SSOAR

Open Access Repository
www.ssoar.info

What's so special about cross-national surveys?

Lynn, Peter; Japec, Lilli; Lyberg, Lars

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Lynn, P., Japec, L., & Lyberg, L. (2006). What's so special about cross-national surveys? In J. Harkness (Ed.), *Conducting cross-national and cross-cultural surveys : papers from the 2005 meeting of the international workshop on Comparative Survey Design and Implementation (CSDI)* (pp. 7-20). Mannheim: GESIS-ZUMA. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49126-2

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

WHAT'S SO SPECIAL ABOUT CROSS-NATIONAL SURVEYS?

PETER LYNN[1], LILLI JAPEC & LARS LYBERG

1 Introduction

1.1 Objectives and Structure of this Paper

For the past fifteen years we have seen an increase in the number of studies comparing countries, regions, and cultures on different dimensions. This is a trend that we predict will continue. As a consequence, we are likely to see even more cross-national surveys in the future. In this paper we use the term cross-national surveys to represent 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.

Cross-national surveys differ from national surveys in a number of important ways. In this paper we will highlight some of the differences and discuss the implications for survey design and implementation. A key difference is the goal of cross-national surveys, since the focus is on comparability. This requires careful consideration of what is meant by comparability between countries (Van de Vijver, 2003; Johnson, 1998). We thus begin with a discussion of the objectives of cross-national surveys (section 2). We then organise our description of the important aspects of cross-national surveys into unique aspects of survey design (section 3), aspects related to survey planning and organisation (section 4), variation in national constraints (section 5), magnification of cultural issues (section 6), unique aspects of standardisation (section 7), control of design and implementation (section 8) and analysis and reporting (section 9). We conclude by discussing how the survey community might progress from this initial identification of issues that are particular to cross-national surveys towards a situation where we are better equipped to design and implement surveys in ways that are more likely to result in comparable data and analyses.

1 We are grateful to workshop participants for helpful comments and discussions.

1.2 Examples of cross-national surveys

Cross-national surveys cover a wide range of topics from economic indicators to literacy. Examples include: the European and World Values Surveys (EVS, WVS), the Afro barometer, the Eurobarometer, the Latinobarometer, the Harmonized European Time Use Survey (HETUS)/ Multinational Time Use Study (MTUS), the Harmonized Indices of Consumer Prices, the International Social Survey Programme (ISSP), the European Community Household Panel Survey (ECHP), EU Statistics on Income and Living Conditions (EU-SILC) and the European Social Survey (ESS). Links to websites for the cross-national surveys discussed in this section are provided in the annex.

Cross-national surveys also vary greatly in the methods used to achieve comparability across countries and cultures. Some surveys specify only a very limited set of survey design constraints. Perhaps one extreme is where only the question wordings are specified (and only in one language, with no guidance on translation), along perhaps with a minimum sample size. Other aspects of survey design are left entirely to the discretion of the participating institution in each country. The WVS takes an approach similar to this. There are clear advantages to such minimal specification. The workload on the central organisation is minimised and the likelihood of countries being able and willing to participate may be maximised.

Also close to this extreme we have European official statistics that cover many countries, cultures and languages with minimal specification. The need for statistics within the EU was recognized early on and Eurostat was set up in 1951, as part of the European Commission, to provide the EU with statistics from the member states. National Statistical Institutes (NSI) in each country carry out surveys and then Eurostat harmonizes the statistics to make them more comparable across countries. Johnson (2003) defines harmonization as

> "a method for equating conceptually similar but operationally different variables that are collected as part of separate surveys fur purposes of cross-cultural or cross national research. Also referred to as 'ex post harmonization'."

