1

The Promise and Challenge of 3MC Research

Timothy P. Johnson[1], *Beth-Ellen Pennell*[2], *Ineke A.L. Stoop*[3], *and Brita Dorer*[4]

[1] *Survey Research Laboratory, University of Illinois at Chicago, Chicago, IL, USA*
[2] *Survey Research Center, University of Michigan, Ann Arbor, MI, USA*
[3] *The Netherlands Institute for Social Research (SCP), The Hague, The Netherlands*
[4] *GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany*

1.1 Overview

Life in the twenty-first century becomes more interconnected daily due in large measure to increasingly complex and reliable communication and transportation networks. This growth in connectivity has also led to increased awareness and, hopefully, greater understanding and respect for individuals who represent diverse cultures, beliefs, and historical experiences. It is within this context that multinational, multiregional, and multicultural survey research, what we refer to as 3MC research, has developed over the past several decades. In addition to basic respect for human diversity, 3MC methods emphasize the importance and address the comparability of survey data across nations, regions, and cultures. These methods represent an evolution of survey methodology away from opportunistic ad hoc international data collection and analysis activities toward more coordinated efforts in which the nations, regions, and cultures of interest have equal representation and share equal responsibility for study planning and leadership.

Although precursors to 3MC research date back to the immediate post-WWII era (see Smith [1], for a brief history of international survey research), the development and expansion of the 3MC research model became possible only more recently. The advent of formal training programs such as the Summer Institute in Survey Research Techniques at the University of Michigan and the founding of international collaborations such as the International Social Survey Programme (ISSP) and the European Social Survey (ESS) that

Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC), First Edition. Edited by Timothy P. Johnson, Beth-Ellen Pennell, Ineke A.L. Stoop, and Brita Dorer.

emphasized comparative fieldwork methods [2] in particular were important precursors. These programs enabled worldwide dissemination of the methodological skills and expertise that would provide a foundation for successful 3MC efforts. Recent technological innovations, many of which are discussed in this volume, have also contributed to the growth and viability of 3MC research across diverse social, political, economic, and physical environments.

A unique contribution of 3MC research is the opportunity it represents to generate comparative knowledge that enhances human understanding and cooperation. Nations and cultural groups that have been historically ignored by the empirical social science community have found opportunities to participate and be represented in 3MC activities. 3MC research has also led to increased development and sharing of methodologies for conducting survey research in international and cross-cultural environments. Evidence for this comes in the form of the annual meetings (since 2003) of the Comparative Survey Design and Implementation (CSDI) (https://www.csdiworkshop.org/) workshop, which focuses on the sharing of innovative methods and strategies for comparative research. It also comes in the form of larger international conferences designed to showcase achievements in the field. The first of these meetings was held in 2008 in Berlin, with a second meeting held in Chicago in 2016.

This is the third volume in the Wiley Series in Survey Methodology that focuses specifically on 3MC research practice. Although the 3MC acronym was first introduced in the 2010 volume [3, 4], the same concerns with multinational, multiregional, and multicultural research were clearly also present in the earlier volume edited by Harkness et al. [5]. We view this current volume as an extension of these earlier works, one that summarizes new 3MC developments over the past decade.

1.2 The Promise

3MC accomplishments have made a rich contribution to our knowledge of best practices for survey methodology, as the advent of work has led to the development of new and modified methodologies. Some of these accomplishments include the now commonly employed questionnaire translation and adjudication protocols first pioneered by Janet Harkness and colleagues [3, 4, 6, 7] and the efforts of Jowell et al. [8] to develop functionally equivalent fieldwork practices. Recent advances in the use of multigroup confirmatory factor analytic modeling for analysis of data from large numbers of nations [9, 10] and the procedures for cross-cultural cognitive interviewing [11] (Chapter 10, this volume) are other examples. Countless additional developments can be found in the 800 pages of the Cross-Cultural Survey Guidelines that are being continuously updated by the University of Michigan (http://ccsg.isr.umich.edu/). In addition, this work supported advancement in the general

field of survey research. The growing availability of large numbers of national-level surveys collected as part of 3MC initiatives, for example, enables for the first time analyses that treat surveys themselves as the unit of analysis, permitting research into basic survey design problems that were not previously possible. Several such examples are presented in this volume. In Chapter 5 of this volume, Koch examines the quality of sample composition across several types of within-household respondent selection procedures using a sample of 153 national surveys conducted across six waves of the ESS. The findings presented make an important contribution to an often overlooked potential source of coverage and nonresponse error. Similarly, Andreenkova (Chapter 14, this volume) examines interview language choice protocols and documentation across multiple comparative projects, providing insights not previously available, and Chapters 43–47 also analyze the quality of comparative surveys across multiple dimensions.

