Data & Documentation Quality Check:
Cross-national Surveys Quality in the Light of Data Harmonization

Ilona Wysmułek, Olena Oleksiyenko, Anastas Vangeli

> To err is human, to forgive divine – but to include errors in your design is statistical
>
> Leslie Kish (1978)

Introduction

The issue of the survey data quality has been the object of heated debate and continuing concern of sociologists ever since the survey method itself has been introduced. Although the researchers agree about the numerous benefits the mentioned method of data gathering gives, the key question of how to control or verify the quality of the obtained information remains. By now there have been quite a few attempts to introduce the standard of survey documentation, aiming at transparency of the each stage of the data gathering and coding. However, there are no universal requirements accepted by all, neither is there the unified institution that has a power (or capacity) to ensure that the following data is of a high quality. Undoubtedly, the bad quality of the survey data undermines the scientific value of the achieved results using this data, thus this question is of a great importance. However, between the clearly good and clearly bad survey there is the whole range of a decent quality which has potentially interesting and important information often from the difficult-to-reach places. To date there has been little agreement on how to compare the quality of surveys, how far the tolerance to human mistakes can go and what is the best way to fix the errors if found.

Within the project "Democratic Values and Protest Behavior: Data Harmonization, Measurement Comparability, and Multi-Level Modeling in Cross-National Perspective" (called in the paper *Data Harmonization Project*), we embarked on a strenuous task of assessing the quality of 22 different cross-national survey projects, having 89 waves, 1720 surveys of the country level, which embeds the information of around 140 countries since 1966 up till 2013. The project *Quality Check* has been initiated by the research team[1] with the aim to introduce the implications of quality control in the process of ex-post data harmonization. We focused on three types of quality assessment: a) the quality of the

[1] Despite the authors of the paper, the project was conducted also by Jakub Wysmułek and Anna Franczak at the beginning stage of the research

documentation b) the quality of the data and c) the quality of the resemblance between the documentation and the data and vice versa. The latter will be a focus of this paper.

In this paper we will start from the theoretical analysis of the different approaches to quality check and proceed with highlighting the importance of quality in ex-post harmonization project. Further on we concentrate on the type of quality check which although being crucial was underestimated in the previous studies, namely quality of the resemblance between the documentation and data. We locate our study in the literature of the present taxonomies of quality checks in surveys. After describing what exactly we were doing and how we approached it, we come to the analysis of the errors and their types that we have found by comparing different sources of documentation and data. The paper finishes with the proposition of quality check index that measures the survey and survey-waves in terms of the average number of errors per wave accounting for the number of items under analysis.

Why Quality Check Matters

One of the deadly sins in the area of science is the lack of objectivity, as to Max Weber. Weber (1949) understood objectivity here as the independence from the special, 'one-sided' view-point. Franciszek Sztabiński (2011) has called this type of objectivity the „external" approach, free of subjective evaluation by the researcher. Another type of objectivity distinguished by Sztabiński is connected with the quality of the obtained scientific results, meaning how trustworthy and clean from errors they are. The ambiguous nature of objectivity in science lays in the fact that although both dimensions are crucial for the legitimization of the results, none of them is possible to reach fully. What does keep science from crisis then? The great achievement of science and the basis of its legitimacy is the awareness of where the error can be, the knowledge of how to check it and the procedures of how to account for it. „The awareness of the measurement error is an evidence of the development of science" (Hyman in Sztabiński, 2011).

