Cross-national Studies: Interdisciplinary Research and Training Program (CONSIRT),
The Ohio State University and the Polish Academy of Sciences

Volume 3, Issue 2
Fall 2017

Editors
Irina Tomescu-Dubrow
and
Joshua K. Dubrow
CONSIRT

consirt.osu.edu/newsletter
ISSN 2392-0858

Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences

Speak Up, Stand Out

Welcome to the latest issue of *Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences*. We share news and knowledge generated by the ever growing community of scholars, institutions, and government agencies who work on harmonizing social survey data and other projects with similar focus. By speaking up and standing out, scholars from across the world strive to make a positive difference in the international social science community.

This issue features news and articles on a variety of topics. First is news of two grants won by members of the Harmonization Project. The **US National Science Foundation** awarded co-PIs Kazimierz M. Slomczynski, Irina Tomescu-Dubrow and J. Craig Jenkins 1.4 million USD for four years to further research on Survey Data Recycling. Meanwhile, Joshua K. Dubrow was awarded 764,000 PLN for three years by **Poland's National Science Centre** for a project on political and economic inequality in cross-national perspective, featuring the SDR dataset.

Stand-out articles by the international survey data community speak to the diversity of harmonization research in the social sciences. **Tom Emery** writes about how to harmonize administrative data with survey data in the cross-national Generations and Gender Survey; **Ilona Wysmułek** presents the issues on the comparability of corruption perception items in major public opinion surveys, and **Marta Kołczynska, Irina Tomescu-Dubrow** and **Kazimierz M. Slomczynski** discuss issues in harmonizing education across nations and time.

As with every issue of Harmonization, we welcome your articles and news. Please send them to the newsletter co-editor Josh Dubrow at dubrow.2@osu.edu.

In This Issue

Acknowledgements

Editors thank Marta Kołczyńska for technical assistance

News

Survey Data Recycling: New Analytic Framework, Integrated Database, and Tools for Cross-national Social, Behavioral and Economic Research (SDR project) received 4-year NSF funding

by Irina Tomescu-Dubrow, the Polish Academy of Sciences and The Ohio State University

Thanks to generous funding by the US National Science Foundation, our team of social science faculty, computer science experts and graduate students from The Ohio State University (OSU) and the Polish Academy of Sciences (PAN), led by Kazimierz M. Slomczynski, Irina Tomescu-Dubrow and J. Craig Jenkins, has new resources to devote toward our ambitious goal of advancing the interdisciplinary field of ex-post survey data harmonization and the methodology underlying it.[1] The SDR project builds on the long-lasting collaboration that OSU and PAN fostered through CONSIRT (consirt.osu.edu), which is reflected in the dual institutional arrangement of the grant.

We are glad – and thankful – that we have an Advisory Board with a varied expertise to enrich our work. Board members include **Claire Durand**, University of Montreal, WAPOR; **Peter Granda**, University of Michigan, ICPSR; **Dean Lillard**, OSU, CNEF; **Malgorzata Mikucka**, Mannheim University; **Pamela Paxton**, UT Austin; and **Markus Quandt**, GESIS.

The SDR project sets out to (1) build theory-informed 'big data' for the social sciences containing information from extant survey and non-survey sources, and metadata for survey quality and for the harmonization process, for use in comparative, cross-national research; and (2) devise methodological tools for analyzing multi-dimensional data structures stemming from ex-post survey data harmonization.

The project builds on the analytic framework of survey data recycling (SDR). Core to SDR are metadata of two types: measures of source data quality, and control indicators for the procedures of ex-post harmonization. SDR involves ways of constructing, testing and using metadata in statistical analyses. The premise is that researchers can employ survey quality and harmonization control variables to account for methodological biases and errors produced in the harmonization lifecycle, from obtaining the source data (from each national survey), through the harmonization decisions (creation of target variables), to cleaning and checking the resulting data file.[2] In doing this work, we extend the scope of the Harmonization Project, including its database, currently available at the Harvard Dataverse (dataverse.harvard.edu).

[1] PTE Federal award 1738502; Project runs from September 2017 to August 31, 2021.

[2] In this context, methodological biases and errors in survey data are understood as consequences of (a) deviations from standards of documenting and preparing survey data suggested in the specialized literature, and (b) inter-survey differences in harmonized items.

Highlights of the SDR project include:

- Individual-level harmonized measures of social capital, wellbeing, and political participation, and their main socio-demographic correlates derived from 24 international survey projects (3,112 national surveys) covering 3.5 million respondents in ca. 140 countries since 1966;
- Metadata on source data quality and on harmonization procedures;
- Macro variables of demographic, political, and economic characteristics of countries across time;
- Methodology for survey quality assessment and ex-post survey data harmonization;
- Detailed documentation of the harmonization process;
- Website interface for customized data downloading, on-line data analysis and visualization.

The community of users will have free access to the project's main outcomes, including the SDR database and analytics for their exploration. For more information, please visit the project's website, at dataharmonization.org.

Political Voice and Economic Inequality across Nations and Time: A New Research Project

by Joshua K. Dubrow, Institute of Philosophy and Sociology, Polish Academy of Sciences

Poland's National Science Centre has awarded a grant for the project, "Political Voice and Economic Inequality across Nations and Time" (2016/23/B/HS6/03916) for the period 2017 - 2020. The Principal Investigator is Joshua K. Dubrow, Professor of Sociology at the Institute of Philosophy and Sociology, Polish Academy of Sciences, and member of the harmonization research group at IFiS PAN. This project builds on empirical research on how economic resources and political voice connect, accounting for how political institutions moderate this connection.

The purpose of the project is to advance the theory, methods, and empirical base for studying the relationship between political inequality and economic inequality. The fundamental research questions are:

(1) How and to what extent are the main components of political voice inequality – political participation and party representation – related to each other once main features of political and economic institutions are accounted for?

(2) How do changes in economic inequality at the macro-level relate to political voice at the micro-level?

(3) At the macro-level, how and to what extent do political voice inequality and economic inequality influence each other?

The social sciences do not have appropriate cross-national and over-time measures of political voice inequality, which left these questions largely unanswered. To facilitate new knowledge, we will create the Political Inequality Database (POLINQ) with different measures of inequality of political voice for over 65 democratic countries from 1990 to 2015. POLINQ will combine two types of data. The first are cross-national surveys containing individual-level indicators of political participation and support for political parties. They feature substantial coverage of varying types of modern democratic countries to provide variation in the degree of institutional efficiency and measurement points (i.e. national survey years) in order to gauge social and political change since the 1990s. The second type of data contains theoretically-informed characteristics of countries from sources such as the Standardized World Income Inequality Database.

Joshua presented this project to the research unit "Institutions and Political Inequality" (IPI) of WZB Berlin Social Science Center, on November 7, 2017. For more information about the project, please visit the project's website, Politicalinequality.org.