The rapid growth in access to high quality 3MC data over the past several decades has also led to many new opportunities for social scientists to rigorously investigate social and policy relevant issues on a much larger scale than has been previously possible. These accomplishments are evident across a variety of fields and disciplines, including political science [12, 13], sociology [14], economics [15], and mental health [16]. One could make the case that the datasets produced from ongoing 3MC initiatives have led to a renaissance of sorts for empirical social science. It is also possible that a century from now these carefully documented survey archives will provide researchers with an essential resource for understanding our period in history.

1.3 The Challenge

The development and assessment of 3MC methods is of course far from, and will likely never be, complete. At the most basic level, the comparability in meaning and interpretation of measures applied across multiple groups will almost certainly continue to be challenged in many research settings. This message that cultural frameworks do not neatly map onto one another is one that readers will find being continually re-emphasized throughout this volume. Demonstration of construct and measurement comparability by investigators will consequently continue to be a necessity. The ongoing accumulation of evidence across multiple initiatives may, however, lead to new approaches to addressing this old problem.

Another ongoing concern is the continual dominance of English as the source language for many 3MC efforts. Although a practical approach to organizing instrument development activities, this nonetheless accords what many would perceive to be undue amounts of influence to one language and cultural tradition. English is known to have a larger lexicon than any other language, which

means that distinctions in wording in English cannot always be replicated in target languages [17]; in addition, the structure of English as source questionnaire language to be translated into multiple languages is challenging: Its ability to condense much information in few words often requires longer and more wordy target versions; and many target languages need to be more specific, e.g. related to gender, numerus, or terms like "the following," and if this additional information is not provided, comparability between the target versions may be impaired. These concerns related to the source language are rarely expressed but will need to be confronted proactively at some point. This brings to the surface a related issue, as survey research itself remains a Western-oriented social scientific methodology that seems most appropriate for applications within liberal democratic political environments. It is important to be sensitive to the concern that 3MC research may be viewed in some quarters as a form of cultural hegemony. Indeed, to participate in 3MC research, some researchers and respondents must submit to modes of communication that make broad assumptions about the nature of social relationships and self-expression that they may see as nonnormative. Understanding varying perceptions of the meaning of information collected via survey research across cultures thus remains an important challenge.

Sadly, another challenge to 3MC research that must be confronted is the growth in nationalism now being witnessed in many nations. We are concerned that many of the policies that will accompany this ideology may lead to weakened relationships and declining interest in cooperation with cross-national, cross-regional, and cross-cultural populations who will inevitably be defined as out-groups. Competition for resources and economic advantage may also undermine national willingness to participate in international research collaborations that are not viewed as bringing immediate returns on investment. Relatedly, political leaders who are willing to discredit public opinion surveys within their own nations for partisan advantage are unlikely to support broader efforts of the type represented by 3MC projects. Unfortunately, many of the social forces that have led to government cynicism, distrust of official statistics, weakened survey climate, and lower response rates in many Western nations may also be weakening public support for 3MC research. Indeed, history and recent events provide instruction regarding the fragile nature of cross-national and cross-cultural relationships. Ironically, 3MC research is likely to be most necessary during precisely those periods in time when it will be most challenging to undertake.

Another ongoing challenge to 3MC research is the need to further develop its theoretical underpinnings. Currently, much 3MC work is accomplished within the invaluable total survey error (TSE) framework [18]. Although important efforts have been made to integrate 3MC concerns into this paradigm (see Chapter 2 in this volume), a generalizable model of *how* culture influences various survey-related error processes has yet to be established.

Some potentially useful cross-cultural frameworks have been developed in other disciplines (the models of Hofstede [19], Schwartz [20], and Triandis [21] are relevant examples), and a few initial steps have been taken in this direction [22, 23], but we are far from a consensus as to how to best proceed. Looking forward, interdisciplinary collaborations similar to those forged between survey methodologists and cognitive psychologists some 30 years ago [24] might be one productive strategy to consider. Working to establish firm theoretical foundations is an important part of 3MC's future that has yet to be addressed.