Tracing back the development of survey methodology it can easily be noticed, that the major attention has been put to the errors occurring during the sampling procedures. While the main puzzle for some time has been the verification of the assumption that sample can represent the whole population, the field of survey quality control was dominated by the sampling challenges. But is it only bad sampling that can undermine the results taken from the survey? In this regards, the work of Deming „On Errors in Surveys" published in 1944 has

been revolutionary, while it underlined that there are different types of errors possible on every stage of survey design that can dramatically influence the end-product. The argument that the good quality of every stage of research equals good quality of the survey was very convincing, and since that time the debate has shifted rather into the adding of some new elements to the existing categorizations or creating some new taxonomy of survey errors. The overall aim became creating the enhancive checklist of all possible errors that are to be checked in order to ensure the high quality of the survey. The approach in which the attention is put to the multiple error sources has been called a total survey error (TSE). However, the quality of survey can be also assessed by the approach called „total quality management" (TQM), which emphasizes that not only all stages of survey production by themselves, but also in correlation to each other they have a direct and indirect influence to the overall quality of the end-product. In TQM compared with TSE approach „Next to accuracy, the relevance, comparability, coherence and completeness come under scrutiny" (Loosveldt, Carton, Billiet, 2004, p.66).

According to Biemer (2003) recent quality assessment are more concerned with control of process quality than evaluation of the final product. The idea behind this shift is that quality of survey data can be achieved through quality control on each stage of the research. This approach is supported by Groves (1989) stating that the picture of total survey error is not complete without different types of errors caused by various mistakes on the stage after data collection, such as "coding, editing, imputation, and other data processing activities that follow the data collection phase" (Groves,1989 p.12). The assessment of the quality of preparation processes is done in order to control for two main types of errors: non-observational (sampling, coverage, non-response) and observational (mode error, instrument error, interviewer error, respondent error). Groves (1989) claims that the total survey error is impossible to describe without assessing the quality of data processing, however his work concentrates more on the coverage error, nonresponse error, sampling error, measurement error, which consists of interviewer's effect, respondent's error, error due to instrument weakness, and error due to the effect of data collection mode.

The question appears: is it possible to control for all possible errors (and their correlations) and if yes, than how? The question of Groves (1989) in this regard went even further: assuming that we want to control for all possible errors, how much money and time will it take? The choice between what phases of the research are to be controlled for quality

and how this control should be done seems to be in practice the balance between the gains of such control and the time and cost spend on it. Thus, we face here the other challenge: not only the quality check is important, but some quality check categories might be more important than the other and the survey principle investigator needs to decide what, how and when to control and check.

Processing Errors Underestimated?

Out of all complicated elements of the preparation and realization of the field work, it seems that for the control of quality of coded data has been put not that much of an attention. Sztabiński (2011) had mentioned that in the stage of data coding, the quality control can be achieved by having two independent research teams coding the data and comparing the end results (p.32-33). However, how often is it done in practice? Unfortunately, in this regard the processing errors differ from the other error categories. Although most TSE taxonomies theoretically include it as a component, „it is too rarely included in models of survey error" (Groves, 2010, p.869). The question arises, then, is it because the processing steps, such as coding, are less error-prone than measurement or nonresponse errors, for example? As the response to this question, Biemer and Lyberg (2003) underline that the risk of making errors in coding are very high and influential, especially in such complex concepts as occupation and education. „For instance, coding error rates or coding disagreement rates can reach 20% levels for some variables, especially if the coding staff is not well trained" (Biemer et al. 2003, p. 219).

All things considered, it seems that researchers aiming for objectivity and lack of bias in the results should care (and usually do care) about the quality of the sources that they are using. In this case, survey as the way of data gathering is a challenging although rich tool for analysis. On the one hand, it is a tool that can give us (with a quite satisfactory accuracy) the information about the whole population. On the other hand, this is the process that contains a number of stages, each of which is prone to errors.

By now, there have been little attention put to the errors that can occur on the stage of data processing, although there are some empirical evidence that at this stage there can occur both systematic and variable errors which might (similarly to measurement errors, sampling errors and nonresponse for example) undermine the overall reliability on the survey.

Quality Check Typologies: What exactly we were doing?

The central point of interest of our project is the observation error, which can occur at the process of data processing and refers to the transparency and consistency of documentation. Being more specific, in this quality analysis we are concentrating on the „processing errors" (category first introduced by Deming in 1944 and presently used in almost all modern taxonomies of errors, including categorization of Anderson et al. (1979), Groves (1989), and Biemer et al. (2003)), which is also called „survey administrative issues" in the work of Herbert Weisberg (2005) and „compiling errors" in Hansen et al. (1953).

