Transparency and Replication in Cross-National Survey Research: Identification of Problems and Possible Solutions

Published as:

Damian, E., Meuleman, B., & van Oorschot, W. (2019). Transparency and Replication in Cross-national Survey Research: Identification of Problems and Possible Solutions. *Sociological Methods & Research*. Doi: 10.1177/0049124119882452

Author information

Elena Damian (corresponding author)

Institute for Social and Political Opinion Research, Centre for Sociological Research

University of Leuven

Parkstraat 45 - bus 3601, 3000, Leuven, Belgium

Email address: elena.damian@kuleuven.be

Bart Meuleman

Institute for Social and Political Opinion Research, Centre for Sociological Research

University of Leuven

Parkstraat 45 - bus 3601, 3000, Leuven, Belgium

Email address: bart.meuleman@kuleuven.be

Wim van Oorschot

Social Policy and Social Work Department , Centre for Sociological Research

University of Leuven

Parkstraat 45 - bus 3601, 3000, Leuven, Belgium

Email address: wim.vanoorschot@kuleuven.be

Author Biographies

Elena Damian is a Postdoctoral Researcher at the Institute for Social and Political Opinion Research, University of Leuven (Belgium). Her research interests include research transparency, replication, applications of multilevel modelling, and cross-national survey research.

Bart Meuleman is Associate Professor and head of department at the Centre for Sociological Research (CeSO) at the University of Leuven (Belgium). His main research interests are cross-cultural survey methodology and cross-national comparisons of value and attitude patterns.

Wim van Oorschot is Professor of Social Policy at KU Leuven (Belgium). He is initiator and co-designer of the Welfare Attitudes modules of the European Social Survey waves 2008 and 2016.

Authors' Note

Data used in this article and syntax to replicate the study's findings are available online.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Acknowledgement

We would like to thank Jaak Billiet, Ruud Luijkx, and Dominique Joye for their helpful comments on this study. We also want to thank Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS) Board of Strategic Advice members for giving us the opportunity to present and get feedback on an earlier version of the checklist for reporting cross-national survey research.

Transparency and Replication in Cross-National Survey Research: Identification of Problems and Possible Solutions

Abstract

In this paper, we examine whether cross-national studies disclose enough information for independent researchers to evaluate the validity and reliability of the findings (evaluation transparency), or to perform a direct replication (replicability transparency). The first contribution is theoretical. We develop a heuristic theoretical model including the actors, factors, and processes that influence the transparency of cross-national studies, and provide an overview of the measures currently taken to improve research transparency. The second contribution is empirical, in which we analyze the level of transparency in published cross-national studies. Specifically, using a random sample of 305 comparative studies published in one of 29 peer-reviewed social sciences journals (from 1986 to 2016), we show that, even though all the articles include some methodological information, the great majority lack sufficient information for evaluation and replication. Lastly, we develop and propose a set of transparency guidelines tailored for reporting cross-national survey research.

Keywords: transparency, replication, cross-national survey research, secondary analysis

1. Introduction

Merton's (1973) scientific norm of communalism makes it clear that transparency is one of the foundations of the scientific method. The disclosure of methodological information has two main functions. Transparency enables readers to evaluate the validity and reliability of a study's findings–what we term *evaluation transparency*–and allows others to attempt to replicate the results; that is, *replicability transparency*. Thus, achieving research transparency is crucial for the overall quality of research and the credibility of findings, and consequently speeds up scientific progress (Munafò et al. 2017). Although the academic community has always recognized its importance, the issue of transparency has moved up the academic agenda over recent years. This emphasis on transparency is a direct response to increasing evidence of a failure to replicate experiments (Baker 2016; Huizenga et al. 2012; Open Science Collaboration 2015), a high prevalence of various questionable research practices (Fanelli 2009; Ioannidis 2005; John, Loewenstein, and Prelec 2012; Simmons, Nelson, and Simonsohn 2011; Steneck 2006), and reports of cases of misconduct, mostly in experimental psychology and health studies (Nelson, Simmons, and Simonsohn 2017).

However, in the field of cross-national survey research–one of the most popular designs in social science research–the tendency toward a greater concern for transparency is largely absent. At first sight, it could be argued that this results from the fact that the majority of cross-national survey studies rely on secondary analysis of publicly available data, which enhances transparency and minimizes the risk of data fabrication. However, we are convinced that the neglect of transparency in the field of cross-national research is unfortunate and problematic. Instead, the central argument in this paper is that the specific features of cross-national survey research bring about particular challenges to transparency. First, the fact that data are collected in multiple countries yields a high level of complexity. Even if extensive documentation is available, survey users–who are usually not involved in the data collection process–might encounter difficulties in understanding and reporting issues that are relevant to assessing data quality. Second, carrying out secondary analyses on multi-national data is often a long and intricate process. In each step of the research, mistakes, deliberate decisions, and unconscious processes can induce bias and consequently distort findings. Third, the availability of a huge amount of data prior to the formulation of hypotheses stimulates what Gelman and Loken (2014) call *data dependent analysis*. An almost infinite number of possible hypothesis tests could be performed. Researchers continuously have to make decisions about which tests to pursue and report, and which tests not to: a situation that has been compared to “a garden of forking paths” (Gelman and Loken 2014, p. 460). Full disclosure about the data and analytical procedures performed is critical for the evaluation of findings and replication, but poses considerable challenges in the case of cross-national research.

This paper opens up the debate about the level of transparency in cross-national survey research. Our first contribution is theoretical, as we develop a heuristic model explaining which actors, factors, and processes influence the level of transparency of published research articles. We argue that current individualistic or social responsibility models for achieving research transparency (Freese 2014) do not fully account for the specificity of cross-national survey data and the actors involved in producing such data. The second contribution is empirical, as we analyze the actual level of transparency in published cross-national research. Using a random sample of 305 comparative studies published in 29 peer-reviewed sociology, political sciences, and cross-cultural psychology journals (from 1986 to 2016), we provide insights into the current reporting practices in cross-national survey research. Lastly, we develop and present a checklist to help researchers achieve greater transparency in their work.

2. Specific transparency issues in cross-national survey research

Despite the relative lack of attention paid to transparency in this field, the particularities of cross-national survey research lead to specific threats to transparency. A number of issues stem from the complexity of cross-national data, in terms of collection as well as analysis.

