tool requirements not needed in other research. Examples are such unique requirements as decentering, or *ex ante* input harmonization (cf. Ehling, 2003). But deliberately designed comparative surveys may also simply bring to the foreground concerns and procedures that are not a prime focus of attention in assumed single-population studies (communication channels, shared understanding of meaning, complex organizational issues, researcher expertise, training, and documentation).

What Lynn and colleagues (2006) conceptualize as a layer can usefully be seen as a central motivation for design and procedures followed, a research and output objective at the hub of the survey life cycle that shapes decisions about any number of the components and procedures of a survey from its organizational structure, funding, working language(s), researcher training, and quality frameworks to instrument design, sample design, data collection modes, data processing, analysis, documentation, and data dissemination. Figure 1.1 is a simplified representation of this notion of comparability *driving* design decisions. For improved legibility, we display only four major components in the circle quadrants, instead of all the life-cycle stages actually involved. The comment boxes outside also indicate only a very few examples of the many trade-offs and other decisions to be made for each component in a comparative-by-design survey.

1.6 RECENT CHANGES IN PRACTICE, PRINCIPLES, AND PERSPECTIVES

The practices followed and tools employed in the design, implementation, and analysis of general (noncomparative) survey-based research have evolved rapidly

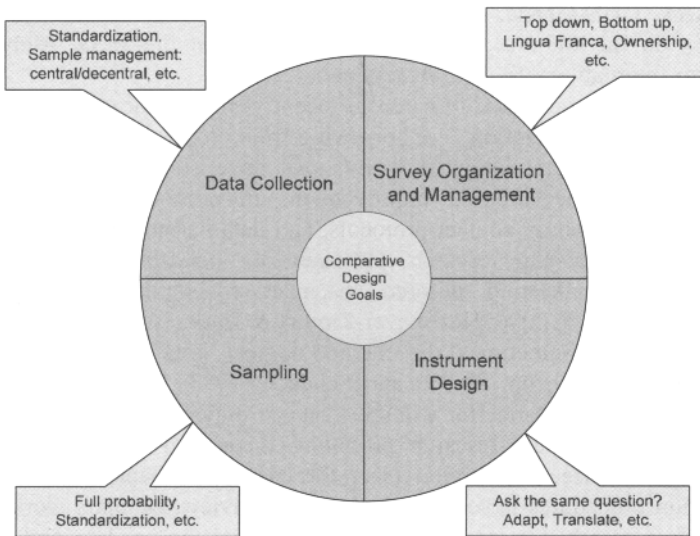


Figure 1.1. "Comparative-by-Design" Surveys

in recent decades. Albeit with some delay, these developments in general survey research methodology are carrying over into comparative survey research. A field of "survey methodology" has emerged, with standardized definitions and (dynamic) benchmarks of good and best practice (cf. Groves et al., 2009, pp. 1–37). Techniques and strategies emerging in the field have altered the way survey research is conceptualized, undertaken, and (now) taught at Masters and PhD level, quietly putting the lie to Scheuch's (1989) claim (for comparative surveys) that "in terms of methodology *in abstracto* and on issues of research technology, most of all that needed to be said has already been published" (p. 147).

The size and complexity of cross-national and cross-cultural survey research have themselves changed noticeably in the last 15–20 years, as have perspectives on practices and expectations for quality. Large-scale comparative research has become a basic source of information for governments, international organizations, and individual researchers. As those involved in research and analysis have amassed experience, the field has become increasingly self-reflective of procedures, products, and assumptions about good practice. In keeping with this, a number of recent publications discuss the implementation of specific projects across countries. These include Börsch-Supan, Jürges, and Lipps (2003) on the Survey on Health, Ageing and Retirement in Europe; Brancato (2006) on the European Statistical System; Jowell, Roberts, Fitzgerald, and Eva (2007b) on the European Social Survey (ESS); and Kessler and Üstün (2008) on the World Mental Health Initiative (WMH). Manuals and technical reports are available on the implementation of specific studies. Examples are Barth, Gonzalez, and

Neuschmidt (2004) on the Trends in International Mathematics and Science Study (TIMSS); Fisher, Gershuny, and Gauthier (2009) on the Multinational Time Use Study; Grosh and Muñoz (1996) on the Living Standards Measurement Study Survey; IDASA, CDD-Ghana, and IREEP (2007) on the Afrobarometer; ORC Macro (2005) on the Malaria Indicator Survey; and, on the PISA study, the Programme for International Student Assessment (2005).