If to import Lynn's (2003) taxonomy, our approach classifies as one seeking "maximum quality" as our work focused on individual variables, and to a lesser extent on the bigger picture; it also used pre-defined notions of quality, and any divergences from this standard were considered an error. For this task, we focused less on consistency and comparability across different waves and countries (those aspects were given more attention in the harmonization work).

Data Harmonization: Implications for Quality

Checking the data quality is important in every survey, but, as to Smith (2009) it is especially significant in international survey projects which aim to be comparable. In this regard, putting all major international surveys together in the one box creates even more concerns in terms of data quality. The concerns deal here with the two main puzzling questions: 1) what exactly are we putting in the common file (meaning what are the strong points and weaknesses of the data sets that we are working with) and 2) if we determine some weaknesses, how can we fix them (and to which extent does it make sense to fix them)? As Słomczyński (2014, May) labeled those concerns, in data harmonization we ask whether we put in 'garbage' and how to 'recycle' it. Without questioning the potential gains of such harmonized data, the challenges of bringing all surveys together has not yet been faced in the academic world. Summing up the pitfalls of harmonization in the context of the quality of the end product, in order for such harmonized data to be usable, it requires a special attention to the following issues, such as a) bringing all surveys to the same standard b) controlling for errors imbedded in surveys c) deciding (or asking) for the right to fix them d) controlling for errors of ex-post harmonization itself and e) having some degree of confidence in cross-national surveys generally and harmonized data of them specifically. All things considered,

there is the sense in doing data harmonization only with the thorough quality check. Thus, a special attention and thorough control should be put in order not to have a „garbage in garbage out“ product, but rather the „data recycling“, meaning the reasonable quality data ready for use for academic analysis.

The Data Harmonization Project itself has a complex “life cycle” as it consists of a “series of interlinked and often iterative processes” (Mohler et al. , 303) and calls for meticulous and incremental implementation. The transition from one step to another was taken with much precaution. The quality assessment task followed a process of identifying variables in the myriad of surveys that could be matched with the target variables considered in the project, and subsequently harmonized. It is throughout this process that the research team has been acquainted with the basic features of the documentation, and contemplated ideas regarding the quality assessment. Thus, there are three main motivations that have caused the appearance of the *Quality Check: between Data and Documentation* project: a) during the process of the selection of needed variables across surveys there have been observed the vast range of 'documentation styles', the evaluation of which often depended on subjective taste of the researcher b) while merging the results of data collection process we have noticed that there is no common understanding of what is the best quality while the criteria of its assessment varies c) there have appeared the need to control how much of a “garbage“ we are putting in and how can we fix it in order for the end result to be usable.

How it all started? Context for Data and Documentation Project

Using the terminology of Wang & Strong (1996), in the *Data Harmonization* project we have assessed the survey projects as consumers, expecting a good quality of the products in order to continue work with them. Wang & Strong (1996) call it "fitness for use", emphasizing that any producer of good must not only take care of the best possible quality of it, according to their own perspective, but also they are obliged to remember to take into consideration consumer’s point of view, because consumer as an ultimate user will decide whether the product fits their needs. There are several important features of the data, which is survey data in our case, Wang & Strong (1996) point out as crucial for quality assessment, they are: accessibility, interpretability, relevance and accuracy. Applying this criteria, we will briefly describe the data that we work with, which is an important context of the *Data and Documentation Quality Check* project.

- **Accessibility:** meaning data must be easily accessible by user without any additional search or obstacles as, for instance, lack of possibility to retrieve data.

In DATA HARMONIZATION PROJECT: All surveys that we examine are academic surveys available in the public domain. Most international surveys are stored in archives with no limitations to use it as a secondary data for research, but in case of Asian Barometer (wave1-3), only questionnaires are available at the survey official web page, the access to data sets requires contact with principal investigators to get permission.