The harmonization efforts include activities such as working group meetings with survey representatives from member countries discussing issues concerning their specific survey. Discussion topics typically include definitions, standards and guidelines. Sometimes the harmonization efforts result in regulations that state, in more detail, what data should be delivered to Eurostat e.g., there is a number of regulations for the Harmonized Indices of Consumer Prices. The survey design, methods and data collection are usually left to each NSI to decide upon. The result of this is that different methods are used in a survey e.g., the EU Labour Force Survey (LFS) is carried out as a face-to-face survey in most countries but in some countries it is a telephone survey.

In a continuing national survey we would be very hesitant to change the survey methodology, e.g., the data collection mode, without a methodological study to assess the effects of the change. The fact that the choice of methods can have an impact on survey results is well documented in the survey literature. However in international surveys, especially where harmonization is the main method for making statistics comparable, we often assume ignorable methodological effects despite the fact that methodological studies showing ignorable effects are very rare. EU-SILC is a recent development where the attempts to achieve comparability go beyond the traditional harmonization efforts to include guidelines on how data should be collected.

Another area where a number of cross-national surveys are carried out is in the field of literacy and learning. Examples of surveys are Trends in Maths and Science Study (TIMSS), Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS) and International Adult Literacy Survey (IALS). These surveys are all deliberately designed to achieve comparability between countries. The IALS measures adults' literacy skills given their formal schooling and reading practices in daily life (OECD 1997). An extensive effort is put into developing tests in each country so that the tests become comparable. However less effort is put into other parts of the design. For example, sampling, estimation, and the implementation of the survey are left to each country.

The European Social Survey (ESS), which collects data on, e.g., trust in politics and religion, satisfaction with government, schools and health care, is at the other end of the scale of cross-national surveys in terms of specification and control. Large efforts are made to achieve comparability. This includes not only detailed specification of design and implementation, but also processes of control, guidance and liaison. Some examples of measures taken in the ESS to achieve comparability are the development of guidelines, creation of a central sampling team that helps out with sampling and estimation in all participating countries, organisation of a translation taskforce that provides guidance on translation procedures, insistence that face to face interviews are carried out in all countries, that interviewers provide information about the interview, that call record data are collected in a standardised way, and that information is collected about issues discussed in the media in each country during the actual data collection period (Stoop, 2003).

2 Objectives of Cross-National Surveys

Goals in cross-national surveys are different from goals in national surveys. Even though data from cross-national surveys are used on national levels, one main use is to compare countries and regions (groups of countries) on different dimensions. The types of comparisons typically made are (1) comparisons of estimates of parameters for different countries or regions or (2) rankings of countries on different dimensions. The parameters in question could be averages, totals, measures of distribution, or measures of association between variables such as model coefficients, for example. Another use of data from cross-national surveys is simply (3) to aggregate estimates from a number of countries in order to provide an estimate relating to a supra-national region such as, for example, the European Union. Each of these three types of use of cross-national survey data obviously requires data that are comparable.

We can not expect meaningful comparisons if we do not take into account the specifics of the countries. To illustrate this we build on an example from Braun & Mohler (2003) about divorce rates in Europe. If the researcher ranks countries (2) on this dimension, Ireland will end up having a very low divorce rate compared to other countries in Europe. The conclusion that should be made from this is not necessarily that the Irish people are happier in their marriages than other Europeans. Instead the low divorce rate can be explained by the fact that the religion in Ireland does not allow divorce but does allow separation. It might be less obvious that we also need to think of comparability when comparing aggregates for a region (3). To illustrate this we continue our example of divorce rates. If we compare divorce rates in Europe with divorce rates in North America, we have to aggregate data for countries within each of these two regions. The same issue of comparability is still present, i.e., we need to make a decision on how to handle separation and divorce rates in these regions, and whether we should aggregate only divorce rates or include separation rates in our measure.

In other words, for cross-national surveys country is the key variable that defines analysis domains or acts as a covariate. All surveys have some important analysis domains and it is always important to ensure comparability between domains. But the difference with cross-national surveys is that one variable is always far more important than the others, and it is always the same one: country. This is the reason that we should pay particular attention to comparability *between* countries, though of course we should additionally pay attention to comparability between sample elements (and therefore groups thereof) *within* countries.