First, with regard to data collection, multi-national surveys–such as the European Social Survey (ESS), the European Values Study (EVS), and the International Social Survey Programme (ISSP)–are known to have an “extra layer of survey design” compared with national surveys (Lynn, Japec, and Lyberg 2006, p. 11). The data collection is conducted by different organizations in diverging national contexts, which means that there are considerable variations in research practices, national laws, regulations for surveys (e.g., data linkage, access to sampling frames, and informed consent), time frames, and budgetary constraints. Furthermore, multiple actors are responsible for the overall coordination, harmonization, archiving, and preservation of the data (e.g., a central coordination team, data archives, etc.). On the upside, such complex data collection processes result in rich datasets that allow comparison between societies on a wide variety of topics. However, the sheer complexity of the process leads to the potential for biases and errors over and beyond the methodological issues that are present in single-nation surveys (e.g., differences in sampling frames, diverging nonresponse patterns, or the challenge of obtaining accurate translations of survey questions: Berry et al. 1992; Harkness, Van de Vijver, and Mohler 2003; Harkness et al. 2010; Van de Vijver and Leung 1997; Wolf et al. 2016). Even if international survey projects manage to produce exhaustive metadata to document the plethora of methodological issues, it is not guaranteed that data users–who are not involved in the data collection–will be familiar with the massive amount of information and will succeed in communicating all the relevant information to their readership. As a result, many errors or data limitations are likely to remain hidden from readers and even from the authors of the studies. Readers and authors risk holding an overly optimistic view of the data quality of well-known cross-national data projects. To give an example, we often see scholars report that the data are representative for all the countries included; a statement that hides the complex reality of covering general populations in a multitude of countries.

Second, the analysis of cross-national survey data also comes with specific challenges to transparency. In the case of most published papers, cross-national survey data are used as secondary data: the data are collected independently from and prior to the formulation of the research question. As a result, scholars face data limitations in that they often have to rely on general proxy operationalizations, because cross-national datasets lack precise and detailed measurements of the specific concepts that are investigated. Furthermore, individual-level survey data are often linked to contextual data at the national or regional level. Details about the nature of contextual variables and their sources need to be reported in the article in order to evaluate the robustness of the findings and enable replication. As a result, cross-national data analysis often requires a complex process of data cleaning, manipulation, and merging. Disclosure about the various steps taken for data preparation (e.g., the choice of particular indicators, recoding of variables, handling missing data, and the sources of the contextual data) is warranted, but requires the provision of a large quantity of information. However, perhaps even more challenging for transparency is the fact that researchers tend to explore a multitude of possible expectations rather than testing a narrow set of pre-defined hypotheses. Such exploratory processes can be highly data dependent. Based on the findings of preliminary analyses, decisions are taken (e.g., to use particular modeling strategies, include or exclude particular respondents, add or remove control variables, include interaction effects, etc.). This process is generally not well documented. The fact that researchers move through a "garden of forking paths" (Gelman and Loken, 2014) basically invalidates reported p-values: Among the multitude of possible tests, some will render significant p-values by chance, even if no real effects are present in the population. Moreover, the preparation and analysis of large-scale, cross-national data, where hypotheses are often formulated after data collection and where numerous decisions have to be taken, increase the risk of using *researcher degrees of freedom* (researcher DFs: Simmons et al. 2011; Wicherts et al. 2016). According to Wicherts and colleagues (2016), researcher DFs are choices that "are arbitrary from a substantive or methodological point of view [...] but could affect the outcome of significance tests applied to the data, and hence the conclusions drawn from the research" (p. 1). Consequently, their presence can seriously increase the chances of false positive findings and affect the size of estimates. Wicherts and colleagues (2016) provide a detailed checklist of 34 researcher DFs found in psychology–but that can also be generalized to other fields–such as selecting a dependent variable out of several alternative measures or operationalizations, hypothesizing after results are known (HARKing) (Keer 1998), or failing to assure reproducibility. Breznau (2016), and Silberzahn and colleagues (2018) illustrate the negative consequences of researcher DFs, and consequently the need for more transparency regarding the research process. Breznau (2016) analyzes two replication studies that use the same secondary data and methods, but fail to reproduce the findings of a well-known cross-national study (Brooks and Manza 2006). Breznau explains that the variation in results is caused by *secondary observer effect bias*. Similar to researcher DFs, this is an error that "comes from the qualities of researchers and their idiosyncratic behaviors while conducting research" (Breznau, 2016, p. 313). In a study by Silberzahn and colleagues (2018), 29 research teams were given the same data and research questions, but had the freedom to use whichever method they considered most suitable to test them. Some 69 percent of the teams found a significant effect, while the others did not. Furthermore, 21 of the 29 teams used a different combination of explanatory variables. Hence, variation in the findings across studies that use the same data or methods is difficult to avoid, and disclosure about the research process is therefore essential.

To sum up, the fact that multi-national survey data are often publicly available does not prevent the field from facing pressing threats to research transparency. As in other types of research, the relevant features of data collection, the resulting weaknesses, and the various steps in data preparation and analysis need to be communicated clearly to the scientific community. In the next section, we discuss the actors, factors, and processes that shape the level of transparency of published research based on analyses of cross-national survey data.

3. Actors, factors, and processes that play a role in the level of transparency of cross-national survey studies

A literature review of research transparency in social sciences indicates that while the academic community acknowledges transparency and replication as major elements of scientific quality, there are still no widely agreed-upon rules or procedures put in place to ensure them (Abbott 2007; Carsey 2014; Firebaugh 2007; Freese 2007; Lucas, Morell, and Posard 2013; Lupia and Elman, 2014). In the absence of enforced transparency policies, the responsibility for what is disclosed in an article, and therefore how much information is given for evaluation and replication, is put in the hands of the author. This so-called *individualistic responsibility model* implies full trust in researchers to keep detailed accounts of their research process and to disclose any additional information on request (Freese 2007; Ishiyama 2014). More recently, scholars such as Freese (2014) have proposed that this current responsibility model does not capture the complexity of the publication process of an article and should be replaced with the *social or community responsibility model*. According to this model, journals–as representatives of the scientific community–should require authors to place the data and replication documents in the public domain. In this way, the published work is separated from the author and can be accessed, verified, or replicated at any point in time. However, we are convinced that–especially when it comes to comparative survey research–both responsibility models underestimate the number of potential factors and actors involved in research transparency. According to our model, published articles are the product of the interplay between authors (who possess particular skills and knowledge), datasets (with varying degrees of data quality and documentation) and journals (that can impose certain requirements or policies before a study is accepted for publication). Hence, as we explained in the previous section, survey data are among the key factors that have a direct influence on the level of transparency of published cross-national studies. Moreover, there are several institutional actors that have an indirect but still salient role, namely data archives, academic professional organizations, educational institutions, and publishing agencies. Below, we describe the direct and indirect influences of these actors, factors, and processes on article transparency (Figure 1).

[Figure 1 about here]

Direct influences on article transparency. As a central actor in the publication process, the author has the most direct influence on research transparency. Conscious or unconscious decisions to include or exclude particular methodological information are informed by the skills, methodological knowledge, and scientific attitudes of authors. For instance, it is likely that authors who invest enough time in learning about the background of the data or who have survey expertise will write more transparent articles. Authors also have other academic roles–e.g., as members of journal editorial boards, academic professional associations, funding committees, or graduate programs–through which they can influence the transparency of articles (these indirect influences are discussed below). In recent years, scholars have launched new initiatives for promoting transparent science and have developed replication and research transparency guidelines and recommendations (Dafoe 2014; Höffler 2017; King 1995). In addition, there are several new professional organizations that dedicate their efforts to promoting research transparency and integrity, such as the Centre for Open Science (COS) and the Berkeley Initiative for Transparency in the Social Sciences. For example, among the COS services are a set of guidelines on how to promote transparency and openness (TOP Guidelines), a free platform for pre-analysis plans (Preregistration Challenge), and journals are encouraged to offer "badges" for data sharing, preregistration of studies, and replication material (Open Science Badges).