- **Interpretability**: consumer must be able to understand and interpret the data. One of the examples is that data cannot be presented in a foreign language.

In DATA HARMONIZATION PROJECT: The important criterion of choice of cross-national surveys for the analysis was the availability of documentation in English. Considering the scope of the project, despite the clear criterion on English translation, we have encountered some problems, such as the lack of sufficient information concerning different questions. For example, in Americas Barometer the data set is available only in Spanish, thought questionnaires are translated to English. Using data from this survey is tricky, because in some questions the English wording of answers differs from Spanish.

- **Relevance:** data must be adequate to consumer's needs and also released on time.

In DATA HARMONIZATION PROJECT: Among cross-national academic surveys, which are publically available and documentation of which is translated into English we have selected the ones fitting our academic interest. We concentrate on surveys of adult population that (among other) contain questions on protest behavior and trust to public institutions, as well as questions on democratic values.

- **Accuracy:** data must be accurate, meaning objective, correct and error-free.

In DATA HARMONIZATION PROJECT: We encountered both insufficient information in codebooks and datasets, and discrepancies between codebooks/questionnaires and data sets such as the lack of labels for missing data, inadequate labels for questions etc. Information about exceptions in regard to omitting some questions in some countries within

the wave was not always provided. In the process of gathering the data and documentation for harmonization, the amount of observed inaccuracies was growing and the idea of the need of assessing the quality of data was contemplated among the members of the research team. The initial phases of our work in the data harmonization project consisted of data mining and extraction of information from the data documentation, and often double-checking the information between the various sources. Such work makes quality issues obvious, and very soon the eye of the researcher becomes sensitive even to the tiniest detail. While this might be a source of frustration, it is very important such quality issues to be recorded, analyzed, and when possible, fixed. This was the driving idea behind our work on quality assessment.

These dimensions do not create an exhaustive list of possible quality indicators from user's perspective. Wang & Strong (1996), based on the result of their research claim that for most data users the parameter which score very high is representational consistency (data are continuously presented in same format, consistently represented, consistently formatted, data are compatible with previous data), traceability (survey data is well-documented and verifiable), and reputation (reputation of the both data source and data). There are also other categories, but our main point in naming some of them is to show that most users are perceiving data and survey quality in far more detailed categories than just accuracy. It is obvious that even the surveys with perfect reputation will not score high on each of the dimensions, but there is a certain "golden standard" that each investigator should try to reach. The problem with standards, thought, is that researchers and organization might agree on a set of certain standards, but it does not guarantee that data delivered by them meets these requirements (Lynn, 2003).

Quality? What did we look at?

„It makes sense to ask questions about the quality of data gathered for a particular research project. Answering these questions, however, is less obvious" (Loosveldt, Carton, Billiet, 2004, p.65). Practically the problem appears because of the great range of the angles from which quality can be assessed. Our quality assessment work was based on close examination of potential human as well as systemic errors concerning the "the expected quality of the deliverables in terms of product characteristics" known as "product quality" (Lyberg & Stukel 2010, 230), and in particular from the perspective of the user, based on the survey metadata made available by the institutions that carried out these surveys. We primarily used "basic comparative documentation" (for definition see Mohler et al. 2010, 311)

such as codebooks as well as the so called “SPSS dictionaries,” documents generated from the actual data sets displaying the basic labels and features of the data. Where codebooks were not available, we resorted to questionnaires. In addition, we also examined not only the labels of the data, but the actual values that appear in the data itself, too. The quality assessment was carried by the means of a Quality Control Sheet (attached in a separate Excel file to this paper) – a tool more rudimentary than for instance a Quality Assurance or Quality Control Program (see Lyberg & Stukel, 2010). The concept we used was rather an import from the software industry (for example see “Product Globalization Services | Pactera” 2014).