Articles

Generations & Gender Survey: The Integration of Administrative Data and Survey Data

by Tom Emery, Manager of the Generations and Gender Programme

The role of social surveys within social statistics and population registries is changing. The increasing abundance of administrative data has raised questions about the need for, and role of, social surveys within modern data infrastructures.

Social surveys are costly to administer and are increasingly affected by low response rates. As the sophistication of national data is growing, it becomes more and more tempting to view administrative data sources such as population registries, tax records, and health records as alternative source of data, especially when they can be linked together. Yet surveys collect many data that are not included in administrative sources, but are intrinsic to social science research. People's attitudes, values, expectations and beliefs, as well as a broad range of sociological, economic, and health concepts are measured only via surveys and are not available as administrative data. Social surveys are also specifically designed for research and therefore acquire explicit consent for data to be incorporated into open data infrastructures that serve research communities. This consent is lacking for administrative records. To increase the breath of empirical analyses calls for integrating survey data with information from administrative records.

The Generations and Gender Survey (GGS) is one example that illustrates the potential of this approach of drawing on administrative data within the fieldwork. The GGS is a cross-national, longitudinal study that focuses on demographic topics that are intensively covered by population registers, but also covers issues like gender values, peoples' family and relationship plans, and social relationships. It interviews respondents every three years to see how their family and relationships evolve over time. Given this, GGS has sought to incorporate administrative data in its data production process, supplementing the data collected in surveys with high quality and relatively inexpensive administrative data covering births, deaths, marriages, divorces, migration, and labour market status. This approach taken by the Swedish team of GGS serves as illustration.

As the sophistication of national data is growing, it becomes more and more tempting to view administrative data sources as alternative source of data.

The Swedish GGS was fielded in 2013 with 9,688 respondents (response rate = 54.7%). Interviews were conducted via CATI and a follow-up postal questionnaire was used for additional modules. The Swedish GGS successfully linked the sample with a wide range of administrative records that were then integrated into the fieldwork process and the publicly available dataset for use by social scientists in conjunction with data from 18 other countries. The administrative data can then also be used to produce 'virtual' longitudinal waves that mirror the GGS's longitudinal, three-year intervals, but without the issue of attrition. Teams in the Netherlands, Norway, and Belgium have also had success with this approach, as linking is relatively easy.

While integrating surveys and administrative records would increase the efficient use of extant data, this approach also raises some challenges. Specifically, can such linked data be disseminated in the open access manner of a survey research infrastructure? To date, the Swedish GGS is the only publicly accessible file that includes administrative data form population registers.

It is hoped that the new General Data Protection Regulation (GDPR), which will come into force in the European Union in May 2018, will standardise definitions of consent and align practices between survey research infrastructures and administrative data holders. The GDPR has clear guidelines for what constitutes consent and, for the first time, provides an unambiguous framework for such work. However, unlike other public and private organizations, statistical offices are largely exempt from the obligation to share personal data on the request of the data subject. Therefore, statistical offices may increasingly recognize the validity of the consent collected by social surveys but will still be able to avoid linking administrative records.

Still, the GDPR could signal the start of increased integration between administrative and scientific survey data that has hitherto been limited. Doing so will require an integration of administrative and scientific metadata standards and transparency in the processing and versioning

of administrative data linkages. This is especially true given the variability in quality of administrative data and its ability to be linked. Nevertheless, overcoming such challenges has huge potential to streamline and reinvigorate scientific data collections like the GGS and many others.

Tom Emery (emery@nidi.nl) is the Manager of the Generations and Gender Programme (www.ggp-i.org) and on the Board of the Open Data Infrastructure for Social Science and Economic Innovations in the Netherlands (www.ODISSEI-data.nl). His research focuses on family dynamics, survey research methods and scientific research infrastructures.

Measuring Perceptions: Comparability of Corruption Items in Cross-national Surveys

by Ilona Wysmułek, Institute of Philosophy and Sociology, Polish Academy of Sciences

The aim of this research note is to illustrate the practical difficulties in establishing measurement comparability of corruption perception items cross-nationally, based on available survey data. The case of corruption items opens the discussion on how to empirically assess comparability of single-variable indicators for European countries, and on how to apply this assessment in substantive analysis. It raises the question of equivalence in the statistical analysis phase of a study – in contrast to design or implementation stages.

This research note represents work in progress, which originated in Wysmułek's (2017a) dissertation research on corruption in public schools in Europe. Initial insights were presented on the GESIS expert meeting in Mannheim, November 2017 (Wysmułek and Słomczyński 2017). We are very thankful for participants' useful comments and for the fact that the Symposium's organizers allowed plenty of time for discussion.

Approaching Comparability

Comparability of survey instruments, necessary for conducting any cross-national analysis, can be considered along several steps. First, the theoretical construct needs to be similarly understood across all countries included in the research. Second, the instrumental definition of the concept should capture the theoretical construct so that there is no country-specific bias. And third, the measurable construct needs to be translated into the specific questionnaire item(s), which is comparable cross countries (Miller, Slomczynski, and Schoenberg 1981). It includes decisions on question wording and response scale of the master questionnaire item, but also the challenges of translation into multiple languages and assumptions of the comparable reaction to, *inter alia*, the question ordering (Krosnick 2017).

For example, in corruption research, on the level of the theoretical construct it is assumed that the definition of corruption as the misuse of public power for private gains is commonly

understood across countries. On the level of the instrumental definition, corruption is often approached through perceptions, which are assumed to uniformly refer to the amount of corruption that respondents believe exists in a given country. On the level of measurement, it can be assumed that disagreeing or agreeing on the 11-point scale with the statement e.g. *Corruption is prevalent in my area's local public-school system* is understood in the same way in different countries.

Przeworski and Teune (1970) once claimed that they see every questionnaire item in a cross-national survey as an equivalence statement. This can be interpreted that, for a questionnaire item to appear in a cross-national survey, the project's academic team should be confident that the understanding of this item and its components is sufficiently similar across groups. This opinion should be based on tests on between- and within-group validity and comparability of measures proposed in the instrument. In this respect, substantive researchers may treat the available cross-cultural questionnaire instrument and the data it yielded as the end-result of a team work. By publishing the cross-national data, the cross-national survey team makes an unwritten statement that the data can be used for comparable analysis.

Nonetheless, the concern with comparability and equivalence is also pertinent at the data analysis stage. As Jia He and Fons van de Vijver underline in their article "Bias and Equivalence in Cross-Cultural Research" (2012), the question of comparability should be put not only at the design and implementation stages of survey life-cycle, but also when analyzing the secondary data, at the statistical analysis stage. These authors, among many others, see all three stages as crucial to minimize bias and achieve equivalence.

There are two dominant types of questions on corruption: measuring corruption perception or experience of giving bribes/gifts.