In the case of cross-national survey research, authors make use of survey data in the production of articles. Several data-related factors constrain or steer authors' decisions to report particular methodological information. Of specific importance are the ways in which the data documentation, metadata, and their overall quality are presented to the public. When survey teams have enough time and money to prepare and disseminate metadata, researchers are more likely to learn about the strengths and limitations of the data and report them in an article. This is particularly important, as there seems to be a general perception that large-scale surveys have high data quality because of the teams of experts involved in the survey design and coordination. In the case of contextual data, it is equally important that international organizations, such as the OECD, the World Bank, and Eurostat provide clear and easily to accessible information about the measurements of their contextual indicators and any future data updates or corrections. Breznau (2016) reveals that these international data banks do not always keep records of updates to their data. For example, the OECD social welfare spending indicator for Switzerland in 2010 was either 17.85 or 26.41, depending on when the data was downloaded. Contacted by email, OECD representatives reported to Breznau that they do not provide information about the maintenance of their database.

As a third direct influence, journals act as gatekeepers, determining what is published, and therefore have the power to shape the transparency of research. In principle, editorial policies have the capacity to encourage or even enforce certain transparency reporting guidelines, offer the option to submit supplementary (online) material, or require separate replication documents and data sharing. Most journal guidelines for authors still include only formatting instructions, and sometimes a brief reminder that researchers have the ethical obligation to provide enough methodological information. However, in the field of political sciences, an increasing number of journals have started to include various transparency guidelines in recent years. A leading example is the American Journal of Political Science (AJPS), which recently adopted a policy requiring all accepted manuscripts to come with detailed replication folders and data upon publication. To facilitate this policy, the journal offers a free online platform to upload all replication material (the Harvard Dataverse) and there is a team of researchers that checks if all replication material functions correctly. The online platform used by AJPS is just one of an increasing number of free data repositories for replication material (e.g., Open Science Framework, figshare, and DataDryad). With this, journals are effectively not constrained by space limitations to ask scholars to share replication material or data. Apart from the editorial boards, peer-reviewers that collaborate with journals also have a decisional role regarding the type of information that goes into an article, and hence its level of transparency. For instance, they can require greater transparency in manuscripts or replication material. Recently, the Royal Society Open Science Journal created *The Peer Reviewers' Openness Initiative*, which has the goal of increasing open access practices (Morey et al. 2015). By joining this initiative, scholars agree to review manuscripts only if they are allowed by the journal editors to apply a number of open research standards (e.g., requiring that data and replication material are made available to the public upon publication, and that more methodological information is reported in the article).

Indirect influences on article transparency. The three direct influences–authors, data, and journals–operate in a broader context of academic and educational institutions that play an indirect but still salient role in the level of article transparency. We identify four indirect influences: data archives, academic professional associations, educational institutions, and publishing agencies.

First, data archives–such as the Roper Center, the Inter-university Consortium of Political and Social Research data archive (ICPSR) at the University of Michigan (USA), and the GESIS Institute for the Social Sciences (Germany)–offer a platform to store data and therefore have a lever to influence the quality of survey data and metadata. The data archives can provide various substantive and technical guidelines and best practices regarding how to standardize and present data and metadata to the public. They are also in the position of requiring a set of minimum methodological information as a condition for storing data. For instance, the Roper Center uses the minimum disclosure requirements of the American Association of Public Opinion Research (AAPOR) Code of Professional Ethics and Practice (2015). Before data are included in the archive, survey researchers have to provide complete questionnaires and codebooks, technical reports, weighting information, response rate calculation, and other documents (Maynard and Timms-Ferara 2011). Furthermore, ICPSR has created a detailed guide–Guide to social science data preparation and archiving–in consultation with experienced archivists and investigators. It includes step-by-step advice from the moment a grant proposal is written until the time the data are shared with the public.

Second, academic professional associations of survey researchers can take initiatives to prompt their members to be transparent in their publications. Various organizations have developed codes of ethics and reporting standards to encourage and guide survey researchers to prepare transparent and complete codebooks, for example the AAPOR Transparency Initiative and the Council of American Survey Research Organizations (CASRO) Code of Standards and Ethics for Survey Research. A review of four U.S. survey research association guidelines–AAPOR, CASRO, National Council of Public Polls (NCPP), and the European Society for Opinion and Marketing Research (ESOMAR)–shows major similarities in terms of reporting criteria, but also several differences. These relate to whether some methodological information should be disclosed when the data are released, or whether this should only happen on request. According to the NCPP, information about response rates, weighting procedures, and respondents' selection procedures can be accessed on written request (Maynard and Timms-Ferara 2011). Another example is the American Political Science Association (APSA), which has updated its ethical guidelines and has developed detailed data sharing and research transparency recommendations for authors (DA-RT - Guidelines for Data Access and Research Transparency in the Quantitative Tradition: Lupia and Elman 2014).

Third, authors' decisions on how to write up a study are influenced by their methodological skills and scientific attitudes. Educational institutions (such as universities or summer school programs) can influence the level of transparency by designing the curricula they offer. Scholars with access to classes on methods, research ethics, and research integrity during their graduate training programs are more likely to value research transparency, and to write transparent articles. The same goes for scholars working for universities with enforced research transparency policies and who have experience in keeping and disclosing detailed information about their research projects. For instance, after Diederik Stapel's data fabrication scandal, Tilburg University (in the Netherlands) put in place new rules and regulations for data handling, and introduced an annual audit committee that checks the quality of data storage and requires the reporting of the research method in a random number of articles. Additionally, researchers in the university are asked to sign an integrity code. Working in such environments might result in, or at least contribute to, a culture of data sharing and of preparing replication material.

Fourth, the decisions of journal editors can be influenced by the decisions made by publishing agencies. Many publishers, such as Elsevier or the Oxford University Press are members of the Committee on Publication Ethics (COPE), and encourage journal editors to become members and follow their guidelines. The main aim of COPE is to offer advice to publishers and editors concerning all matters of publication ethics and how to deal with research misconduct. For example, they offer guidelines and courses on topics such as plagiarism, authorship, conflicts of interest, and misconduct (COPE Code of Conduct for Journal Editors 2011).