In order to check whether there is the processing error in 89 survey waves of 22 survey projects, we have decided to choose the variables that we will check in details, which in a way represent the quality of the documentation/data quality of the survey for us. Due to the large amount of tedious checking we needed to divide this work in the research team into smaller segments and check each of the chosen variables in each of the survey used for harmonization by using the same template. As the result we have created the file which contained the information on the exact information about such variables as gender, age (year of birth, age in years), education (level of education and schooling years), trust in parliament and participation in demonstrations from a) codebook b) questionnaire c) SPSS dictionary d) data itself. Out of each of mentioned sources we have tried to extract the maximum available information about the variables meaning a) variable name b) question number c) exact question formulation d) variable label e) value labels f) exact values in the data. If there was a discrepancy between any of these above mentioned elements, we have recorded it and discussed whether this discrepancy is significant and whether it can lead to errors.

Taxonomy of errors

The important step of the quality assessment task is to find an appropriate framework that allows for recording and analyzing the quality issues, or what we call “errors” in the ways the documentation and the data work together. The simplest way to do this is to create taxonomy of potential errors, and then to operationalize it. This in itself is a ground-breaking task, as to our knowledge, no one ever did anything that approximates the nature of the work done towards completing our quality assessment task.

The categories of errors that we considered are as follows:

1. **Suspicious Value (SV)**

   Meaning: the value that appeared in data looks suspicious and was not explained in codebook or SPSS dictionary. Mostly refers to 'age' and 'schooling years', while there is no obligation to describe all values, nor there is an obligation to write minimum and maximum values, so it can lead to unexpected 130 years old respondents in data.

   Values for years of schooling that are not suspicious: 4-28 years (everything above and below is suspicious). Values for age that are not suspicious is 15 – 97. Everything above and below is suspicious.

2. **Variable Value Discrepancy (VVD)**

   Meaning: there is a different value label in the codebook and SPSS dictionary. Refers to one or couple values coded with an error.

3. **Contradictory Value Labeling (CVL)**

   Meaning: there is a completely different labeling in codebook and SPSS dictionary. Refers to the cases with reversed labels and to cases were all values in one variable are coded differently in two sources of documentation.

4. **Lack of Labels (LL)**

   Meaning: a) There is no information about how "Don't know", "Refused to Answer", "Missing Data" was coded, while there is no description in either codebook (CB) or SPSS dictionary. Ex. we see in data report, that there is a value -1, and we can assume it is some coding for missing value, but there is no information in CB or SPSS dictionary what exactly it is. b) There is absolutely no missing data coding (neither in CB nor in SPSS dictionary or data). Ex. 'nulls' appear in data and are not explained c) there is a number in variable values list that is not defined in codebook/SPSS dictionary or any other source.

5. **Misleading Values (MV)**

   Meaning: refers to cases when in the codebook and/or questionnaire there is not a complete information about what a user can expect to have in the data, like giving only age categories, although in data there is the information about the exact age.

6. **Misleading Variable Label (ML)**

   Meaning: meaning the differences between the question and the label are not clear and logical label. This is a minor mistake, could be called as '*niechlujstwo*'

7. **Insufficient Information (II)**

Meaning: the situation when the general codebook does not include important information, like country specific differences in asking/not asking/coding some question.

8. **Translation Issues**

Meaning: Not all documentation is translated into English. Ex.: English codebook, but Spanish data with data labels and values.

The categories mainly referred to the usability and the clarity of the variable and value labels, as well as to the property of the data being easily understood using the documentation available. At the same time, these categories discern between the nuances of the potential error.

Examples

In the tables below, there are examples for every error category. In the example 1 (Suspicious Value), in the data there appear such values for the age variable as 108 years old for example and it is not explained in the codebook. In the example 2 (Variable Value Discrepancy), one can see different wording of the value labels in the codebook compared to the SPSS dictionary. In example 3 (Contradictory Value Labeling), the individual values have different meanings according to the codebook, as opposed to the SPSS dictionary.