Comparative methodological research in recent years has turned to questions of implementation, harmonization and, borrowing from cross-cultural psychology, examination of bias. Methodological innovations have come from within the comparative field; evidence-based improvements in cross-cultural pretesting, survey translation, sampling, contact protocols, and data harmonization are a few examples. Recent pretesting research addresses not just the need for *more* pretesting, but for pretesting tailored to meet cross-cultural needs (see contributions in Harkness, 2006; Hoffmeyer-Zlotnik & Harkness, 2005; Chapters 5 and 6, this volume). Sometimes the methodological issues have long been recognized—Verba (1969, pp. 80–99) and Scheuch (1993, 1968, pp. 110, 119) could hardly have been clearer, for example, on the importance of context—but only now is methodological research providing theoretical insights into how culture and context affect perception (see, for example, Chapters 10–12, this volume). Design procedures have come under some review; scholars working in Quality of Life research, for instance, have emphasized the need to orchestrate cross-cultural involvement in instrument design (Fox-Rushby & Parker, 1995; Skevington, 2002).

The increased attention paid to quality frameworks in official statistics comprising, among others, dimensions such as relevance, timeliness, accuracy, comparability, and coherence (Biemer & Lyberg, 2003; Chapter 13, this volume), combined with the “total survey error” (TSE) paradigm (Groves, 1989) in survey research, is clearly carrying over into comparative survey research, despite the challenges this involves (cf. Chapter 13, this volume). Obviously, the comparability dimension has a different meaning in a 3M context than in a national survey and could replace the TSE paradigm as the main planning criterion in such a context, as Figure 1.1 also suggests.

Jowell (1998) remarked on quality discrepancies between standards maintained in what he called “national” research and the practices and standards followed in cross-national research. Jowell’s comments coincided with new initiatives in the International Social Survey Programme to monitor study quality and comparability (cf. Park & Jowell, 1997a) as well as the beginning of a series of publications on comparative survey methods (e.g., Harkness, 1998; Saris & Kaase, 1997) and the development of a European Science Foundation blueprint (ESF, 1999) for the European Social Survey (ESS). The ISSP and the ESS have incorporated study monitoring and methodological research in their programs; both of these ongoing surveys have also contributed to the emergence of a body of researchers whose work often concentrates on comparative survey methods.

Particular attention has been directed recently to compiling guidelines and evidence-based benchmarks, developing standardization schemes, and establishing specifications and tools for quality assurance and quality control in comparative survey research. The cross-cultural survey guidelines at <http://www.ccsq.isr>.

umich.edu/ are a prominent example. Numerous chapters in the volume treat such developments from the perspective of their given topics.

Initiatives to improve comparability and ensure quality are found in other disciplines too. The International Test Commission has, for example, compiled guidelines on instrument translation, adaptation, and test use (http://www.intestcom.org/itc_projects.htm/); and the International Society for Quality of Life Research (ISQoL) has a special interest group on translation and cultural adaptation of instruments (<http://www.isoqol.org/>). The European Commission has developed guidelines for health research (Tafforeau, Lopez Cobo, Tolonen, Scheidt-Nave, Tinto, 2005). The International Standards Organization (ISO) has developed the ISO Standard 20252 on Market, Opinion, and Social Research (ISO, 2006). One of the purposes with this global standard is to enhance comparability in international surveys. We already mentioned the cooperation between Eurostat and national agencies on the European Statistical System (http://epp.eurostat.ec.europa.eu/portal/about_eurostat/european_framework/ESS/).

New technologies are increasingly being applied to meet the challenges of conducting surveys in remote or inhospitable locations: laptops with extended batteries, “smart” hand-held phones and personal digital assistants (PDAs) that allow transmission of e-mail and data, phones with built-in global positioning systems (GPS), pinpointing an interviewer’s location at all times, digital recorders the size of thumb drives, and geographic information systems (GIS) combined with aerial photography that facilitate sampling in remote regions, to name but a few. It is easy to envision a future when these technologies become affordable and can be used much more widely for quality monitoring in cross-national research.