1.4 The Current Volume

This volume contains four dozen chapters distilled from the 2016 3MC conference held in Chicago. They are organized into sections that focus on a wide variety of topics relevant to ongoing developments in applied 3MC research. In addition to this chapter, the first section includes a conceptual piece by Tom Smith (Chapter 2) that considers TSE within the context of 3MC research. In doing so, he elaborates on the concept of "comparison error," which we anticipate will become an important element of the 3MC TSE model. Chapter 3, contributed by Jose-Luis Padilla, Isabel Benitez, and Fons J. R. van de Vijver, addresses notions of equivalence and comparability from a mixed methods perspective.

Two chapters examine sampling issues. Chapter 4, by Stephanie Eckman, Kristen Himelein, and Jill Dever, provides insights and examples of the effective use of geographic information system (GIS) technology as part of household sample designs in developing nations. As mentioned earlier, Koch examines various methods of within-household respondent selection and their effects on data quality in Chapter 5.

The section on cross-cultural questionnaire design and testing presents a number of important innovations. Ana Villar, Sunghee Lee, Ting Yan, and Brita Dorer first provide an overview of questionnaire design and testing within the 3MC context (Chapter 6). This is followed by a contribution from Anna Andreenkova and Debra Javeline, who discuss strategies for detecting and addressing differences in question sensitivity in a comparative context (Chapter 7). An online multinational study, designed to re-evaluate a series of classic split-ballot questionnaire experiments previously conducted in monocultural settings, is presented in Chapter 8 by Henning Silber, Tobias Stark, Annelies Blom, and Jon Krosnick. In Chapter 9, Mengyao Hu, Sunghee Lee, and Hongwei Xu discuss the use of anchoring vignettes and provide an empirical example that includes an innovative sensitivity analysis. Cognitive interview methods for evaluating question comparability are reviewed in Chapter 10 by Kristen Miller, and Hyunjoo Park and Patricia Goerman consider best approaches to conducting cognitive interviews with

non-English-speaking respondents in Chapter 11. Patricia Goerman, Mikelyn Meyers, Mandy Sha, Hyunjoo Park, and Alisu Schoua-Glusberg investigate, in Chapter 12, the degree to which monolingual and bilingual cognitive testing respondents are able to identify the same issues with survey questionnaires. The final chapter in this section (Chapter 13), by Timothy Johnson, Allyson Holbrook, Young Ik Cho, Sharon Shavitt, Noel Chavez, and Saul Weiner, investigates the usefulness of behavior coding as a method for comparing the cognitive processing of survey questions across cultural groups.

A section concerned with languages, translation, and adaptation includes four chapters. As mentioned earlier, Anna Andreenkova (Chapter 14) explores available procedures and documentation concerning the interview language selection process in 3MC surveys, a topic that has previously received little attention but has important ramifications for sample coverage, respondent cooperation, and measurement error. In Chapter 15, Emilia Peytcheva reviews the effects of interview language on respondent answers. Dorothée Behr, Steve Dept, and Elica Krajceva discuss the documentation of a sophisticated survey translation and monitoring process in Chapter 16, and Diana Zavala-Rojas, Willem Saris, and Irmtraud Gallhofer consider, in Chapter 17, strategies for preventing differences in translated survey items using the Survey Quality Prediction (SQP) system.

In the following section, three chapters address issues relating to mixed modes and methods within the 3MC context. The first of these is Chapter 18 by Edith de Leeuw, Tuba Suzer-Gurtekin, and Joop Hox, who provide an overview of methods for the design and implementation of mixed-mode surveys. Chapter 19, by Tuba Suzer-Gurtekin, Richard Valliant, Steven Heeringa, and Edith de Leeuw, provides an overview of design, estimation, and adjustment methods for mixed-mode surveys. In Chapter 20, Nathalie Williams and Dirgha Ghimire discuss new technologies for mixed methods data collection in 3MC surveys.

In the next section, another three chapters focus on issues of response style variability across cultures. In the first of these (Chapter 21), Sunghee Lee, Florian Keusch, Norbert Schwarz, Mingnan Liu, and Tuba Suzer-Gurtekin examine the cross-national comparability of response patterns to subjective probability questions. In Chapter 22, Mingnan Liu, Tuba Suzer-Gurtekin, Florian Keusch, and Sunghee Lee compare multiple methods for the detection of acquiescent and extreme response styles. Ting Yan and Mengyao Hu evaluate the effects of translation on respondent use of survey response scales when responding to a generic self-rated health question in Chapter 23.