At present, research on measurement comparability advances in the direction of developing methods of testing whether a set of indicators reflects the expected concept/construct. According to Billiet (2016, 2017), such approach allows for the interplay between the theoretical definition of a concept and the empirical verification of it, which he sees as a continuous dynamic process. Within this approach, the currently most common method to assess equivalence is multi-group confirmatory factor analysis (for review, see Cieciuch, Davidov, Schmidt, and Algesheimer 2016). However, this approach is not applicable when the theoretical concept is measured via a single-item indicator.

Within-country validity and reliability of a given single-item construct needs to be established for all countries involved in a given international survey project. Additionally, as Wolf, Schneider, Behr and Joye (2016) reflected in relation to any construct, "In comparative research, in addition to validity and reliability, the challenge of creating comparative measures has to be met" (p. 503).

Indeed, for single-item constructs, the methodological issue is: Are the same answers on a given scale comparable if the distribution of these answers differ very much across countries? For example, could we claim that the "completely agree" answer has the same substantive meaning if in one country only three percent of respondents chose it and in another country three times more respondents do the same?

We underscore here that the same answer in different countries may correspond to different points in the country-specific distributions, changing its substantive meaning. For this reason, we opt for finding a common metric for ordinal scales, claiming that answer x of person *i* in country *A* (x_{iA}) corresponds to an answer x of individual *m* in country *B* (x_{mB}) if the location of these answers in the entire distributions is the same in respective countries. The proposed distributional metric tells us in what percentile an individual can be located in each country. Like in the case of z-scores, the metric is country-specific but comparable across countries. Percentile ranks are computed for each country separately, but their meaning across countries is the same.

Even if we establish the common metric of answers, we need to understand the extent to which particular countries differ – with respect to the relationship of the construct and "criterion variables" – from the average of all countries included in the analysis. In this note we propose a new coefficient of comparability, based on the deviation of explained variance of the construct by criterion variables from the explained variance that characterizes all countries together.

Research Focus and Objectives

Previous research shows that for the 1989-2013 time-span, it is possible to identify 895 questions on corruption dispersed in 63 survey waves of 19 international survey projects, such as Global Corruption Barometer, International Crime Victims Survey, Eurobarometer, Life in Transition Survey and other (Wysmulek 2017b). There are two dominant types of questions on corruption, measuring corruption perception or experience of giving bribes/gifts. More often, researchers attempt to measure perceptions of corruption prevalence, but there is no standard approach to this task. Only few measure corruption in specific institutions, although there are strong theoretical and empirical arguments to do so.

In this note, we focus on the *perception of corruption in the education sector* as measured in the Quality of Government Survey 2013, conducted in 25 European countries (see Figure 1). The question on perception of corruption in the education sector is followed by questions on perception of corruption in other institutions, such as police and health sector. There is growing empirical evidence for large differences in corruption levels in different institutions within countries.

The education sector stands out as a particularly interesting case, as the contact with public officials in schools is repetitive, long-term and often compulsory. There are theoretical reasons to expect that the perception of corruption in schools refers to different types of behavior and evokes different associations than corruption in police, health sector or political parties. For scholars interested in analyzing perceptions of corruption in the education system, combining items on

perceptions of corruption in different institutions into a single factor would pose substantive limitations. This does not mean that the corruption perception items do not correlate. Since available survey data do not allow us to disentangle different forms and characteristics of corruption in public schools, which would facilitate constructing a composite factor, we rely on a single-item measure of peoples' perception of corruption in the education sector.

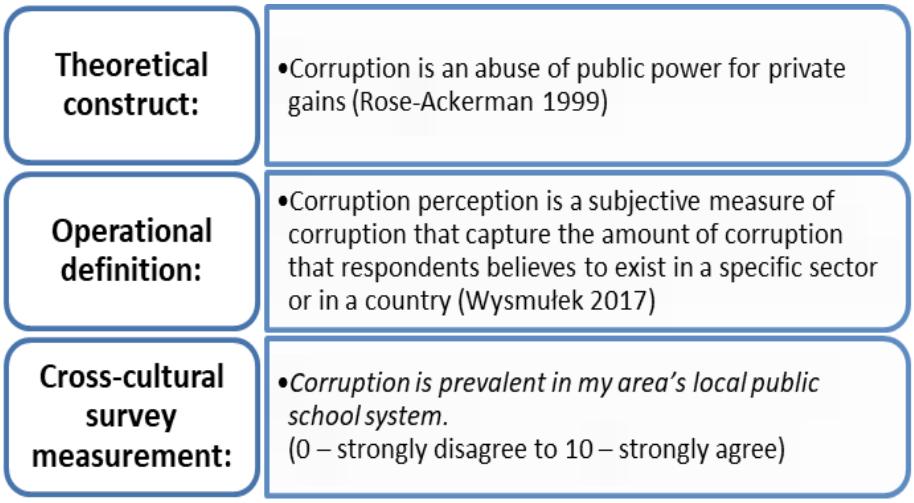


Figure 1. Theoretical Concept, Operational Definition, and Cross-cultural Survey Measurement of Corruption Perception.

One of the ways to analyze the comparability of a survey measure composed of a single variable is to establish its relationship with a criterion variable. In the domain of researching opinions, finding a reference point is a challenging task, yet probably not more challenging than disentangling opinions into components. In this note we aim to bring attention back to the possibilities that criterion validity can give in establishing cross-country measurement comparability.

In short, we aim to: (a) establish a common metric for cross-country comparability of the survey item on perception of corruption in education; and (b) asses cross-national comparability of this item, by using regression models with and without a criterion variable.

Establishing a Common Metric

The questionnaire of the 2013 Quality of Government Survey contains a statement that corruption prevails in a local public-school system. Respondents reacted to this statement using an 11-point scale, ranging from 0 for completely disagree to 10 for completely agree. This is an ascending, bipolar scale. Although the length and the direction of the scale are identical across the 25 countries the project includes, it is reasonable to ask whether the answers are cross-nationally comparable.

We assume that the theoretical construct, operational definition, and the questionnaire instrument are valid. However, we question that the same answer could mean the same thing in different distributional contexts. For example, if in one country most people use only three lowest points on the scale and in another country only a small fraction uses the same points, one wonders

whether these points of the scale are functionally the same. To overcome this problem, we transform the responses from points of scale to percentile values. This procedure assigns a percentile rank to each respondent, which indicates his or her position on the distribution, relative to other respondents in the same country. It overcomes the problem of equal interval distances and assumes that the position of an individual across distributions is comparable, while allowing distributions themselves to be different. For example, if a given respondent choses the 7th point on the scale, and this score was greater or equal to the scores of 82% of all respondents in the sample, then the percentile rank of this respondent is 82.

Figure 2 illustrates the transformation of points on scale from 0 to 10 into percentile ranks. Note that point 2 on the scale 55th percentile rank in Poland and to 65th percentile rank in Germany. This makes the cross-national difference in functional meaning of point 2 on the scale.