Van de Vijver (2003) discusses different aspects of comparability and equivalence. He points out that cultural specifics in comparative studies constitute a potential problem if care is not taken to distinguish these from more universal aspects. This means that we

need to ascertain which aspects are universal and which are culture specific. Having done that, we should aim to develop measurement tools that provide unbiased estimates of the *relevant* concepts. This requires valid and reliable instruments (*realism*) and unbiased *representation* with respect to the population distributions of all survey measures. The latter can only be achieved by avoiding coverage, sampling and non-response bias. Additionally, it is necessary to be able to estimate the variance of survey estimates. Meaningful comparisons between countries cannot be made if the precision of estimates is unknown. We do not develop further here the discussion of how equivalence might be defined, but recognition of the importance of the concepts of relevance, representation and realism should inform identification of relevant issues in the sections that follow.

3 Unique Aspects of Survey Design

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. The first crucial component is to decide which countries are included. For some surveys, this is relatively uncontroversial, such as with European official statistics which must cover all EU member states. But even here there can be discussions about the inclusion of accession and candidate countries and other important European countries such as Switzerland, Norway and Iceland. Some surveys impose no geographical constraints on participation (e.g. WVS), or no constraints within a particular region (e.g. ESS, Afrobarometer), with participation often dependent on national funding, capacity and enthusiasm. Others identify a limited set of countries for inclusion, perhaps based on a combination of substantive interests and practical constraints. However, the set of countries included is important as it can have great influence on findings, for example regarding between-country variation, or where regions of the world or other supra-national aggregates are to be compared based on only a sample of the countries in each region or aggregate (e.g. see Brown et al., forthcoming).

The second component of design that is unique to cross-national surveys is the choice of how to distribute the sample over countries. Often, reflecting the recognition of countries as key analysis domains and between-country differences as key estimates, the choice is to aim for equal sample sizes in each country. In fact, the precision of such estimates is likely to be maximised by attempting to achieve approximately equal *effective* sample size per nation, where this takes into account the differential effects of design features such as sample clustering, stratification and variation in selection probabilities on precision. (Although this is of course by no means an easy thing to control – Lynn et al., 2004.) Some surveys, such as ESS and EU-SILC, specify national sample size requirements in terms of effective sample size.

Alternatively, some analysts may prefer to view nation as an important covariate, the aim of comparative analysis being to identify the extent and nature of variation that is explained by nation. From this perspective, it can be argued that it is the total variation in the group of nations under study that is the dependent variable and that the sample should be distributed in proportion to each nation's contribution to the total variation. The latter perspective is one of identifying variation over nations in general, rather than differences between specific nations. These two perspectives provide very different design conclusions if the nations under study vary greatly in size (as they typically do).

Whichever perspective is used, a further possible refinement is that important analysis domains may be identified within some or all countries and that sample size requirements may be specified with respect to those domains. If those domains are geographical, there may be more of them in some countries than others, leading to variation in national sample sizes where no variation would have been implied by the requirements for national estimates alone. EU-SILC is an example of a survey where this kind of refinement has been applied. However, the question of sample size per country is one that is often answered on pragmatic or resource grounds, rather than in reaction to the statistical requirements of the survey.

The third component of design that has special characteristics in the case of cross-national surveys is the identification of meaningful relevant concepts and items to study. The concepts and items must be relevant cross-nationally and are consequently often not the same ones that would be chosen for a national survey on the same topic. Both the criteria for choosing concepts and items and the processes by which the choices are made are typically very different than in the case of national surveys.