4. Checklist for achieving higher transparency in cross-national survey research

In this section, we propose a checklist for reporting cross-national research and, in the subsequent section, this instrument is used to provide empirical evidence regarding the current reporting practices in cross-national studies. As a starting point for our checklist, we distinguish between the two types of transparency mentioned before: evaluation transparency and replication transparency. Even though most of the information needed to achieve the two types overlaps, there are some key differences. For example, to evaluate the findings of a study, a reader needs to have background information about the survey data (e.g., response rates, sampling frame, etc.), while this information is not needed for a direct replication. In order to conduct a direct replication, a scholar would need replication material with codes, while these are not needed to assess the validity and reliability of the results. Our checklist is furthermore based on the Total Survey Error (TSE) framework designed by Groves and colleagues (2004). This framework provides a comprehensive overview of the possible error sources in estimating survey statistics, and thus makes clear the methodological information that is necessary to assess the presence of biases and errors. In addition, we reviewed several publication and transparency guidelines in sociology, political science, and public opinion research (e.g., AAPOR Code of Ethics 2015; Code of Ethics of the American Sociological Association Committee on Professional Ethics 1997; DA-RT - Guidelines for Data Access and Research Transparency in the Quantitative Tradition: Lupia and Elman 2014), data archives guidelines (e.g., ICPSR provides a detailed guide: Guide to social science data preparation and archiving), and journals' submission guidelines. The final tool includes 23 items indicating what we consider to be the most crucial methodological information in cross-national survey analysis, covering among other things information regarding the survey data used, sampling, and individual and contextual variables (Table 1). We believe that this information is required for an article to be sufficiently transparent. As Table 1 makes clear, some pieces of information are required to evaluate the validity and reliability of the conclusions (column 2), while other parts are necessary for replication of the analysis (column 3). A second distinction can be made among the 23 indicators in the checklist. Some are generic for cross-national data collection (such as the sampling design, the survey mode, or the response rate), while others are specific for particular analyses (e.g., dealing with missing data, and the selection of variables).

Information about the survey data used. An obvious starting point is that authors have to clearly include the correct name for the survey data used in the article. For instance, the names of the European Values Study and the World Values Survey tend to be reported incorrectly. Further to the name, articles should mention the survey years or waves included in the analysis, and the version of the dataset(s), preferably in the form of a reference to the version. The dataset version is particularly important, because international surveys continue to be updated and it is common to find several versions of the same dataset. There are also integrated datasets with several waves available (e.g., the European Values Study Longitudinal Data File 1981–2008) or more than one survey (i.e., the Integrated Values Surveys 1981–2014 data file).

Information about sampling. According to the TSE framework (Groves et al. 2004), sampling is one of the two major components of errors in surveys. As a consequence, published research should provide sufficient information regarding the representation of the research population, as well as potential challenges (Häder and Lynn 2007). Some international surveys can have different age composition (e.g., the minimum age is 18 in the European Values Study and 15 in the European Social Survey) and the age categories of the country samples can also vary. For instance, Häder and Lynn (2007) explain that although the target population for the second wave of the European Social Survey comprised individuals aged 15 and above, it was not always possible to cover this population because some countries do not work with that low an age, or that they had an upper age limit of 75–80 years (e.g., Sweden). Therefore, as a starting point, comparative studies should give a description of the population to which conclusions are generalized. Otherwise, comparing countries with a different population composition can threaten the validity of findings.

Crucial information regarding the way in which a sample is drawn also needs to be provided. This includes disclosure of information about the sampling frame(s), the sample design, survey mode, and sample size (Harkness 1999; Lynn 2003; Siegfried and Häder 2010; Tille and Matei 2010). The sampling frame contains the population that has a non-zero likelihood of being selected, and accordingly it informs readers about how much of the population of interest (i.e., the target population) is covered by the survey. The sampling design (i.e., the concrete rules and procedures employed to select respondents based on the frame) are the basis for the representativeness and quality of the data. Because of constraints in time and budget, as well as differences in local conditions, cross-national surveys often contain variations in the survey design and survey mode, which can induce bias and as a result distort cross-national comparisons. To give an example, answers to sensitive questions such as personal income can significantly differ when the question is asked in a face-to-face interview, a telephone interview, or in a self-administered questionnaire (Lavrakas 2008). Surprisingly, researchers rarely report the differences in survey design across countries.

In order to give readers the opportunity to make a general assessment of the representativeness of the data, articles should contain information on the response rate(s) and indications of the presence of nonresponse and selection bias, as well as the eventual use of weights in the analysis (either to correct for selective nonresponse or to correct for unequal selection probabilities in the sampling design: Billiet, Koch, and Philippens 2007; Stoop 2010; Lavalle and Beaumont 2010). In addition to unit nonresponse, item nonresponse can also be an issue. Articles should therefore be transparent about the number of missing values and how they were treated (e.g., multiple imputation or *listwise deletion:* Bethlehem and Schouten 2010; Spiess 2010). After all, the choice of a particular treatment of missing data relies on different assumptions about the causes of missing data and could therefore influence the conclusions (Rubin 1976).

Information about the measurement of variables. In addition to sampling, measurement is the second key component of the TSE approach. Transparency thus implies that sufficient information regarding the concrete operationalization of the concepts in the study is disclosed. As a minimum, articles should include the variable names (as mentioned in the dataset) as well as a short description of all the variables and their meaning. However, this is not sufficient. Because survey measurement is highly susceptible to many types of error (Groves et al. 2008), it is good practice to disclose to the readers the exact formulation of survey questions and the answer categories. In cross-national survey research, individual survey measurements are often linked with contextual information at the country or regional level. For these contextual indicators, it is equally important to report similar information or an exact reference for all the contextual variables used from external sources (Breznau 2016; Fortin-Rittberger et al. 2016). Indispensable information includes the exact names of indicators, the year of the data, the source of the data, and when the data was retrieved.

Replication material. To achieve full replication transparency, an article should ideally include replication material with syntax codes for the various steps in the data preparation and analyses. Moreover, any additional information should be included about the analytical procedures performed from the moment the data were downloaded until it was ready for analysis. The syntaxes should be clean and clear, with short comments about the steps taken, if necessary. Researchers can upload all replication material on a public website or a free data repository, such as Dataverse or the Open Science Framework.

In principle, the information proposed in the checklist can be communicated to the scientific community in various ways: In the article itself, as online supplementary material on a website (e.g., of the journal), or by providing a reference to the source. In our opinion, the specific medium is of secondary importance, and what counts is that the information can be readily accessed.

[Table 1 about here]

5. Current reporting practices in cross-national survey studies: Evidence from 305 peer-reviewed articles (1986–2016)

In the previous section we proposed a checklist of 23 pieces of key information that cross-national studies should report in order to be fully transparent. The question remains, however, to what extent current reporting practices in cross-national survey research fulfill these requirements of transparency. This section provides empirical evidence regarding what type of and how much methodological information researchers disclose in articles reporting on their cross-national survey studies. We use the checklist presented in Table 1 as a benchmark, and thus focus on the type of information readers would need in order to evaluate the validity and reliability of a study's findings (evaluation transparency) as well as to carry out a direct replication (replicability transparency).