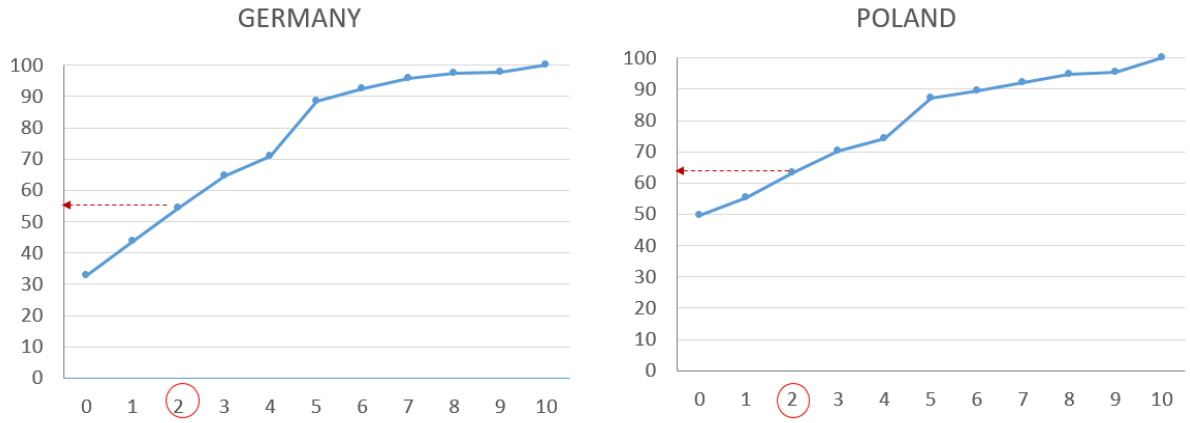


Figure 2. Illustration of the Transfer of 11-Point Scale to Distributional Scale (Percentile Ranks).

Empirical Models: In Search for a Criterion Variable

Establishing a common metric via distributional scales allows us to apply regression analysis with and without criterion variables. The first model explores the effect of social position. Based on extant literature, we assume that the relation of social position with perception of corruption in the education sector varies across countries, so we account for this effect in Model 1.

Model 1: Social position effect

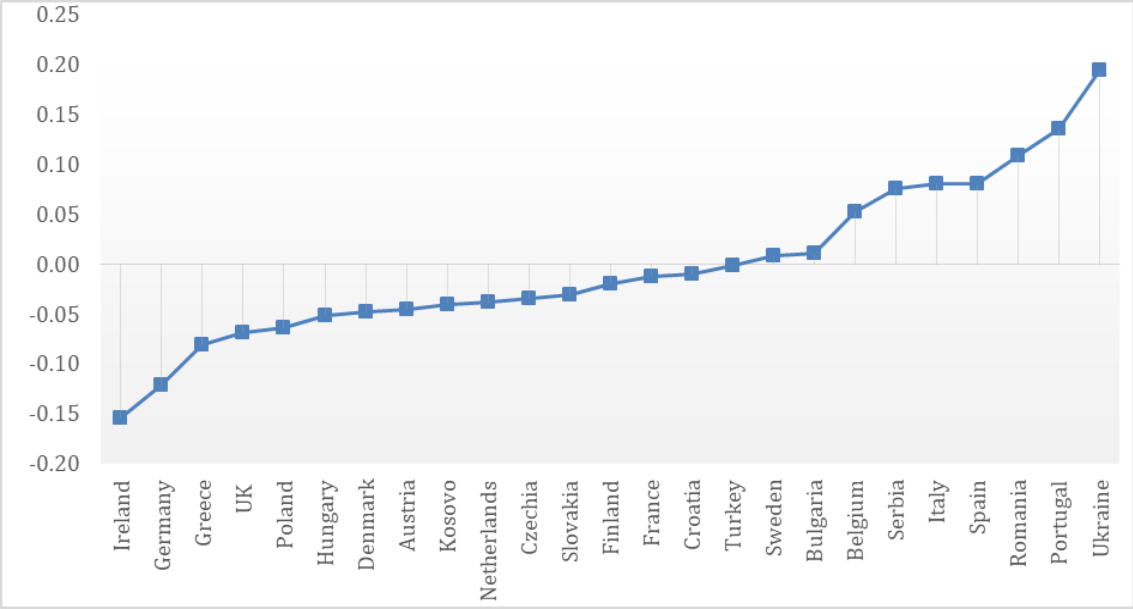
$$\text{Corr\_edu}_i = a + \gamma_1 * \text{female}_i + \gamma_2 * \text{rural}_i + \gamma_3 * \text{age}_i + \gamma_4 * \text{tertiary}_i + e_i$$

In the next model, we add the criterion variables. The criterion variables should be chosen so that they reflect a *priori* knowledge about their strong relation to the construct. In the domain of perceptions, we assume that respondents' opinions depict a certain general worldview they hold. In any given country, those people who agree that corruption in the education sector exists would be more likely to also sense it in other public institutions, including the police and health care, since public servants are perceived to function in similar ways. In Model 2 we include perception of corruption of police and health institutions.

Model 2: Adding criterion variables for perception on corruption in schools as a set of variables of corruption perception in different institutions

$$Corr\_edu_i = a + \gamma_1 * female_i + \gamma_2 * rural_i + \gamma_3 * age_i + \gamma_4 * tertiary_i + corr\_health + corr\_police + e_i$$

To establish cross-national comparability of the concept, we ask both the question of how much explanatory power is brought by Model 2 over Model 1, and of the inter-country difference between these models. In particular, it is assumed that societies have a similar relation between the construct and the criterion. Figure 3 presents the results on how countries differ with respect to the explanatory power of criterion variables from the expected value, that is, the mean.



Source: Quality of Government 2013

Figure 3. The Relative Impact of Criterion Variables (RIC) for 25 Countries.

Using the relative impact of criterion variables (RIC), we can construct a cross-national comparability coefficient (CCC) as

$$CCC = 1 - |RIC|$$

According to RIC (provided in Figure 3), most countries have CCC smaller than 0.95, which shows that they are very similar with respect to the relative impact of the criterion variables. These are Bulgaria, Sweden, Turkey, France, Finland, Slovakia, Czech Republic, Netherlands, Kosovo, Austria, Denmark, and Hungary. In contrast, Ukraine and Ireland deviate from the average impact of criterion variables, with CCC smaller than 0.85.

The choice of criterion variables should be done with caution and requires substantive knowledge about the relation of the studied concept to other concepts. In our illustration we selected perception of corruption in police and health institutions. This choice was taken for its simplicity of methodological exercise rather than for substantive analysis.

Ilona Wysmułek is a Ph.D. student at the Polish Academy of Sciences, and a member of the Harmonization research team.

References

Billiet, Jaak. 2016. "What Does Measurement Mean in a Survey Context?" In Ch. Wolf, D. Joye, T. Smith, and Y. Fu (eds.), *The SAGE Handbook of Survey Methodology*. London: SAGE, p. 193-209.