A large section, containing 10 chapters, explores issues of data collection in 3MC surveys. In Chapter 24, Kristen Cibelli Hibben, Beth-Ellen Pennell, Sarah Hughes, Jennifer Kelley, and Yu-chieh Lin present an informative set of case studies that highlight challenges to cross-national data collection and potential solutions. Data collection challenges specific to sub-Saharan Africa are

discussed by Sarah Hughes and Yu-chieh Lin in Chapter 25. Justin Gengler, Kien Trung Le, and David Howell, in Chapter 26, focus on data collection challenges unique to fieldwork in the Arab Gulf region. In Chapter 27, J. Daniel Montalvo, Mitchell Seligson, and Elizabeth Zechmeister provide a similar overview of their data collection experience in Latin American and Caribbean nations. Issues conducting survey research in India and China are discussed in Chapter 28 by Charles Lau, Ellen Marks, and Ashish Kumar Gupta. In Chapter 29, Nicole Watson, Eva Leissou, Heidi Guyer, and Mark Wooden present best practices for panel maintenance and retention. Luzia Weiss, Joseph Sakshaug, and Axel Börsch-Supan provide an overview of the use of biomarkers and other biometric data in 3MC research in Chapter 30, and Yfke Ongena, Marieke Haan, and Wil Dijkstra discuss the multinational use of event history calendars in Chapter 31. Finally, Julie de Jong provides a broad overview of ethical considerations in the conduct of 3MC research in Chapter 32, and Kirstine Kolsrud, Katrine Segadal, and Linn-Merethe Rød focus on ethical and legal issues surrounding the linking of survey and auxiliary data in Chapter 33.

Three chapters examine quality control and monitoring. Lesli Scott, Peter Mohler, and Kristen Cibelli Hibben discuss the organization and management of 3MC surveys from a TSE perspective in Chapter 34. In Chapter 35, Zeina Mneimneh, Lars Lyberg, Sharan Sharma, Mahesh Vyas, Dhananjay Bal Sathe, Frederic Malter, and Yasmin Altwaijri provide multiple case study examples of best practices for the monitoring of interviewer behaviors in 3MC research. In Chapter 36, Michael Robbins provides an overview of strategies for preventing and detecting falsification in 3MC surveys.

Survey nonresponse is also considered in a separate section containing three chapters. In the first of these (Chapter 37), James Wagner and Ineke Stoop discuss nonresponse and nonresponse bias from a comparative perspective. In Chapter 38, Matt Jans, Kevin McLaughlin, Joseph Viana, David Grant, Royce Park, and Ninez Ponce investigate cultural correlates of nonresponse in the California Health Interview Survey, and Oliver Lipps and Michael Ochsner consider, in Chapter 39, the degree to which offering respondents a greater choice of languages for completing interviews improves, or not, the representativeness of survey samples.

In the next section, two chapters address current advances in the analysis of data from 3MC surveys. In Chapter 40, Deana Desa, Fons van de Vijver, Ralph Carstens, and Wolfram Schulz discuss measurement invariance problems and solutions in international large-scale assessments of educational achievement. In Chapter 41, Kimberley Lek, Daniel Oberski, Eldad Davidov, Jan Cieciuch, Daniel Seddig, and Peter Schmidt present an empirical application of approximate measurement invariance in 3MC research.

Another section examines data harmonization, documentation, and dissemination. An overview of these topics is presented in the introductory Chapter 42 by Peter Granda. This is followed by five chapters contributed

by researchers at the CONSIRT (Cross-National Studies: Interdisciplinary Research and Training) program at the Polish Academy of Sciences and Ohio State University. Chapter 43, by Kazimierz Slomszynski and Irina Tomescu-Dubrow, discusses basic principles of survey data recycling. Data harmonization and data documentation quality in 3MC surveys are discussed by Maria Kolczńska and Matthew Schoene in Chapter 44. The identification of processing errors is discussed in Chapter 45 by Olena Oleksiyenko, Ilona Wysmutek, and Anastas Vangeli. In Chapter 46, Marta Kolczńska and Kazimierz Slomszynski examine the potential usefulness of item metadata as controls for ex post harmonization in cross-national survey projects. In Chapter 47, Marcin Zielinski, Przemek Powalko, and Marta Kolczńska focus on the application of statistical weights in cross-national survey projects.