4 Survey Planning and Organisation

It is rare that cross-national surveys are carried out by a single survey agency. Typically, different survey organisations will carry out the field work in each nation, each with different ideas and working practices. It is therefore necessary for there to be an entity in charge of the entire planning and implementation phase, separate from the survey agencies. The responsibilities of such an entity can vary from simply providing instructions to individual countries to actually assisting and supervising the individual survey organizations involved (see section 8). Experience tells us that the involvement from such a central entity must be extensive if equivalence is to be achieved. Otherwise national considerations will sometimes take precedence over survey design requirements, and there will be an increased risk of simple misunderstanding of the instructions.

Ambitious cross-national studies are enormous undertakings. The design, quality control, analysis, and dissemination are on completely different levels as compared with a national survey. This has implications for almost every stage of the survey process. Considerable resources are required and skillful resource allocation is needed. Communication and liaison between multiple agencies (who will often have different working languages) also becomes a major factor in the success of a cross-national survey.

5 Variation in National Constraints

Methodological and financial resources often differ extensively between countries, as can the experience and capabilities of survey agencies. This can impose considerable constraints on the nature and extent of standardization that can be insisted upon or achieved. At the margins this can lead to trade-off decisions between relaxing one or more requirements for a particular country (with the disadvantage of possibly lesser comparability, as well as the risk of setting a precedent for permitting relaxations of the rules) or forcing the exclusion of the country from the survey (with the disadvantage of not being able to include the country in any analyses).

Legal aspects might affect what it is possible to achieve in a given country. As laws are inherently national (or even in some cases sub-national), this creates differences between countries in the types of design and implementation features that are possible. Relevant laws can include those which relate to data linkage, access to potential sampling frames or auxiliary data, informed consent, rights of privacy and freedom from harassment or unsolicited approaches.

There can be big differences between nations in sampling constraints, such as the nature of available sampling frames and legislation governing their use. This inevitably leads to a choice between permitting and dealing with controlled variation in sample design or imposing a lowest-common-denominator sample design suitable for use in all participating countries, but probably not the best possible approach in many countries.

6 Magnification of Cultural Issues

There is always a variation among individuals in a population regarding their perceptions of questions and concepts. This variation is due to differences in education, experience, cultural norms, language, sensitivity to topic and other personal traits. What happens is that the outcome of the response process is affected by this variation and questions must be worded and administered in such a way that this variation is minimized. In cross-national surveys, in addition to this within-country variation there is also between-country

variation reflecting more profound, systematic, differences between countries. Examples of such differences are that the meaning of concepts, topic sensitivity, perception of response scales, interviewer-respondent interaction, survey climate, computer access, and telephone penetration rate might vary considerably between countries. Basically one has to develop a questionnaire and a mix of data collection modes that is so robust that differences of this kind do not have a devastating effect on the comparability. In essence this means that questions should have the same meaning and be interpreted in a similar fashion across countries (and modes, when necessary).

In addition, there is typically no longer one dominant culture and language amongst the target population. Instead, several very different cultures may be of approximately equal importance to the survey (in terms of their representation in the sample(s) and impacts on survey estimates). And differences in culture (and language) make it impossible to use identical instruments in each country. In consequence, developing and testing cross-national questionnaires becomes a very complex task.

To complicate matters further all survey materials including questions and questionnaires must be translated. This is not only expensive and time-consuming, but can also have a big impact on the measurement characteristics of survey items. There is evidence that traditional translation procedures, such as back-translation, have serious drawbacks and must be replaced by more efficient and effective procedures.

7 Unique Aspects of Standardisation

Typically, national surveys are thought of as having a single “design”. By this, we mean that one sample is selected (to a specified design), one set of survey instruments are used (with a particular design), the field work is organised as one unified operation, and so on. Standardisation is applied (or at least strived for) at several levels. Each sample element receives a standard treatment to encourage participation, is administered a standard questionnaire in a standard way, by interviewers who have received the same training and instructions, and so on. With cross-national surveys, there are several reasons why this is not the case.