	Example 1	Example 2	Example 3
Unique ID	IVS_1_9_ X003[2]	ASB/1_q010	CNEP/3/ZA_L.Education
Error type	**Suspicious Value**	**Variable Value Discrepancy**	**Contradictory Value Labeling**
Question	WVS: This means you are __ __ years old. IVS: Master question in EVS 1981 (ZA4438, age): Age in year of	I'm going to name a number of institutions. For each one, please, tell me how much trust you have in them. Is it a great deal of trust, quite a lot of trust, not very much trust, or none at all? Parliament	What is the highest level of education you have completed? [Do not read options. Code from answer.]

[2] Here and later in the paper the acronyms for survey names are used. For the exact name of the survey please see Appendix 1

	respondent Master question in EVS 1990 (ZA4460, Q719b): [Can you tell me your year of birth, please] This means you are ... years old		
Label in codebook	Age	None	None
Label in SPSS dictionary	AGE	How much trust do you have in Parliament?	Q67. HIGHEST LEVEL OF EDUCATION
Values from codebook	.-5 other missing -4 question not asked -3 not applicable -2 no answer -1 don't know	1 = A great deal of trust 2 = Quite a lot of trust 3 = Not very much trust 4 = None at all	No formal schooling (cannot read or write) 0 Some primary schooling 1 Primary school completed 2 Some secondary school / high school 3 Secondary school completed (vocational or commercial school) 4 Secondary school completed / high school (general educational track) 5 Incomplete university education / Other post-secondary qualifications (e.g. diploma / degree from a technikon or college) 6 University completed 7

			Post-graduate 8 Don’t know [Do not read] 99
Values from SPSS dictionary	.-5|Missing; Unknown -4|Not asked in survey -3|Not applicable -2|No answer -1|Don't know	0|Not sure 1|None at all 2|Not very much trust 3|Quite a lot of trust 4|A great deal of trust 98|Don't know 99|No Answer	1.00|No formal schooling (cannot read or write) 2.00|Some primary schooling 3.00|Primary school completed 4.00|Some secondary school / high school 5.00|Secondary school completed (vocational or commercial school) 6.00|*Secondary school completed / high school (general education 7.00|Incomplete university education /Other post-secondary qualif 8.00|University completed 9.00|Post-graduate 10.00|Don't know [Do not read]
Data	.-5 -4 -3 -2 -1 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 **98 99 100 101 103 108**	1 2 3 4 98 99	1 2 3 4 5 6 7 8 9

Target	Schooling_years	Tr_parli	Edu

Table 1. Examples for Suspicious Value, Variable Value Discrepancy and Contradictory Value Labeling.

In example 4 (Lack of Labels), we see that in the data, aside from the labeled values (such as 1=yes and 5=no) , there also appears value 3, undescribed in any other source of documentation. Without the label for value 3 we do not know what it means. In example 5 (Misleading Values), we see a different logic of the coding in the SPSS dictionary and the values in the data. In example 6 (Misleading Variable Label), we have a question concerning trust in parliament, which in the SPSS dictionary and codebook is coded as "confidence".

	Example 4	Example 5	Example 6
Unique ID	PPE7N_NL_V345	CB/2012_RESPDOB	LB/2010_ P20ST.A
Error type	**Lack of Labels**	**Misleading Values**	**Misleading Variable Label**
Question	335 Have you ever taken part in a demonstration or protest march?	In what year were you born?	Please look at this card and tell me, how much trust do you have in each of the following groups/institutions. Would you say you have a lot (1), some (2), a little (3) or no trust (4) National Congress/Parliament
Label in codebook	PROTEST	None	None
Label in SPSS dictionary	PROTEST PARTIC IN DE	Respondent - year of birth	Confidence in National Congress / Parliament
Values from codebook	(1) YES (5) NO (8) DON'T KNOW (9) NA	TABLE A1 _ _ _	0 No answer 1 A lot of confidence 2 Some confidence 3Little confidence