To evaluate current research practices, we analyzed a sample of 305 international survey-based articles. To select these articles, we first made a broad selection of 29 social science journals that regularly publish research based on cross-national survey data and that are written in English (see Appendix 1). From these journals, we then selected all the articles that (1) use as main data one of seven prominent international surveys that offer free data access, namely: Afrobarometer, Eurobarometer, European Social Survey, European Values Study, International Social Survey Program, World Values Survey, and Latinobarometer; (2) make a comparison across at least five countries; and (3) do not have a purely methodological goal, but answer a substantive research question. The resulting articles can cover one or multiple waves (time points), and use a wide variety of analytical tools. With regard to the selection of surveys, we focus on general-topic, general-population opinion surveys that are conducted cross-nationally and offer free data access (Smith and Fu 2010). Therefore, we excluded any international surveys that focus on specific populations and/or specialized topics, because they have specific reporting practices that might be customary in particular subfields, for example the Survey of Health, Ageing, and Retirement in Europe (SHARE) and the International Reading Literacy Study (PISA). The data collection took place between November 2016 and January 2017. Three students were trained and closely supervised in manually checking approximately 3700 articles that mention the name of one of the surveys in any part of their text and then we selected only those that met the above-mentioned criteria. This resulted in a sample of 1007 comparative studies, from which we then drew a random sample of 305 articles. In the period from March to May 2017, one of the authors together with one master's student read each article and recorded the methodological information that was present in the articles. Concretely, we checked for each of the 23 indicators in the checklist (see section 4) whether the required information was present (score 1) or absent (score 0) in the article, supplementary material, or as a reference to an external source. Our focus was on the mere presence of methodological information, and we did not check whether the disclosed information was accurate or complete. Intercoder reliability was assessed by computing the Krippendorff's alpha and percentage agreement scores using the software ReCall (Freelon 2010). The Krippendorff's alpha intercoder reliability between the two coders is 0.81, which is satisfactory, ranging from 0.47 to 1 across variables, and the average percentage agreement between the two coders is 94 percent, ranging from 80 to 100 percent across the variables.

To provide an indication of the overall level of transparency in cross-national research, we also created an evaluation transparency index (ETI) and a replicability transparency index (RTI), which are calculated as the mean of the various items in the checklist (Table 1). For the measurement of ETI we used items number 2, 5, 6, 7, 8, 9, 11, 13, 14, 15, 17, 18, 21, and 22, and for RTI we used items number 2, 3, 11, 13, 14, 15, 17, 18, 21, 22, and 23.[1, 2] The resulting indices range from 0 to 1, with a higher value representing higher transparency.

Results. Table 2 and Table 3 provide per indicator the percentage of articles that mention the relevant methodological information regarding the survey data set used, the sampling/representation and the measurement of the variables.[3] These descriptive results are striking, in the sense that a large proportion of the articles fail to provide the information that is crucial to evaluate the representativeness of the survey data. Specifically, Table 2 shows that most articles mention how many countries are included in the comparison (100 percent), provide a list of countries analyzed (88.9 percent), and disclose the total size of the sample (81.6 percent). The mode of data collection, however, is only mentioned in 16.4 percent of the studies, notwithstanding the fact that the survey mode is a key element in the collection of survey data.

Information on how the sample was obtained is often missing as well. Only one out of five studies provides a description of the population to which the conclusions are generalized, 38.7 percent provide at least some information regarding the sampling design, and the sampling frame(s) used are almost never mentioned (2.6 percent). The response rate–although a crucial indicator of data quality–is disclosed by less than one article out of ten, and only 2.6 percent of the articles reflect on possible nonresponse bias. In a similar vein, a mere 9.5 percent of the analyzed articles mention the number of missing values (item nonresponse), and about one third (30.5 percent) describe how these missing values were dealt with. Only 28.5 percent of the studies mention the use of weights. This is surprising given that most cross-national surveys advise users to include weights. For instance, the European

Social Survey has a relevant guide–Weighting European Social Survey Data (2014)–which clearly states from the first page that users should always use weights if they want to obtain accurate estimates. In sum, these findings reveal a surprising lack of transparency regarding sampling or representation. Only a small minority of the published cross-national studies provide the information that is necessary for the reader to assess the representativeness of the data.

The results for the indicators regarding measurements are shown in Table 3. With regard to the measurement of individual and contextual data, we can see that a substantial proportion of the articles fail to include the necessary methodological information (although the lack of transparency is less severe than in the case of sampling issues). Some 70 percent of the articles report the exact survey question by which the dependent variables are measured, while 58 percent report the full range of the answer categories. These values decrease when it comes to aggregated individual variables, with 34 percent providing information about the measurements and the answer categories for all the variables. With regard to contextual variables from external sources, it is interesting to note that only 35 percent give the year of the data, 78 percent the source of the data, and 43 percent an exact reference to the data sources for all the contextual variables included in the study.

Our results furthermore show that sharing replication material is not customary in cross-national survey research. Of the 305 articles in our study, only 14 (4.6 percent) refer to the availability of full or partial replication material. In three cases, partial replication material is included in the article or online supplementary material, in five other cases the article mentions that full or partial replication material is available from the author(s) on request. In the six remaining cases, a reference is provided to an external website with replication material (e.g., Dataverse or the author's website). However, of the six articles that provided such a reference, we were able to access only two while for the other four articles the link did not work or we could not find any replication material on the website mentioned. For example, in one article it was stated that a folder with all replication material had been uploaded to Dataverse, but we could not find it.

[Table 2 about here]

[Table 3 about here]

Lastly, the various indicators are combined into two transparency indices, reflecting the evaluation transparency (ETI) and the replicability transparency (RTI). For the overall level of evaluation transparency index of cross-national studies, Graph 1 shows that 69 percent of the articles have an ETI below 0.5. This means that more than half of the articles provide only half or less of the information needed to evaluate the validity and reliability of the findings. Furthermore, Graph 2 indicates that 46 percent of the articles provide half or less of the information needed to replicate the study. It is also interesting to mention here that of the 15 articles with an ETI between 0.8 and 1, five were published in Public Opinion Quarterly (POQ). Of the 23 articles with an RTI between 0.8 and 0.9, three were published in the same journal (results not shown in the tables or graphs). One explanation is that out of all the 29 journals included in the study, POQ is the only one that states in its submission guidelines that authors *must* include methodological information about the data used (e.g., response rate, survey mode, etc.). These findings are very startling, as they show an overall picture of the level of transparency in this type of research. Even though all the articles include some methodological information, the great majority lack sufficient information for evaluation and replication.

[Graph 1 about here]

[Graph 2 about here]

6. Discussion and conclusions

For the first time, this study provides empirical evidence of the state of transparency in cross-national survey research in social sciences. Based on a sample of articles published in the last 30 years, our findings reveal that most cross-national studies provide little of the methodological information needed to evaluate the validity and reliability of their findings or to reproduce them. Information that is of paramount importance to assess data quality, such as the response rate, survey mode, or sampling frame, is only reported in a small proportion of published articles. Moreover, providing replication material is still exceptionally rare in this field. We show that even though the data are public, the particular character of cross-national research poses specific challenges to research transparency. These challenges are, among other things, connected with the separation between data collection and data analysis (the two tasks are often performed by different teams), and the sheer complexity of collecting and analyzing cross-national survey data. As a result, the risk exists that researchers fail to disclose key methodological information about the survey data and the steps taken in data preparation and analysis.