Both the software tools and the related strategies for analysis have also changed radically for both testing and substantive applications. Statistical applications and models such as Item Response Theory (IRT) and Differential Item Functioning (DIF) have gained popularity as tests for bias, as have, in some instances, Multitrait Multimethod (MTMM) models. The increased availability of courses in instruction, also online, makes it easier for researchers to gain expertise in the new and increasingly sophisticated software and in analytical techniques.

Documentation strategies, tools, and expectations have greatly advanced. One needs only to compare the half-page study reports for the ISSP in the late 1980s with the web-based study monitoring report now required to recognize that a sea change in requirements and transparency is underway. Proprietary and open access databanks help improve consistency within surveys across versions and speed up instrument production, even if the benchmarks for question or translation quality remain to be addressed.

The improved access to data—which itself tends to be better documented than before—is also resulting in a generation of primary and secondary analysts who are better equipped, have plentiful data, and have very different needs and expectations about data quality, analysis, and documentation than researchers of even a decade ago.

Critical mass can make an important difference; the current volume serves as one example: In 2002, regular attendees at cross-cultural survey methods symposia held through the 1990s in ZUMA, Mannheim, Germany, decided to form an annual workshop on “comparative survey design and implementation.” This is

now the International Workshop on Comparative Survey Design and Implementation (CSDI; <http://www.csdiworkshop.org/>). CSDI's organizing committee was, in turn, responsible for organizing the 2008 international conference on Multinational, Multicultural and Multiregional Survey Methods referred to throughout this volume as "3MC" (<http://www.3mc2008.de/>) and were also the prime movers for this present volume. Moreover, work groups at CSDI were the primary contributors to the University of Michigan and University of Nebraska CSDI initiative on cross-cultural survey guidelines mentioned earlier (<http://www.ccsr.isr.umich.edu/>). Finally, although the survey landscape has changed radically in recent years (see Table 1.1), readers not familiar with vintage literature will find much of benefit there and an annotated bibliography is under construction at CSDI (<http://www.csdiworkshop.org/>).

Table 1.1 outlines some of the major developments that have changed or are changing how comparative survey research is conceptualized and undertaken. The abbreviations used in the table are provided at the end of the chapter.

Some of these changes are a natural consequence of developments in the general survey research field. As more modes become available, for example, comparative research avails itself of them as best possible (see Chapter 15, this volume). Other changes are a consequence of the growth in large-scale multipopulation surveys on high-stake research (health, education, policy planning data). The need to address the organizational and quality needs of such surveys has in part been accompanied by funding to allow for more than make-do, ad hoc solutions. Developments there can in turn serve as models for other projects. Finally, the increasing numbers of players in this large field of research and the now quite marked efforts to accumulate and share expertise within programs and across programs are contributing to the creation of a body of information and informed researchers.

1.7 CHALLENGES AND OUTLOOK

3M survey research remains challenging to fund, to organize and monitor, to design, to conduct, and to analyze adequately than research conducted in just one or even two countries. We can mention only a few examples of challenges related to ethical requirements by way of illustration. For example, countries vary widely in official permissions and requirements, as well as in informal rules and customs pertaining to data collection and data access. Heath, Fisher, and Smith (2005) note that North Korea and Myanmar officially prohibited survey research (at the time of reporting), while other countries severely restricted data collection on certain topics or allowed collection but restricted the publication of results (e.g., Iran).

Regulations pertaining to informed consent also vary greatly. American Institutional Review Boards (IRBs) stipulate conditions to be met to ensure respondent consent is both informed and documented. IRB specifications of this particular kind are unusual in parts of Europe, although as Singer (2008) indicates, European regulations on ethical practice can be rigorous. Some European survey practice standards recognize that refusals to participate must be respected, but definitions of what counts as a reluctant or refusing respondent differ