			4 No confidence at all 8 Don´t know
Values from SPSS dictionary	None	-9|Break-off -7|Legal skip -5|No such member in the household -3|Interviewer error -2|Refuse to answer -1|Don't know 1|1900-1909 2|1910-1919 3|1920-1929 4|1930-1939 5|1940-1949 6|1950-1959 7|1960-1969 8|1970-1979 9|1980-1989 10|1990-1999 11|2000-2009 12|2010-present 9000|Since 2000	.-4|Not asked -3|Not applicable -2|No answer/Refused -1|Don´t know 1|A lot of confidence 2|Some confidence 3|Little confidence 4|No confidence at all
Data	0 1 **3** 5 8	-3 1911 1916 1918 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982	-2 -1 1 2 3 4

		1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994	
Target	PR_demonst	Birth_year	PR_demonst

Table 2. Examples for Lack of Labels for Missing Data, Misleading Values and Misleading Variable Label.

The example 7 (Insufficient Information) refers to a variable for which we have only the name of the source variable in the SPSS dictionary and values in data, but we do not have any other information about the variable (we neither have value labels in codebook nor in SPSS dictionary, which makes it impossible to interpret whether 1 is male and 2 is female or vice versa). Finally, in the example 8 (Translation Issues) includes the situation when the documentation which claims to be in English does not include all necessary information translated and in SPSS dictionary the only label given to variable is in Spanish.

	Example 7	Example 8
Unique ID	LITS/2_respondentgender	AMB/1_ b13
Error type	**Insufficient Information**	**Translation Issues**
Question	None	To what extent do you trust the National Congress?
Label in codebook	None	None
Label in SPSS dictionary	None	¿Hasta qué punto tiene confianza usted en el Congreso Nacional?
Values from codebook	None	1 NOT AT ALL 2 3 4 5 6 7 A LOT (8) DN

Values from SPSS dictionary	None	1|1 Nada 2|2 3|3 4|4 5|5 6|6 7|7 Mucho 888888|DK 988888|NR 999999|N/A
Data	-1 1 2	1 2 3 4 5 6 7 888888 988888 999999
Target variable	Gender	Tr_parli

Table 3. Examples for Insufficient Information and Other type of error.

Analysis of the outcomes

In the following part we will present the results of the analysis of the data and documentation comparison done in all 22 surveys (89 waves) for 7 chosen variables, among which are basic background variables (gender, age, year of birth, education, schooling years), attitudinal variable (trust in parliament) and factual variable (participation in demonstration). Firstly, we will present the general characteristics of the sample and distribution of errors across the whole analyzed sample. Secondly, we will have a closer look on how variables vary as to the type of error, which we can meet in them. Thirdly, we will come into the comparison of survey projects, showing how they differ in terms of the quality of data and documentation. Here we will have a closer look on the quality of each wave of each international survey project in the study, and then compare the whole survey projects, accounting for the number of waves in the analysis.

Sample and Distribution of Errors

For the task of assessment of the resemblance between documentation and data, we have worked on a sample of variables which we had matched with target variables specified in the Data Harmonization Project during the previous steps of work in the project. In

particular, we focused on the following seven target variables: age (Age), year of birth (Birth_year), gender (Gender) education level (Edu), years of education (Schooling_years), participation in demonstrations (Pr_demonst) and trust in parliament (Tr_parli). We were working with all available in English sources of documentation for the projects, extracting from them the information on corresponding variables in each of the survey-wave.

Our main unit of the analysis was the variable in the survey wave. While we have 89 projects and we are checking seven target variables in each, the expected amount of total variables for the analysis has been 693. However, the total number of items of observation in the project is larger than that and reaches 814 cases. The increased sample of variables is due to two things: a) there were situations in which more than one source variable matched the target variable[3] b) some surveys had only country specific information available[4]. Thus, all matching source variables, together with country specific variables if other was not available, were captured in the sample. Out of 814 items that we have checked, 687 were matched to the target variables and in 127 cases we did not find any corresponding source variable to our target variable. The complete set of all seven target variables matched with at least one source variable occurred only in 14 waves in 5 survey projects, namely ESS/1-6, ASB/3, CB/2012, EVS/3-4 and WVS/2-5. Some target variables are more frequently matched than the others; while for instance all but one[5] survey projects contained a variable on gender, only 12[6] contained a “birth year" variable.