Billiet, Jaak. 2017. "The interplay between measurement and conceptual validity in the development, and evaluation of measurement instruments." Paper presented at the GESIS Symposium on Advances in Scale Development in the Social Sciences: Issues of Comparability, Mannheim, Germany, 14-15 November.

Cieciuch, Jan, Eldad Davidov, Peter Schmidt, and Ren Algesheimer. 2016. "Assessment of Cross-Cultural Comparability." In Ch. Wolf, D. Joye, T. Smith, and Y. Fu (eds.), *The SAGE Handbook of Survey Methodology*. London: SAGE, p. 630-648.

He, Jia, and Fons van de Vijver. 2012. "Bias and Equivalence in Cross-Cultural Research." *Online Readings in Psychology and Culture* 2(2): 3-19.

Krosnick, Jon. 2017. "Challenges Impeding Comparability of Measurement Across Countries: Findings from an International Project Documenting Response Effects." Paper presented at the GESIS Symposium on Advances in Scale Development in the Social Sciences: Issues of Comparability, Mannheim, Germany, 14-15 November.

Miller, Joanne, Kazimierz M. Slomczynski, and Ronald J. Schoenberg. 1981. "Assessing Comparability of Measurement in Cross-National Research: Authoritarian-Conservatism in Different Sociocultural Settings." *Social Psychology Quarterly* 44(3): 178-91.

Przeworski, Adam, and Henry Teune. 1970. The Logic of Comparative Social Inquiry. New York: Wiley.

Wolf, Christof, Silke L. Schneider, Dorothe Behr, and Dominique Joye. 2016. "Harmonizing Survey Questions Between Cultures and Over Time." In Ch. Wolf, D. Joye, T. Smith, and Y. Fu (eds.), *The SAGE Handbook of Survey Methodology*. London: SAGE, p. 502-524.

Wysmułek, Ilona. 2017a. *Informal Payments in Public Schools: Determinants of Corruption Perception and Behavior in Europe.* Unpublished doctoral dissertation.

Wysmułek, Ilona. 2017b. *Cross-national Surveys featuring Corruption Items in Europe (1989-2013)*, doi:10.7910/DVN/WRVUTZ, Harvard Dataverse, V2, UNF:6:t1ybE8+JV2GxDiYQe2KfIQ==

Wysmułek, Ilona and Kazimierz M. Slomczynski. 2017. "Measuring Perceptions: Comparability of Corruption Items in Cross-National Surveys." Paper presented at the GESIS Symposium on Advances in Scale Development in the Social Sciences: Issues of Comparability, Mannheim, Germany, 14-15 November.

On the Relationship between Two Harmonized Measures of Education in the SDR Dataset

by Marta Kołczyńska (IFiS PAN), Irina Tomescu-Dubrow (IFiS PAN), and Kazimierz M. Slomczynski (Cross-national Studies: Interdisciplinary Research and Training program, CONSIRT)

Education, as an important explanatory factor for many sociological phenomena, is an almost omnipresent variable in statistical analyses using survey data. Although seemingly straightforward, education proves hard to measure accurately in surveys. This, in turn, raises difficulties for using the

education variable in comparative studies, and provides additional challenges for harmonizing education measures ex-post.

Generally, non-specialized surveys (i.e., those with broad topic coverage) carry two types of indicators of education: levels or types of education completed (ordinal or partially ordered), and the number of years of schooling (metric). Researchers employ one or the other depending on item availability and substantive research interest. The general expectation is that both indicators, while having specific properties, capture respondents' educational attainment.

The purpose of this research note is to examine the association between the measure of education levels and years of schooling in cross-national surveys. We use the Survey Data Recycling dataset (SDR, Slomczynski et al. 2017), which pools information from 22 international survey projects. Of the 1721 national surveys in the SDR dataset, 1410 record respondents' education in terms of highest completed level, and 1494 provide data about respondents' years of schooling. For each type of education indicator, we created a harmonized variable, which is comparable across surveys.[3] In this note, we analyze the 1189 surveys of the SDR dataset that contain both target variables for completed education – levels and years.

Although seemingly straightforward, education proves hard to measure accurately in surveys.

Target Variables: Levels of Education and Years of Schooling

In the original (source) surveys, the item on "level of education" generally asks respondents about their highest level of completed education, and appears either as an ordinal or a partially ordered variable.[4] In the process of harmonization in the SDR project, we mapped the categories of the source variables onto categories of the 2011 International Standard Classification of Education (ISCED 2011, UNESCO Institute for Statistics 2012). In the SDR dataset, the resulting target variable "level of education," referring to ISCED levels (i.e., one-digit codes), is an ordinal variable.[5]

The Survey Data Recycling framework postulates that properties of source survey items that would be lost in the process of harmonization be recorded in separate control variables and made available

[3] Harmonization of "education levels" is independent from that of "years of schooling" in that it involves different source variables. See SDH Team 2017 for further information about the harmonization and coding of both target variables.

[4] The partial ordering in the case of some source variables refers to separate categories for pursuing different educational paths within the same educational level. It should also be noted that, while surveys generally ask about "highest completed level of education", among the coded responses there is frequently a category "some university", as well as – less often – "some primary school", "some junior high school" or "some high school." See Slomczynski et al. 2017 (SDR_Master_File_Variable_Report_EDU_1_0.tab) for details about the source variables and the recoding schemes.

[5] The values of the target variable, corresponding to the coding scheme of levels of education, are as follows: 0 stands for *less than primary*, 10 is *primary level*, 20 is *lower secondary*, 30 is *upper secondary*, 40 is *post-secondary non-tertiary*, 50 is *short-cycle tertiary*, 60 is *Bachelor's or equivalent*, 70 is *Master's or equivalent*, and 80 is *doctoral or equivalent*; 90 corresponds to *not elsewhere classified*.

as metadata in datasets created via ex-post survey harmonization. In the SDR dataset, there are three binary harmonization controls for the target variable "level of education," which we created to record whether the original coding of the given education level refers to (a) completed and higher level, (b) vocational training, and (b) incomplete level.

The target variable "schooling years" is defined as the number of schooling years the respondent completed. Its values range from 0, no schooling, to 18, which means 18 years of schooling or more.[6] In the SDR dataset, the following dichotomous harmonization controls accompany this target variable, to record whether: (a) the target value was derived by the SDR team from the question about respondents' age at completion of full-time education,[7] (b) the lowest category of the source-response scale was "open-ended," representing the given value together with lower ones, (c) the highest category of response was open-ended, (d) the respondent was still in school at the time of the survey, and (e) the source value exceeds 18 years and was top-coded. A further harmonization control variable records (f) the number of distinct values the schooling years variable takes in the source data.[8]

In this note, we analyze the 1189 surveys of the SDR dataset that contain both target variables for completed education – levels and years.