These include the variation in national constraints discussed above in section 5, and differences in language and culture that make it impossible to use identical instruments in each country. There is therefore pressure for differences in survey design (Lynn, 2003). It is rarely possible to ensure that every aspect of survey design is replicated in identical fashion in every nation. Neither is it necessarily desirable to do this. Rather, the challenge is to identify which aspects of design need to be identical, which should be allowed (encouraged) to vary – and within what parameters – and which may be less important, in the

sense that relevant characteristics of the survey data may be insensitive to variations in design. At present there is insufficient research evidence to provide much guidance on this question. We simply do not know the relative effects of variation in different aspects of design on the equivalence of cross-national data.

If the objective of a survey is to achieve equivalence in each nation (Jowell, 1998), for some aspects of design this may best be achieved by standardisation of design. For example, equivalence of measurement may best be achieved by administering an item in the same mode, using the same wording and same visual stimuli in every nation (though it should not be assumed that this will always be the case). On the other hand, there are some dimensions for which equivalence may best be achieved by conscious variation in design. For example, to achieve representation of an equivalent population by a sample of equivalent precision, may require the use of different sampling frames and different selection methods in different nations (Lynn et al., 2004; Adams & Wu, 2002 chapter 4). However, even in these two areas of measurement and sampling there is a paucity of knowledge about the effects of different design options on equivalence for different types of survey measures. For other aspects of survey design and implementation, such as field work practices, almost nothing is known.

The extent to which it is possible to standardise the design may depend as much on the infrastructure and processes for coordination and control (see section 8 below) as on statistical considerations. For example, with respect to measurement instruments it may be recognised that the instruments should ideally be pretested in each participating nation and in each language in which they will be administered. And that the findings of the prestests should be discussed collaboratively and collective decisions made about revisions to the instruments. But resource and logistical constraints may prevent this approach from being implemented, whereas other design aspects may be more standardised because it is easier to do so, even though they may be less important.

8 Control of Design and Implementation

On any survey, the design and the data collection and processing procedures – once decided and agreed – require control to ensure that they are carried out as intended. Survey organisations have processes and mechanisms by which to implement this control as routine, often involving regular reporting and liaison with the single survey funder. But with a cross-national survey there are typically several survey organisations and a separate co-ordinating organisations, and possibly other funding organisations too. This requires special control procedures. Without them, there is a strong risk that the design and implementation plan will not be followed adequately.

Different models have been used for the co-ordination and control of survey implementation. At one extreme, it is possible to set up a large central co-ordination team who are able to liaise closely with, and monitor the activities of, each national team throughout the implementation process. At the other extreme, the central co-ordinator may simply issue some written instructions on implementation and leave each nation to follow the instructions. The first of these two models is obviously resource intensive and requires substantial funding for the central activities.

In practice, the co-ordination model adopted may be influenced by the nature of survey funding. If each nation is funding its own participation, there is likely to be pressure for the national funds to be spent on country-specific activities, rather than contributing to a central pot. There may or may not be a separate funding source for centralised activities. However, the choice of co-ordination model is likely to have important implications for the degree of supervision and quality control that is possible. For example, the analysis of regular process data during the field work period is an important tool for monitoring adherence to certain aspects of the survey specification. If the central team is unable to perform this analysis, gross failures to adhere to the specification (which could be either wilful or accidental) may not be identified until long after the survey is completed, if ever (e.g. Park & Jowell, 1997).

As with any survey, the monitoring and control process should feed back into the survey process in a cycle of continuous quality improvement wherever this appropriate. The distinctive aspect of cross-national surveys is that this is likely to involve multiple agencies and be a relatively slow process.

9 Analysis and Reporting

The “distance” between the user and the producer is obviously greater in cross-national surveys compared with national surveys. This puts additional pressure on developing adequate documentation as the analyst will often be reliant upon the documentation to understand relevant aspects of the design and implementation.

Local knowledge (familiarity with national circumstances and contexts) is crucial as input both in the design phase and in the analysis phase. It is particularly difficult to draw informed inference from cross-national comparisons without an even and informed knowledge of the relevant context in each nation. This can be very difficult to obtain and typically requires collaboration of researchers in several nations.