The final chapter (48) in this volume, by Lars Lyberg, Lilli Japec, and Can Tongur, discusses some prevailing problems in 3MC research and looks forward to the future of comparative survey research. These 48 chapters collectively address both the promise and the challenges of 3MC research.

References

1 Smith, T.W. (2010). The globalization of survey research. In: *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (ed. J.A. Harkness, M. Braun, B. Edwards, et al.), 477–484. Hoboken, NJ: Wiley.

2 Jowell, R. (1998). How comparative is comparative research? *American Behavioral Scientist* 42: 168–177.

3 Harkness, J.A., Braun, M., Edwards, B. et al. (ed.) (2010). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. Hoboken, NJ: Wiley.

4 Harkness, J.A., Villar, A., and Edwards, B. (2010). Translation, adaptation, and design. In: *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (ed. J.A. Harkness, M. Braun, B. Edwards, et al.), 117–140. Hoboken, NJ: Wiley.

5 Harkness, J.A., van de Vijver, F.J.R., and Mohler, P.P. (ed.) (2003). *Cross-Cultural Survey Methods*. Hoboken, NJ: Wiley.

6 Harkness, J.A. and Schoua-Glusberg, A. (1998). Questionnaires in translation. In: *Cross-Cultural Survey Equivalence* (ed. J.A. Harkness), 87–126. Mannheim: ZUMA.

7 Harkness, J., Pennell, B.-E., and Schoua-Glusberg, A. (2004). Survey questionnaire translation and assessment. In: *Methods for Testing and Evaluating Survey Questionnaires* (ed. S. Presser, J.M. Rothgeb, M.P. Couper, et al.), 453–473. Hoboken, NJ: Wiley.

8 Jowell, R., Roberts, C., Fitzgerald, R., and Eva, G. (2007). *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*. Los Angeles, CA: Sage.

9 Davidov, E., Schmidt, P., and Billiet, J. (2011). *Cross-Cultural Analysis: Methods and Applications, Second Edition*. New York: Routledge.

10 Davidov, E., Cieciuch, J., Meuleman, B. et al. (2015). The comparability of measurements of attitudes toward immigration in the European Social Survey: exact versus approximate measurement equivalence. *Public Opinion Quarterly* 79: 244–266.

11 Willis, G. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly* 79: 359–395.

12 Dalton, R.J. and Welzel, C. (2014). *The Civic Culture Transformed: From Allegiant to Assertive Citizens*. New York: Cambridge University Press.

13 Inglehart, R. and Welzel, C. (2005). *Modernization, Cultural Change and Democracy: The Human Development Sequence*. New York: Cambridge University Press.

14 Breen, M.J. (2017). *Values and Identities in Europe: Evidence from the European Social Survey*. New York: Routledge.

15 Blanchflower, D.G. and Oswald, A.J. (1992). *The Wage Curve*. Cambridge: MIT Press.

16 Kessler, R.C. and Üstün, T.B. (2008). *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York: Cambridge University Press.

17 Harkness, J., Pennell, B.-E., Villar, A. et al. (2008). Translation procedures and translation assessment in the World Mental Health Survey Initiative. In: *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders* (ed. R. Kessler and B. Üstün), 91–113. New York: Cambridge University Press.

18 Biemer, P.P. and Lyberg, L. (ed.) (2010). Special issue: total survey error. *Public Opinion Quarterly* 74 (5): 817–1045.

19 Hofstede, G. (2001). *Culture's Consequences* (2). Thousand Oaks, CA: Sage.

20 Schwarz, S. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In: *Advances in Experimental Social Psychology* (ed. M.P. Zanna), 1–65. San Diego, CA: Academic Press.

21 Triandis, H.C. (1996). The psychological measurement of cultural syndromes. *American Psychologist* 51: 407–417.

22 Schwarz, N., Oyserman, D., and Peytcheva, E. (2010). Cognition, communication, and culture: implications for the survey response process. In: *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (ed. J.A. Harkness, M. Braun, B. Edwards, et al.), 177–190. Hoboken, NJ: Wiley.

23 Uskul, A.K., Oyserman, D., and Schwarz, N. (2010). Cultural emphasis on honor, modesty, or self-enhancement: implications for the survey-response process. In: *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (ed. J.A. Harkness, M. Braun, B. Edwards, et al.), 191–201. Hoboken, NJ: Wiley.

24 Jabine, T.B., Straf, M.L., Tanur, J.M., and Tourangeau, R. (1984). *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press.