TABLE 1.1. Changes in Comparative Survey Research

Developments	Examples	Effects
Programs & Projects		
Size, number, ubiquity of 3M survey projects <i>See Chapters 25–31</i>	ESS; ISSP; WHO-WMH; SHARE; PISA; PIRLS; PIACC; TIMSS; Barometer survey families; WVS/EVS	Critical mass of (1) researchers engaged in ongoing programs, (2) methods research, (3) organizational experience results in capacity building
Frameworks		
Organizational structures <i>See Chapters 13, 15, 25–31</i>	Top-down, bottom-up models; centralized & decentralized models, specifications and organization	Shared ownership; public access; cumulative knowledge & capacity-building for all involved
Quality assurance & control <i>See Chapters 13, 16</i>	Quality movement (conferences); SHARE central monitoring system	Application of quality frameworks & total survey error paradigm; continuous quality improvement
Question Design & Translation <i>See Chapters 2–9</i>	Deliberate procedures & steps; teams; tailored pretesting; translation protocols based on theory and practice	Improved designs, testing, and versions; improved comparability
Sampling <i>See Chapter 14</i>	Proof for probability samples; notion of effective sample size; sampling expert panels	Cumulative knowledge about efficacy of sample designs & implementations; improved comparability
Data Collection <i>See Chapter 15</i>		
Kinds of information collected in addition to answers	Biomarkers; physical measurements; performances; collateral data to enrich analysis; metadata; paradata	Simpler, more affordable; also on general populations, contributing to prevalence information, policy planning, and so forth.
Where data are collected	Almost everywhere in the world	Theoretically founded selection of populations possible
How data are collected	Emerging, new & mixed modes; C-A modes; responsive designs	Rapid technological advances enhance options for isolated locations; integrated QA/QC; real-time monitoring & responsive designs possible
Documentation <i>See Chapters 16, 17</i>		
Kinds of data; documentation standards; related tools	C-A documentation; integral documentation of survey data, metadata, paradata	Required standards, user expectations and access, & survey practice all change. Documentation better and easier; new research & interventions facilitated

TABLE 1.1. *Continued*

Developments	Examples	Effects
Tools Simple & sophisticated tools for phases of survey life cycle and activities	Sample management systems; CARI; control charts; documentation tools like SMDS; DDI structure & codes; data wizards; translation tools; question databanks; data bank systems for version production; procedural guidelines	Diverse tools enhance implementation and QA/QC; shorter production time; enhanced version consistency and documentation
Analysis <i>See Chapters 20–24</i>	Forms of analysis to test & analyze; training options; soft- and hardware	Tailored analyses possible; more choice of procedures; software improved and more accessible, user know-how and options increased
<i>Access to data</i> <i>See CSDI website</i>	Online; download low cost or free; metadata & paradata included	Worldwide access to better documented & easier to use data & to better software; growth in scientific competition & review
Initiatives <i>See Section 1.6 & CSDI website</i>	CSDI, CCSG, DDI; ZUMA symposia, ESS, ECQOS conferences, ISO 20252; ITC Guidelines	Growth in methodological knowledge & critical mass, initiatives, & evidence-based methods research
<i>User/research expertise</i>	ESSi & EU PACO courses, ESF QMSS seminars; 3MC, ECQOS conferences; MA and PhD degrees	Improved research

across countries and this affects options for “refusal conversion.” Multinational panels or longitudinal studies are a further special case. Challenges met here include variation across countries in stipulations regulating access to documents and to data that would enable a respondent to be tracked from wave to wave.

Although much remains to be done, expectations about quality are rightly increasing and these call for greater efforts at every level. At the same time, tools to help deal with the complexity of 3M surveys are evolving and experience in dealing with large-scale projects is growing. Some of the challenges once seen as basic—such as translation—can be much better handled today, provided projects decide to do so. Even the core issues underlying “translation”—those of meaning, meaning in context, and comparability—are being approached differently (see Chapters 2, 3, 10, and 11, this volume). In numerous areas, research from other disciplines is being brought to bear on surveys. Such cross-disciplinary work typically reveals a multitude of new considerations and may seem to complicate the picture. Much in the way the cognitive aspects of survey methods movement once seemed overwhelming to some, researchers may not relish embracing fields such as discourse analysis, linguistics, sociolinguistics, cultural theories, or content

analysis into question design. Ultimately, however, the expectation is that these will help identify both problems and viable solutions.