There is a risk that some participating countries might challenge the results from cross-national surveys. One reason is that in ranking studies the outcome is inevitably that some

countries will be the lowest ranked. In other comparative studies specific countries might be described in a less than favorable light. National pride is obviously at stake and the typical scenario is that representatives from countries that do not do well in the studies question methods or survey organizations involved. Therefore analysis and documentation are especially crucial in cross-national surveys in order to demonstrate the scientific credentials of the survey.

10 Concluding Discussion

Much methodological development regarding cross-cultural survey methods (Harkness et al., 2003) concerns question development and translation, how to achieve equivalence in different dimensions, how to perform secondary analysis, and the effect of cultural bias. As mentioned these issues become very complex in cross-national survey settings. There is considerable activity in these methodological and conceptual areas. However, other aspects are receiving less attention and here we end our discussion by adding a few thoughts on methodological issues that might benefit from some research in a cross-national survey context.

As mentioned (section 2), there is no definition of the ideal design properties of cross-national surveys. Without an approximate definition of ideal design properties it is very difficult to know how best to allocate resources. It seems as if input harmonization, as we know it, can take us only so far. And post hoc harmonization is even more limited. We need to develop a cross-national survey methodology that emphasizes management issues and other issues that become especially problematic in cross-national surveys. The survey community is currently learning how to manage large cross-national surveys such as the European Social Survey. It would be useful to assemble the management experiences from recent efforts to form a consensus on how to run these kinds of surveys. Without decent management we will not get decent measurements.

Based on the ideal design properties, we then need to establish design principles that can guide us in the planning and trade-off situation. Advancing input harmonization from merely design specifications to a management system that includes planning, cooperation, training, quality control, assistance, analysis, documentation and continuous improvement seems to be a logical next step. The literature on so-called house effects shows that sometimes survey organizations sharing a common data collection effort produce significantly different results (Smith, 1982). The encouraging fact is that the same literature also point to data collections where no house effects have been discovered (Cohen & Potter, 1990), which means that it is possible for individual survey organizations to sometimes collect comparable data. Cross-national surveys resemble a situation with a potential for exten-

sive house effects, i.e., survey organizations in different countries are supposed to perform identical tasks in such a way that result differences between countries cannot be attributed to the survey organization's standard practices and performance. Thus, these house effects should be minimized.

Translation of survey materials is emerging as a discipline in its own right. Current standard practices such as back-translation seem to be problematic. A more efficient translation procedure is to use team translation. Guidelines need to be developed, implemented and disseminated.

It is widely agreed that good product quality can only be achieved by using underlying processes that are free from unnecessary variation, i.e. stable and predictable. One way of keeping track of the processes is through the choice, measurement and analysis of key process variables (Couper & Lyberg, 2005). Cross-national survey processes need specific sets of process data so that the complex production processes can be controlled. Also it might be worthwhile to explore the concept of responsive designs (Groves & Heeringa, 2004). Here process data are used to inform cost and quality tradeoff decisions in real time. The ability to monitor both process data and regular survey data creates an opportunity to alter the design during the survey data collection in order to improve survey cost efficiency and achieve more accurate estimates.

Any survey, national or cross-national, faces problems regarding general design issues, how to allocate resources effectively, how to perform quality control activities, and how to analyze and disseminate the results to the users. There is no comprehensive theory of survey planning and management and most survey work is based on a mixture of theories, design principles resulting from experiments and experiences, and constraints in terms of costs, time and other user demands. Of course, in a cross-national survey these general problems remain and new ones appear as well. We have attempted to identify in this paper the likely sources of such new problems. However, we know relatively little about their effect on comparability. The identification of such properties is a research priority. When more is known about the relationship between the various unique characteristics of cross-national surveys identified in this paper and between-country comparability, the research community will be in much stronger position to understand how best to design and implement cross-national surveys.