Achieving greater transparency in this field is a shared responsibility between various actors in the academic community. The contents of an article are not only shaped by its authors, but also directly affected by the editorial policies of journals and features of the datasets analyzed. After all, the quality and completeness of the documentation of survey data is a key factor that influences the level of transparency. Furthermore, these three direct influences on research transparency–authors, journals, and data–operate in a broader institutional context. Educational institutions, data archives, professional organizations, and publishing agencies have the power to shape opportunities, but also to create obstacles to research transparency.

The main message of this paper is that in the absence of mandatory requirements, there are still many ways in which members of the academic community can help to increase the level of transparency of scientific works. In this regard, our checklist for reporting cross-national survey research can be a useful tool to aid researchers in writing up their own papers, and when they act as peer-reviewers, members of editorial journal boards, funding agencies, professional organizations, or educators. We hope that our findings will also motivate journals and publishing agencies to adopt this tool and create stronger transparency requirements for publication. In this regard, we also encourage journals to use additional ways to promote research transparency and open science practices in general. For example, by using a badge for a high evaluation or replicability transparency index, or the existing Open Science Badges for data sharing, preregistration of studies, and replication documents. Furthermore, as a solution to the extensive data and metadata available, and the difficulty of going through a tremendous amount of information, survey organizers could use our checklist to create a brief fact sheet containing the key information about data collection and data quality (e.g., response rates, missing values, the availability of weights, etc.). This can be an extra resource for researchers to understand the strengths and weaknesses of their data, and provide all relevant information to their readership. Educational institutions and funding agencies can also play a key role in promoting research transparency by adding open science practices to the list of requirements for obtaining tenure, a promotion, or research grants. Additionally, universities can offer training to both their staff and students on the topic, such as workshops about the preparation and storage of replication material or hands-on replication assignments in graduate methods courses (e.g., using ReplicationWiki database and teaching materials: Höffler 2017).It is important to mention here that even though our study offers information about transparency in cross-national survey research, the findings and the applicability of our checklist could be generalized to other fields using publicly available survey data with nested structures (e.g., national surveys that contain data about individuals nested in municipalities or neighborhoods). Furthermore, we believe that the low levels of transparency observed here are indicative of a general lack of transparency in the fields of sociology and political sciences. This situation can be partly explained by the fact that, currently, academics continue to be highly rewarded (e.g., career advancement and research grants) by the quantity of academic work published and the impact factor of the journals in which they publish, while there are still no real rewards for investing time in producing transparent research. We hope that our study, along with other research into the lack of transparency and reproducibility of results in the social sciences, can contribute to a change in this culture.

Notes

1. We did not collect information about the measurement of independent and control variables, as this would have significantly increased the time spent on each article. However, having information about the measurement of dependent and aggregated variables represents a good indication of the information given for the other variables.
2. Not all the items included in the evaluation and replicability checks were used to create these two indices. The item regarding the name of the survey was left out because it was one of the initial selection criteria for articles, and therefore we had this information for all the articles beforehand. Nonresponse and selection bias, the number of missing values, and the description/meaning of variables were also left out, as they are very similar to other items in the list. For example, for the items regarding response rate and nonresponse, we chose to include the former, for the items regarding the description of the variable and the exact item/indicator we chose the latter, and so on. Furthermore, the checklist items regarding the independent variables were skipped as well, as we did not recode this type of information (see note 1).
3. Table 2 and Table 3 include some items about additional methodological information that are not found in the checklist (Table 1). As the first step in our study, we were interested in offering insights into what kind of and how much information authors provide in the empirical sections of their articles. For example, with regard to the measurement of a variable, we recorded whether an article provides a description of the operationalization, the exact question, question number/code, and/or a reference to the questionnaire. However, we do not expect researchers to provide all this information in an article. The checklist in Table 1 contains the basic information that we recommend researchers to report.

References

Abbott, Andrew. 2007. "Notes on Replication." *Sociological Methods & Research* 36(2): 210–219.

American Association for Public Opinion Research. 2015. *AAPOR Code of Professional Ethics and Practices*. Oakbrook Terrace, IL: AAPOR.

American Sociological Association. 1997. "Code of Ethics of the American Sociological Association Committee on Professional Ethics." Retrieved April 5, 2018 (http://www.asanet.org/sites/default/files/code_of_ethics_aug_2017_2_1.pdf).

Baker, Monya. 2016. "Is There a Reproducibility Crisis?" *Nature* 533:452–454.

Berry, John W., Ype W. Poortinga, Marshall H. Segall, and Pierre R. Dasen. 1992. *Cross-Cultural Psychology: Research and Applications*. 2nd ed. New York: Cambridge University Press.

Bethlehem, Jelke and Barry Schouten. 2010. "Nonresponse error: Detection and correction," pp. 558–578 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Billiet, Jaak, Achim Koch, and Michel Philippens. 2007. "Understanding and improving response rates," pp. 113–138 in *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, edited by Roger Jowell, Caroline Roberts, Rory Fitzgerald, and Gillian Eva. Los Angeles: Sage Publications.

Breznau, Nate. 2016. "Secondary observer effects: idiosyncratic errors in small-N secondary data analysis." *International Journal of Social Research Methodology* 19(3):301–318.

Brooks, Clem and Jeff Manza. 2006. "Social policy responsiveness in developed democracies." *American Sociological Review* 71:474–494.

Carsey, Thomas M. 2014. "Making DA-RT a Reality." *Political Science and Politics* 47:72–77.

Committee on Publication Ethics. 2011. *Code of Conduct and Best Practice Guidelines for Journal Editors*. London: COPE.

Dafoe, Allan. 2014. “Science Deserves Better: The Imperative to Share Complete Replication Files.” *Political Science & Politics* 47:60–66.

European Social Survey. 2014. “Weighting European Social Survey Data.” Retrieved March 20, 2018 (https://www.europeansocialsurvey.org/docs/methodology/ESS_weighting_data_1.pdf).

Freelon, Deen. 2010. “ReCal: Intercoder Reliability Calculation as a Web Service.” *International Journal of Internet Science* 5(1):20–33.

Firebaugh, Glenn. 2007. “Replication Data Sets and Favored-hypothesis Bias. Comment on Jeremy Freese (2007) and Gary King (2007).” *Sociological Methods & Research* 36(2):200–209.

Fanelli, Daniele. 2009. “How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-analysis of Survey Data.” *PlosOne* 4(5):e5738.

Freese, Jeremy. 2007. “Replication Standards for Quantitative Social Science: Why not Sociology?” *Sociological Methods & Research* 36(2):153–72.

Fortin-Rittberger, Jessica, David Howell, Stephen Quinlan, and Bojan Todosijevic. 2016. “Supplementing cross-national survey data with contextual data,” pp. 670–80 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Gelman, Andrew and Eric Loken. 2014. “The Statistical Crisis in Science. Data-dependent Analysis—A “Garden of Forking Paths”— Explains Why Many Statistically Significant Comparisons Don't Hold up.” *American Scientist* 102(6):460–65.

Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2004. *Survey Methodology*. New York: Wiley.

Harkness, Janet. 1991. “In pursuit of quality: Issues for cross-national survey research.” *International Journal of Social Research Methodology* 2(2):125–40.

Harkness, Janet A., Fons J. R. Van de Vijver, and Peter P. Mohler. 2003. *Cross-Cultural Survey Methods*. New York: Wiley.

Harkness, Janet A., Michael Braun, Brad Edwards, Timothy P. Johnson, Lars E. Lyberg, Peter P. Mohler, Beth-Ellen Pennell, and Tom W. Smith. 2010. *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. Hoboken, New Jersey: Wiley.

Häder, Sabine and Peter Lynn. 2007. "How Representative Can a Multi-Nation Survey Be?" pp. 33–52 in *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, edited by Roger Jowell, Caroline Roberts, Rory Fitzgerald, and Gillian Eva. Los Angeles: Sage Publications.

Höffler, Jan H. 2017. "ReplicationWiki—Improving Transparency in Social Sciences Research." D-Lib Magazine 34(3/4):1-6.

Huizenga, Hilde M., Ruud Wetzels, Don van Ravenzwaaij, and Eric-Jan Wagenmakers. 2012. "Four Empirical Tests of Unconscious Thought Theory." *Organizational Behavior and Human Decision Processes* 117:332–40.

Ioannidis, John P.A. 2005. "Why Most Published Research Findings are False." *PLoS Medicine* 2(8):696–701.

Ishiyama, John. 2014. "Replication, Research Transparency, and Journal Publications: Individualism, Community Models, and the Future of Replication Studies." *PS: Political Science & Politics* 47:78–83.

John, Leslie K., George Loewenstein, and Drazen Prelec. 2012. "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling." *Psychological Science* 23(5):524–32.

King, Gary. 1995. "Replication, Replication." *PS: Political Science & Politics* 28:444–52.

Lucas, Jeffrey W., Kevin Morrell, and Marek Posard. 2013. "Considerations on the 'Replication Problem' in Sociology." *The American Sociologist* 44:217–32.

Lavalle, Pierre and Jean-Francois Beaumont. 2010. "Weighting: Principles and practicalities." Pp. 460-476 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Lavrakas, Paul J. 2008. *Encyclopedia of Survey Research Methods*. Sage Publications: USA.

Lupia, Arthur and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science & Politics* 47:19–42.

Lyberg, Lars E. and Herbert F. Weisberg. 2010. "Total survey error: A paradigm for survey methodology," pp. 27–40 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Lynn, Peter, Lilli Japec, and Lars Lyberg. 2007. "What's so Special About Cross-National Surveys?" pp. 7-20 in *Conducting Cross-national and Cross-cultural Surveys*, edited by Janet Harkness Manheim: ZUMA.

Lynn, Peter. 2003. "Developing quality standards for cross-national survey research: Five approaches." *Social Research Methodology*, 6:323–36.

Maynard, Marc and Lois Timms-Ferrara. 2011. "Methodological Disclosure Issues and Opinion Data." *Journal of Economic and Social Measurement* 36:19–32.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Morey, Richard D., Christopher D. Chambers, Peter J. Etchells, Christine R. Harris, Rink Hoekstra, Daniël Lakens, Stephan Lewandowsky, Candice Coker Morey, Daniel P. Newman, Felix D. Schönbrodt, Wolf Vanpaemel, Eric-Jan Wagenmakers, and Rolf A. Zwaan. 2016. "The Peer Reviewers' Openness initiative: Incentivizing Open Research Practices through Peer Review." *Royal Society Open Science* 3:150547.

Munafò, Marcus R., Dorothy V. M. Bishop, Katherine S. Button, Christopher D. Chambers, Nathalie Percie Du Sert, Uri Simonsohn, Eric-Jan Wagenmakers, Jennifer J. Ware, and John P. A. Ioannidis. 2017. "A manifesto for reproducible science." *Nature Human Behaviour* 1:1–9.

Nelson, Leif D., Joseph Simmons, and Uri Simonsohn. 2017. "Psychology's Renaissance." *Annual Review of Psychology* 69:17.1–17.24.

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." pp. aac4716 in *Science 349*.

Pennell, Beth-Ellen, Janet A. Harkness, Rachel Levenstein, and Martine Quaglia. 2010. "Challenges in Cross-National Data Collection," pp. 269–98 in *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, edited by Janet A. Harkness, Michael Braun, Brad Edwards, Timothy P. Johnson, Lars Lyberg, Peter P. Mohler, Beth-Ellen Pennell, and Tom W. Smith. New Jersey: Wiley.

Rubin, Donald B. 1976. "Inference and missing data." Biometrika, 63: 681-592.

Siegried, Gabler and Sabine Häder. 2010. "Special challenges of sampling for comparative surveys," pp. 346–56 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Silberzahn, Raphael, Eric L. Uhlmann, Daniel P. Martin, Pasquale Anselmi, Frederik Aust, Eli C. Awtrey, Štěpán Bahník, et al. 2018. "Many Analysts, One Dataset: Making Transparent How Variations in Analytical Choices Affect Results." *Advances in Methods and Practices in Psychological Science* 1:337–356.

Smith, Tom W. and Yang-chih Wu. 2010. "The globalization of surveys," pp. 680–92 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Simmons, Joseph, Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allow Presenting Anything as Significant." *Psychological Science* 22:1359–66.

Spiess, Martin. 2010. "Dealing with missing values," pp. 597–612 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Steneck, Nicholas H. 2006. "Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions." *Science and Engineering Ethics* 12:53–74.

Stoop, Ineke A. L. 2016. “Unit Nonresponse,” pp. 409–24 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

Tille, Yves and Alina Matei. 2010. “Basics of sampling for survey research,” pp. 311–28 in *The SAGE Handbook of Survey Methodology*, edited by Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-chih Fu. Los Angeles: Sage.

van de Vijver, Fons J. R. and Kwok Leung. 1997. *Methods and Data Analysis for Cross-Cultural Research*. Thousand Oaks, CA: Sage.

Wicherts Jelte M., Coosje L. Veldkamp, Hilde E. M. Augusteijn, Marjan Bakker, Robbie C. M. van Aert, and Marcel A. L. M. van Assen.2016. “Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking.” *Frontiers in Psychology* 7:1–12.

Wolf, Christof, Dominique Joye, Tom W. Smith, and Yang-chih Fu. 2016. *The SAGE Handbook of Survey Methodology*. Los Angeles: Sage.