At national and supranational levels of funding and research planning, recognition of the need for global research is strong. Recent initiatives at national and international level specifically target training and capacity-building courses, as well as research into how better to deal with cross-national and within-country cross-cultural research (see Table 1.1). Much “national” research is rightly seen as comparative at both subnational and supranational levels (cf. Smith, forthcoming). We are thus cautiously optimistic that it will become easier to find funding for methodological cross-cultural and 3M research than in the past, when funders at the national level were less aware of the national benefits and need for such research.

APPENDIX

Table 1.1 refers to surveys, initiatives, tools, organizations, and other entities often in abbreviated forms which could not be explained within the framework of the table. In the alphabetical list below we provide the full form of abbreviated terms and, as relevant, a few words of explanation or a website for further information.

Abbreviation	Full form	Additional information
C-A	Computer-assisted	As in CAPI (computer-assisted personal interview) and CATI (computer-assisted telephone interview)
CARI	Computer Audio-Recorded Interviewing	
CCSG	Cross Cultural Survey Guidelines	The guidelines posted at http://ccsg.isr.umich.edu/ are designed specifically for cross-cultural research, aiming to support and develop best practice for comparative surveys.
CSDI	The International Workshop on Comparative Survey Design and Implementation	An annual workshop on comparative survey methodology; http://www.csdiworkshop.org/
DDI	Data Documentation Initiative	An initiative to provide a generic structure and codes for documenting surveys; http://www.icpsr.com/DDI/
ECQOS	European Conference on Quality in Official Statistics (e.g., “q2006”)	These conferences are held about every two years, the latest in Finland in 2010.
EFTA	European Free Trade Association	The association was set up in 1960; http://www.efta.int/ .
ESS	European Social Survey	An EU-supported biennial survey which covers between 25 and 30 countries; http://www.europeansocialsurvey.org/
Eurostat	The European Statistical Agency	http://ec.europa.eu/eurostat/

Abbreviation	Full form	Additional information
EU PACO	The European Union Panel Comparability	Training workshops organized in connection with longitudinal cross-national panels; ftp://ftp.cordis.europa.eu/pub/tmr/docs/socecolong970252.pdf
IMF	International Monetary Fund	http://www.imf.org/external/index.htm
ISSP	International Social Survey Programme	An ongoing annual survey on five continents; http://www.issp.org/
ISQoL	International Society for Quality of Life Research	http://www.isoqol.org/
ISO	International Standards Organization	http://www.iso.org/iso/home.htm/
OECD	Organization for Economic Co-operation and Development	http://www.oecd.org/
PISA	Programme for International Student Assessment	This international triennial survey began in 2000; http://nces.ed.gov/surveys/pisa/
QA	quality assurance	
QC	quality control	
SHARE	Survey on Health, Ageing and Retirement in Europe	http://www.share-project.org/
TIMSS	Trends in International Mathematics and Science Study	http://timss.bc.edu/
LSMS	Living Standards Measurement Study	The survey was set up by the World Bank in 1980; http://www.worldbank.org/lms/
MIS	Malaria Indicator Survey	http://www.searo.who.int/EN/Section10/Section21/Section1365_11100.htm/
MTUS	Multinational Time Use Study	This survey started in the early 1980s; http://www.timeuse.org/mtus/
ITC	International Test Commission	http://www.intestcom.org/
UNESCO	The United Nations Educational, Scientific and Cultural Organization	http://www.unesco.org/
	UNESCO Statistics	http://www.uis.unesco.org/
WHO	World Health Organization	http://www.who.int/en/
WMH	World Mental Health Initiative Survey	A global survey begun in the late 1990s http://www.hcp.med.harvard.edu/wmh/
WFS	World Fertility Survey	An international survey in 41 countries in the late 1970s–early 1980s.
ZUMA	The Center for Survey Research and Methodology (Zentrum für Umfragen, Methoden, und Analysen), Mannheim, Germany.	Now part of GESIS-Leibniz Institute for the Social Sciences.
3MC	International Conference on Multinational, Multiregional, and Multicultural Contexts	http://www.3mc2008.de/
3M	Multilingual, multicultural, multi-national	A more economic way to refer to multilingual, multicultural, multi-national surveys.