Overall Covariation

Completed levels of education and number of years of schooling share the expectation that both measure the same concept - respondents' educational attainment. Hence, the association between the two target variables in the SDR dataset should be positive and strong, although not perfect. In this part of the note, we present the overall covariation of the two variables, without accounting for the effects of control variables discussed in the previous sections. These effects, although substantively and statistically significant, do not change the pattern of associations between the two education variables which we draw attention to in this section. We return to examining the effects of harmonization controls in the following section.

We start by calculating the Spearman's rank correlation between levels of education and schooling years for the pooled sample of 1,524,897 respondents. The result is 0.77. Indeed, the two target variables are strongly related, but their association is not as strong as one would expect for

[6] Top-coding eliminates unrealistically high values. The trimming point was chosen at 18 years of schooling, which includes 12 years of primary and secondary education combined, and 6 additional years of post-secondary education. See Slomczynski et al. 2017 for details about the source variables and the recoding schemes.

[7] The number of schooling years was obtained by subtracting 6 from the age at completion of full-time education; 6 is currently the most common age for entering school worldwide. Of the 1419 surveys, 659 contain direct questions about respondents' years of schooling, while 835 record respondents' age at completion of education.

[8] This control variable takes values ranging from 2 to 19.

closely related variables. It should be noted that in the European Social Survey Round 6, in the source data (prior to ex-post harmonization), Spearman's *rho* between ISCED and schooling years is 0.81 overall, with country-specific correlations between 0.66 in Switzerland, and 0.96 for Albania and Kosovo. Results from the ESS can be interpreted as providing a 'realistic' benchmark for the association of education levels and years of schooling, given measurement difficulties in both.

Figure 1 presents the variation in correlations between education levels and schooling years across the 1189 national surveys. In 36 surveys, the correlation of the two target variables is below 0.5, and among them, four surveys yield negative coefficients.[9] In 78 of the surveys, Spearman's rank correlation exceeds 0.95. Altogether, the correlation coefficients range from -0.222 (New Baltics Barometer, Wave 1, Lithuania) to 0.9995 (International Social Survey Programme, 1991, Hungary).

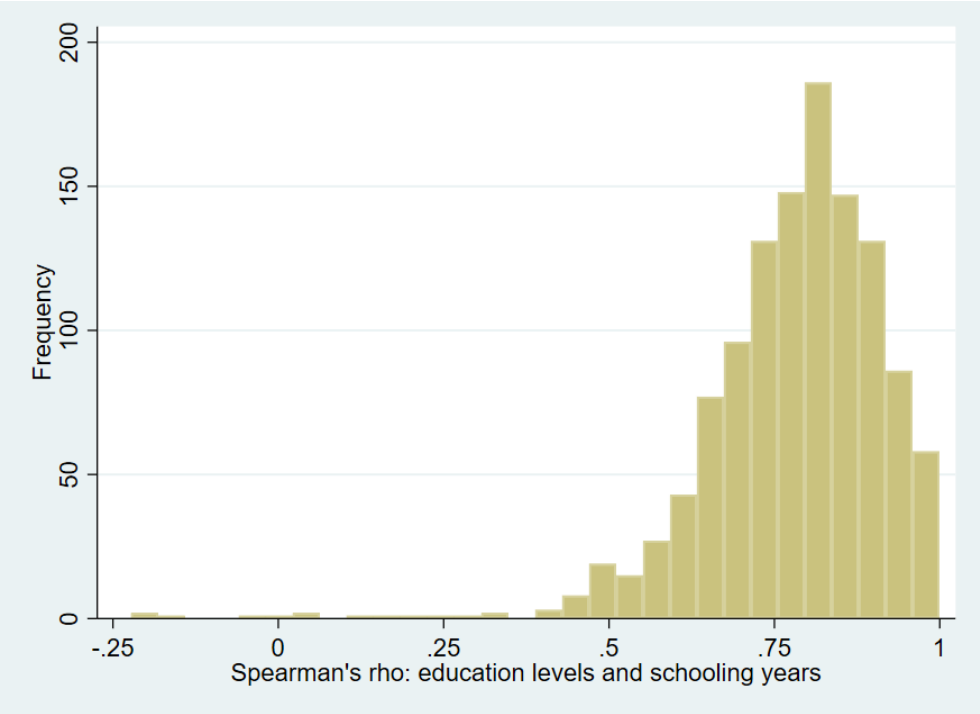


Figure 1. Distribution of Spearman's *rho* Between Education Levels and Schooling Years, by Survey.

Negative correlations imply that more schooling years are associated with completing a lower educational level than a higher one. Taken as such, this would contradict the sequential character of formal education systems worldwide. In practice, the result indicates the presence of errors, which could stem from the source data or the process of harmonization.

Renewed scrutiny of the source survey documentation and the SDR documentation might shed some light on these problems. The low, negative, correlations between education levels and years of schooling in Estonia, Latvia and Lithuania in the 1st wave of the New Baltic Barometer (NBB) can serve as illustration. Checking the harmonization procedures employed in the SDR project did not reveal any recoding errors. The NBB codebook (NBB 2010) indicates that codes assigned to education levels in wave 1 are reversed compared to codes assigned in later waves:

[9] Latinobarometro, wave 2006, Argentina, and all surveys in New Baltic Barometer wave 1: Estonia, Latvia, and Lithuania.

initially, university degree was assigned the value 1, "no qualifications" was coded as 7; in later waves higher values correspond to higher education levels. If this switch came with a coding error in the data, a negative correlation with years of schooling would be plausible.[10] However, it would not account for the low value of the correlation coefficient.

Positive correlations that are very close to perfect are also problematic, since they imply that the education system has a near zero drop-out rate at all levels. In practice, such correlations could stem from the fact that values of one variable – years of schooling, especially – are calculated based on the values of the other variable, instead of respondents providing answers to each question separately. A close review of the documentation may provide evidence of such procedures. For example, in the case of ISSP/1991/Hungary, where the correlation between education levels and schooling years equals 0.9995, the codebook notes that in that survey "Years in school included part-time education, [and were] derived from educational attainment" (ISSP 1993: 142).

Years of Schooling by Levels of Education

In this section we engage the harmonization controls corresponding to "schooling years" and "education levels," to see how including them affects the association between the two target variables. In doing so, let us first take a look at the variation of "schooling years" across "education levels," when controls are not applied.

Table 1 shows a pattern of increasing mean schooling years for consecutive education levels, with less than primary average having 5.0 years and doctoral having 17.1. We notice that the mean years of schooling are lower for *post-secondary* than for *short-cycle tertiary*, 13.5 and 14.7 respectively. *Upper secondary* has 12.0 years, and *Bachelor's degree* almost 16, as usually coded in transforming levels of education into years of schooling.

Table 1. Schooling Years by Level of Education.