Appendices

Appendix 1. Number of articles by journal

Journal	Number of articles
Acta Politica	10
Acta Sociologica	11
American Journal of Political Science	8
American Journal of Sociology	9
American Sociological Review	11
Comparative European Politics	9
Comparative Political Studies	10
Cross-Cultural Research	11
European Journal of Political Research	11
European Political Science	3
European Societies	15
European Sociological Review	22
International Journal of Comparative Sociology	10
International Journal of Public Opinion	11
International Journal of Social Welfare	11
International Political Science Review	11
Journal of Cross-Cultural Psychology	11
Journal of European Public Policy	8
Journal of European Social Policy	10
PlosOne	8
Political Studies	9
Politics and Policy	7
Public Opinion Quarterly	7
Social Forces	11
Social Indicators Research	26
Social Problems	6
Social Science Research	11
The Journal of Politics	10
West European Politics	8
Total	**305**

Figure 1. Conceptual framework of the actors, factors, and processes playing a role in the transparency of an academic article

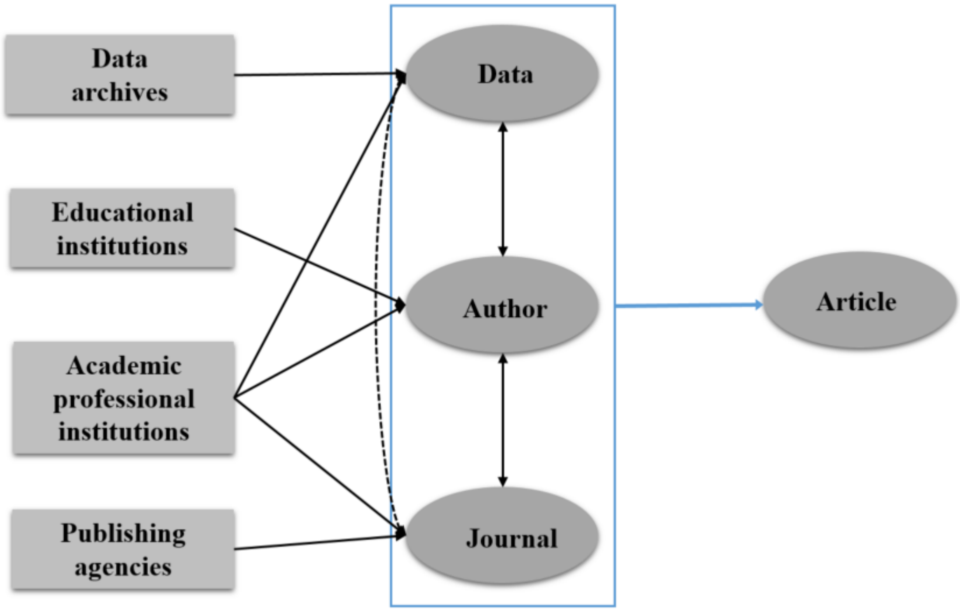


Table 1. Checklist items for reporting cross-national survey research

No.	Item	Evaluation transparency checks	Replicability transparency checks
	Information about survey dataset		
1	Name of the survey	√	√
2	Survey years used	√	√
3	Dataset version(s)		√
4	Reference to the dataset		√
	Information about sampling/representation		
5	Description of the population	√	
6	Sampling frame	√	
7	Sample design	√	
8	Survey mode	√	
9	Response rate	√	
10	Nonresponse and selection bias	√	
11	Weighting procedures used	√	√
12	Number of missing values	√	
13	Treatment of missing values	√	√
14	Study sample size	√	√
15	List of countries	√	√
	Measurement of individual variables		
16	Description/meaning of the variables	√	
17	Exact question or item	√	√
18	Full range of answer categories	√	√
	Measurement of contextual variables		
19	Description/meaning of the variables	√	
20	Code of the original indicator		√
21	Year of the data used	√	√
22	Source of the data	√	√
	Replication materials		
23	Syntaxes of data preparation and statistical analyses		√

Table 2. Percentage of articles disclosing information about data and sampling (N=305)

No.	Items	% of articles mentioning the relevant information
5	Description of the population	19.3
6	Sampling frame	2.6
7	Sample design	39.7
8	Survey mode	16.4
9	Response rate	9.5
10	Nonresponse and selection bias	2.6
11	Weighting procedures used	28.5
12	Number of missing values	9.5
13	Treatment of missing values	30.5
14	Study sample size	81.6
-	Study sample size per country	27.2
-	Number of countries included	100
15	List of countries	88.9

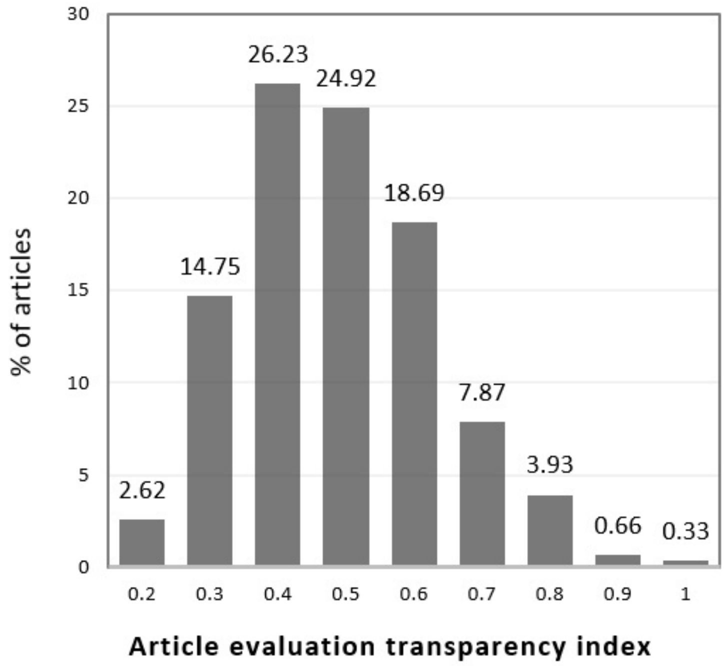
Note: The item numbers correspond to the ones in the checklist (Table 1).

Table 3. Percentage of articles disclosing information about measurements

No.	Items	Yes, for all variables %	Yes, for some variables %	No %	N
	Dependent variables				
16	Description of the operationalization	97.4	0	2.6	305
17	Exact question or item	70.2	5.6	24.3	305
-	Question number	13.4	0	86.6	305
18	Full range of answer categories	57.7	6.2	36.1	305
-	Reference to the questionnaire	0.7	0	99.3	305
	Contextual variables (aggregated)				
-	Description of the operationalization	100	0	0	67
-	Exact question or item	34.3	6	59.7	67
-	Question number	10.5	1.5	88.1	67
-	Answer categories	34.3	9	56.7	67
-	Reference to the questionnaire	0	0	100	67
	Contextual variables (external sources)				
19	Description of the operationalization	94.6	4.9	0.5	203
20	Code of the original indicator	1.5	0	98.5	203
21	Year of the data used	35	33.5	31.5	203
22	Source of the data	77.8	18.2	4	203
-	Reference to the data source	42.9	28.1	29.1	203

Note: The item numbers correspond to the ones in the checklist (Table 1).

Graph 1. Percentage of academic articles by evaluation transparency index (N=305)



Graph 2. Percentage of academic articles by replicability transparency index (N=305)