References

Adams, R., & Wu, M. (Eds). (2002). *PISA 2000 Technical Report.* Paris: Organisation for Economic Co-operation and Development.

Braun, M., & Mohler, P. (2003). Background Variables. In J. A. Harkness & F. J.R. Van de Vijver (Eds.), *Cross-Cultural Survey Methods*. New York: Wiley.

Brown, G., Micklewright, J., Schnepf, S., & Waldman, R. (forthcoming) *Cross-national surveys of learning achievement: How Robust are the Findings?*

Cohen, S., & Potter, D. E.B. (1990). Data Collection Organization Effects in the National Medical Expenditure Survey. *Journal of Official Statistics, 6*(3), 275-293.

Couper, M., & Lyberg, L. (2005). *The Use of Paradata in Survey Research.* Invited paper presented at the 56th ISI Session, Sydney, April.

Groves, R. M., & Heeringa, S. G. (2004). *Responsive Design for Household Surveys: Tools for Actively Controlling Survey Nonresponse and Costs.* University of Michigan, unpublished paper.

Harkness, J., Van de Vijver, F. J.R., & Mohler, P. Ph. (2003). *Cross-Cultural Survey Methods.* Wiley.

Johnson, T. (1998). Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research. In J. A. Harkness (Ed.), *ZUMA-Nachrichten Spezial No. 3. Cross-Cultural Survey Equivalence.* Mannheim: ZUMA.

Johnson, T. (2003). Glossary. In J. A. Harkness & F. J. R. Van de Vijver (Eds.), *Cross-Cultural Survey Methods*. New York: Wiley.

Jowell, R. (1998). How Comparative is Comparative Research? *American Behavioural Scientist, 42*(2), 168-177.

Lynn, P. (2003). Developing Quality Standards for Cross-National Survey Research: five approaches. *International Journal of Social Research Methodology, 6*(4), 323-337.

Lynn, P., Gabler, S., Häder, S., & Laaksonen, S. (2004). *Methods for Achieving Equivalence of Samples in Cross-National Surveys.* Working Paper 2004-09 of the Institute for Social and Economic Research. Colchester: University of Essex.

OECD (1997). *Literacy Skills for the Knowledge Society.* Further Results from the Adult Literacy Survey.

Park, A. & Jowell, R. (1997). *Consistencies and Differences in a Cross-National Survey.* London: SCPR.

Smith, T. (1982). House Effects and the Reproducibility of Survey Measurements: A Comparison of the 1980 GSS and the 1980 American National Election Study. *Public Opinion Quarterly, 46,* 54-68.

Stoop, I. (2003). *Collecting Event Data*. Presentation at the Comparative Survey Design and Implementation Workshop in Brussels. (http://www.gesis.org/Mitarbeiter/zuma/harkness/csdi)

Van de Vijver, F. J.R. (2003). Bias and Substantive Analysis. In J. A. Harkness & F. J.R. Van de Vijver (Eds.), *Cross-Cultural Survey Methods* (pp. 207-233). New York: Wiley.

Links to websites

Afrobarometer. http://www.afrobarometer.org/

Eurobarometer. http://europa.eu.int/comm/public_opinion/index_en.htm

European Community Household Panel Survey. http://epunet.essex.ac.uk

European Official Statistics. Eurostat. http://europa.eu.int/comm/eurostat/

European Social Survey. http://www.europeansocialsurvey.org/

European Value Survey. http://www.europeanvalues.nl/

International Adult Literacy Survey. http://www.statcan.ca/english/Dli/Data/Ftp/ials.htm

International Social Survey Programme. http://www.issp.org

Latinobarometer. http://www.latinobarometro.org/

Multinational Time Use Survey. http://www.iser.essex.ac.uk/mtus/

Programme for International Student Assessment. http://www.pisa.oecd.org/

Trends in Maths and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). http://timss.bc.edu/

World Values Survey. http://wvs.isr.umich.edu/