Level of education	Schooling years		
	Mean	Std. Dev.	*N*
Less than primary	5.0	4.22	119,189
Primary	7.6	2.79	242,745
Lower secondary	10.2	2.22	166,871
Upper secondary	12.0	2.41	563,553
Post-secondary	13.5	2.37	99,921
Short-cycle tertiary	14.7	2.58	48,114
Bachelor's or equivalent	15.9	2.37	254,037
Master's or equivalent	16.1	1.86	30,442
Doctoral or equivalent	17.1	0.49	25
Not elsewhere classified	9.7	5.95	1,760
Total	11.5	4.165	1,526,657

[10] We compared the data from the 1993 and 1995 waves of NBB; for each country, the difference in distribution of education levels between these waves is large and does not provide a clue about the presence of absence of the coding error.

One would expect that respondents' whose education levels are incomplete have, on average, fewer years of schooling than respondents with completed levels. In turn, respondents coded into a given level of education should have lower means of years of schooling than respondents with completed-and-higher levels. Vocational training usually requires some additional time.

Generally, results conform to these expectations. In Table 2 we show how control variables describing educational levels influence the mean years of schooling, for a subset of education levels that are most comparable across countries: *primary education*, *upper secondary*, and *Bachelor's degree or equivalent*. For example, *incomplete primary* has the lowest mean number of years of schooling. *Upper secondary* with vocational training has a higher mean than the same level without this training. For *Bachelor's or equivalent*, the completed-and-higher level carries a higher mean than the reference mean.

Table 2. Schooling Years for Primary, Upper Secondary, and Bachelor's or Equivalent Levels, with Control Variables.

Level	Control variables*	Mean	Std. Dev.	N
Primary	Incomplete level	6.8	3.51	7986
Primary	*Reference***	7.6	2.75	231732
Primary	Completed and higher	7.3	2.83	2135
Primary	Completed and higher, vocational training	10.0	2.37	892
Upper secondary	Incomplete level	11.1	2.70	85756
Upper secondary	Incomplete level, vocational training	10.5	2.64	45133
Upper secondary	*Reference***	12.3	2.14	346039
Upper secondary	Vocational training	12.5	2.50	86625
Bachelor's or equivalent	Incomplete level	14.5	2.61	43985
Bachelor's or equivalent	Incomplete level, vocational training	14.4	2.80	499
Bachelor's or equivalent	*Reference***	15.8	2.18	61828
Bachelor's or equivalent	Vocational training	16.1	2.10	12140
Bachelor's or equivalent	Completed and higher	16.3	2.19	134878
Bachelor's or equivalent	Completed and higher, vocational training	15.2	2.60	707

* All control variables – completed and higher, vocational training, and incomplete level – are dichotomies, with 1 for fulfilling the condition and 0 otherwise. Realized combinations of controls differ among levels of education.
** Reference category means that the value for each control is equal to 0.

We also examined the effects of harmonization controls for the target variable "years of schooling" on the correspondence between years of schooling and education. These analyses, presented in Table 3, show that within education levels, the mean value and variance of schooling years is higher when the "schooling years" variable was derived from source variables asking about "age at completion of education" than when the source question asked directly about years of schooling. For some education levels, the differences in means are not very large, in particular for *upper secondary*, *Bachelor's* and *Master's*, which likely are the most standard levels of education across countries. In other cases – *primary*, *lower secondary*, *post-secondary* – the differences in means between the "derived" and "not derived" categories are almost 1 year, and almost 2 years for *short-cycle tertiary*.

The effect of harmonization controls can also be analyzed in a regression framework. We estimated a set of linear regression models predicting the number of schooling years with individual respondents as cases and standard errors clustered around national surveys (Table 4). The models

show that, for a given level of education, in surveys where the lowest category of schooling years was open ended (e.g., 5 years or less), respondents tend to have fewer years of schooling than when the lowest category was not open (Model 1). This negative effect is stronger in surveys with fewer distinct values in the variable "schooling years" that are realized in the data (Model 2). The same is true for surveys where the highest category in "schooling years" was open (e.g., 16 years or more, Model 3). These patterns reflect the harmonization decision that in the case of schooling years coded in ranges (brackets), the respondent is assigned the lower value from that range.

Table 3. Schooling Years by Education Level and the Control Variable "Derived from Age Completed Education."

Level of education	Schooling years			
	Derived		Not derived	
	Mean	Std. Dev.	Mean	Std. Dev.
Less than primary	5.146	4.378	4.520	3.738
Primary	8.051	2.831	7.189	2.688
Lower secondary	11.020	2.687	10.073	2.145
Upper secondary	12.191	2.667	11.849	2.014
Post-secondary	14.340	2.918	13.376	2.293
Short-cycle tertiary	15.636	2.497	13.799	2.330
Bachelor's or equivalent	16.006	2.603	15.766	2.146
Master's or equivalent	17.178	1.791	16.813	1.865
Doctoral or equivalent	17.600	0.894	16.950	0.224
Not elsewhere classified	13.385	3.776	9.673	5.957
Total	11.427	4.488	11.530	3.883

Based on the same models, respondents who at the time of the interview were "still at school", on average have one year of schooling more than those who were out of school. Surveys that have fewer distinct values of the "schooling years" variable realized in the data tend to have a lower value of schooling years in the harmonized variable. Finally, respondents whose number of schooling years was top-coded at 18 still have a higher value on that variable than those without top-coding, and this positive effect is greater in surveys where the "schooling years" variable is derived from age at completion of education (Model 4).

Conclusion

In 920 out of the 1189 surveys containing both harmonized education variables, the correlation of education levels and years of schooling is greater than 0.7, which according to most benchmarks indicates a strong relationship. We do find, however, instances where the association is (a) lower than expected for two operationalizations of the same concept, or (b) higher than expected for survey questions asked separately. These results require further investigation. Referring to the lower associations, some of the "mismatch" between the measures of "education levels" and "schooling years" reflects real differences in number of years and highest level of education obtained, for example in the case of drop-outs or people who complete an education level ahead of their peers. The remaining discrepancies between education levels and schooling years stem from measurement

error in the original data and differences in the reliability of measurement – higher reliability is expected of the education levels items, given the cognitive difficulties that respondents encounter when trying to recall the number of years of education they completed (Schneider 2016), and differences in the formulation of the questionnaire items across surveys (Hoffmeyer-Zlotnik and Warner 2007).

Table 4. Models Predicting the Number of Years of Schooling with the Level of Education and Harmonization Control Variables for the Target Variable "Schooling Years."

	Model 1	Model 2	Model 3	Model 4
Education level: Less than primary (reference)[a]				
Primary	2.621***	2.617***	2.615***	2.610***
	(0.122)	(0.122)	(0.122)	(0.122)
Lower secondary	5.203***	5.197***	5.194***	5.179***
	(0.128)	(0.128)	(0.128)	(0.129)
Upper secondary	6.877***	6.869***	6.858***	6.868***
	(0.121)	(0.121)	(0.120)	(0.121)
Post-secondary	8.328***	8.318***	8.312***	8.311***
	(0.139)	(0.138)	(0.138)	(0.139)
Short-cycle tertiary	9.065***	9.065***	9.046***	9.033***
	(0.147)	(0.147)	(0.146)	(0.148)
Bachelor's or equivalent	9.969***	9.962***	9.953***	9.958***
	(0.132)	(0.132)	(0.132)	(0.132)
Master's or equivalent	10.690***	10.704***	10.699***	10.794***
	(0.140)	(0.140)	(0.140)	(0.137)
Doctoral or equivalent	12.858***	12.563***	11.864***	12.811***
	(0.822)	(0.766)	(0.626)	(0.846)
Harmonization controls for "schooling years"				
Open minimum	-0.707**	-1.123**	-0.194	-0.698**
	(0.266)	(0.350)	(0.267)	(0.267)
Open maximum	-1.466***	-1.094***	-1.970***	-1.473***
	(0.270)	(0.326)	(0.262)	(0.271)
Number of options (centered 8)	-0.118***	-0.125***	-0.133***	-0.117***
	(0.012)	(0.013)	(0.013)	(0.012)
Open minimum x Number of options		0.115**		
		(0.042)		
Open maximum x Number of options			0.174***	
			(0.037)	
Still at school	1.017***	1.012***	1.017***	1.023***
	(0.061)	(0.060)	(0.061)	(0.062)
Derived from age	0.121*	0.128*	0.123*	0.071
	(0.057)	(0.058)	(0.058)	(0.061)
Top-coded	3.963***	3.957***	3.957***	3.434***
	(0.053)	(0.052)	(0.053)	(0.060)
Top-coded x Derived from age				0.784***
				(0.082)
Constant	6.008***	6.064***	6.150***	6.031***
	(0.182)	(0.183)	(0.187)	(0.183)
N	1,524,897	1,524,897	1,524,897	1,524,897
Number of clusters (surveys)	1189	1189	1189	1189
R^2	0.668	0.669	0.669	0.669

[a] Cases with the level of education coded as "not elsewhere classified" have been excluded. *p<0.05; **p<0.01; ***p<0.001

Our analyses also explored the effects of harmonization control variables, which bring in information about the design of the survey questions or ways of coding responses that would be lost in the process of harmonization. Further research can address possibilities of constructing harmonization controls to improve the quality of measurement of education in cross-national surveys harmonized ex-post.

Marta Kołczyńska received her Ph.D. in sociology from The Ohio State University (OSU) and is a member of the Harmonization research team. Marta has a post-doctoral position in the NCN funded project, "Political Voice and Economic Inequality across Nations and Time," at IFiS PAN.

Irina Tomescu-Dubrow is Associate Professor at the Institute of Philosophy and Sociology, Polish Academy of Sciences (PAN). She is Program Manager at Cross-National Studies: Interdisciplinary Research and Teaching Program (CONSIRT), of OSU and PAN.

Kazimierz M. Slomczynski is Professor Dr. Hab. at IFiS PAN and Emeritus Professor, Department of Sociology, OSU. He directs the Cross-National Studies: Interdisciplinary Research and Training program (CONSIRT) of OSU and PAN.

References

Hoffmeyer-Zlotnik, Jürgen H.P. and Uwe Warner. 2007. "How to Survey Education for Cross-National Comparisons: The Hoffmeyer-Zlotnik/WarnerMatrix of Education." *Metodološki zvezki* 4(2): 117-148.

ISSP (International Social Survey Programme). 1993. *International Social Survey Programme: Religion I - ISSP 1991. Codebook*. GESIS Data Archive, Cologne. ZA2150. https://dbk.gesis.org/dbksearch/download.asp?db=E&id=15590 (accessed December 2, 2017).

NBB (New Baltics Barometer). 2010. *New Baltic Barometer (NBB) Trend File Code Book*. UK Data Archive Study Number 6510 - New Baltic Barometer I-VI Trend Dataset, 1993-2004. http://doc.ukdataservice.ac.uk/doc/6510/mrdoc/pdf/6510nbb_trend_codebook.pdf (accessed December 2, 2017).

Schneider, Silke L. 2016. *The Conceptualisation, Measurement, and Coding of Education in German and Cross-National Surveys. GESIS Survey Guidelines*. Mannheim, Germany: GESIS Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_020.

SDH Team (Survey Data Harmonization Team): Kazimierz M. Slomczynski, J. Craig Jenkins, Irina Tomescu-Dubrow, Marta Kołczyńska, Ilona Wysmułek, Olena Oleksiyenko, Przemek Powałko, Marcin W. Zieliński. 2017. *Master file documentation. SDR Master Box Version 1.0*. Harvard Dataverse, V1, DOI.

Slomczynski, Kazimierz M., Craig J. Jenkins, Irina Tomescu-Dubrow, Marta Kołczyńska, Ilona Wysmułek, Olena Oleksiyenko, Przemek Powałko, and Marcin W. Zieliński. 2017. "SDR Master Box." doi:10.7910/DVN/VWGF5Q, Harvard Dataverse, V1, UNF:6:HIWud4wueVRsU8wTN+lySg==

UNESCO Institute for Statistics. 2012. *International Standard Classification of Education: ISCED 2011*. Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf (accessed November 14, 2017).

Harmonization would like to hear from you!

We created this *Newsletter* to share news and help build a growing community of those who are interested in harmonizing social survey data. We invite you to contribute to this Newsletter. Here's how:

1. Send us content!

- Send us your announcements (100 words max.), conference and workshop summaries (500 words max.), and new publications (250 words max.) that center on survey data harmonization in the social sciences;
- Send us your short research notes and articles (500-1000 words) on survey data harmonization in the social sciences. We are especially interested in advancing the methodology of survey data harmonization. If we have any questions or comments about your items, we will work with you to shape them for this *Newsletter*.

  Send it to: Joshua K. Dubrow, dubrow.2@osu.edu.

2. Tell your colleagues!

To help build a community, this *Newsletter* is open access. We encourage you to share it in an email, blog or social media.

Support

This newsletter is a production of Cross-national Studies: Interdisciplinary Research and Training program, of The Ohio State University (OSU) and the Polish Academy of Sciences (PAN). The catalyst for the newsletter was a cross-national survey data harmonization project financed by the Polish National Science Centre in the framework of the Harmonia grant competition (2012/06/M/HS6/00322). This newsletter is now funded by the US National Science Foundation (NSF) under the project, "Survey Data Recycling: New Analytic Framework, Integrated Database, and Tools for Cross-national Social, Behavioral and Economic Research" (SDR project - PTE Federal award 1738502). The SDR project is a joint project of OSU and PAN. For more information, please visit dataharmonization.org.

Copyright Information