In the table below we present the general outlook of the sample, were we show the exact amount of the items of our observation for quality check. The variables matched, meaning the amount of source variables corresponding to our target variables in 89 waves of 22 survey projects is divided into two subcategories – items in which we found discrepancies between data and documentation and items without an error found. The number of unmatched variables shows in how many cases there were no source variable corresponding to target variable.

Sample	Count

[3] For example 2 variables on education in one survey-wave.
[4] There were three exceptions, namely CNEP, PPE7N, VPCPCE, in which due to the lack of the merged documentation for the given wave we have looked at the documentation on the country level . In one case (ASES) the information about education was only country specific.
[5] LITS/1
[6] ASB/3, CB/2009-2012, CDCEE/1-2, CNEP/3/MX-TW-UY, ESS/1-6, EVS/1-4, ISJP/1-2, ISSP/2010-2011, PA2, PA8NS, PPE7N_NL, WVS/1-5

Variables matched	Items of observation, no error	490
	Items of observation with error	197
Unmatched variables (“blanks”)		127
	Total N	814

Table 4. General outlook of the sample.

Throughout the process we have identified a total of 250 errors in 197 variables, out of the total 687 variables matched on the level of the survey-wave. Some variables contained more than one error. For example, 49 variables contained two kinds of errors, while two variables contained three errors (see Table 5), which are overviewed below (see Table 6).

Distribution of errors		Frequency
Number of errors	1	146
	2	49
	3	2
Total number of variable containing errors		**197**
Total number of errors		**250**

Table 5. Distribution of errors

Unique ID	LB/1997_s2	LITS/2_respondentage
Target Variable	Age	Age
Number of errors	3	3
Question	How old are you? (Write the number of years that respondent is)	None
Label in codebook	None	None
Label in SPSS dictionary	AGE	None
Values from codebook	[1] 18-24 years old [2] 25-34 years old [3] 35-44 years old [4] 45-54 years old [5] 55-64 years old [6] 65 and older	None

	[0] No answer	
Values from SPSS dictionary	None	None
Data	**null** 0 2 7 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 **97 99**	**-1 2 7 8 9 10 13 14** 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 96 **97 98 99**
Errors	Suspicious value, Lack of Labels, Misleading Values	Suspicious value, Lack of Labels , and Insufficient Information

Table 6. Examples of the two most problematic matched variables, LB/1997_s2 and LITS/2_respondentage.

Variation across target variables

The errors that we have found were not distributed equally across all of the target variables, as seen in Table 7 and Figure 1. There were some variables that often included more than one error, for example 58 of the age variables contained a total number of 78 errors, compared to 44 of the education variables containing a total of 46 errors.

		Number of Errors			**Total number of variables with errors**	**Total number of errors**
		1	**2**	**3**		
Target Variable	Age	40	16	2	58	78
	Birth_year	8	0	0	8	8
	Edu	42	2	0	44	46
	Gender	10	1	0	11	12
	PR_demonst	22	12	0	34	36
	Schooling_yea	10	12	0	22	44

	rs					
	Tr_parli	14	6	0	20	26
Total		146	49	2	197	250

Table 7. Errors by target variables.

It is interesting that the greatest amount of errors we have found in the background variables, which are essential in different kinds of analysis and are often used in research. The inconsistencies in different sources of documentation and data, when we look at questions on age, schooling years and education of respondent can be a cause to at worth wrong empirical judgments or at best hours of tedious work in getting to understanding what exactly the respondent was asked and how these variables can be used. Moreover, 'age' was the only variable were we have often found more than one type of inconsistency.

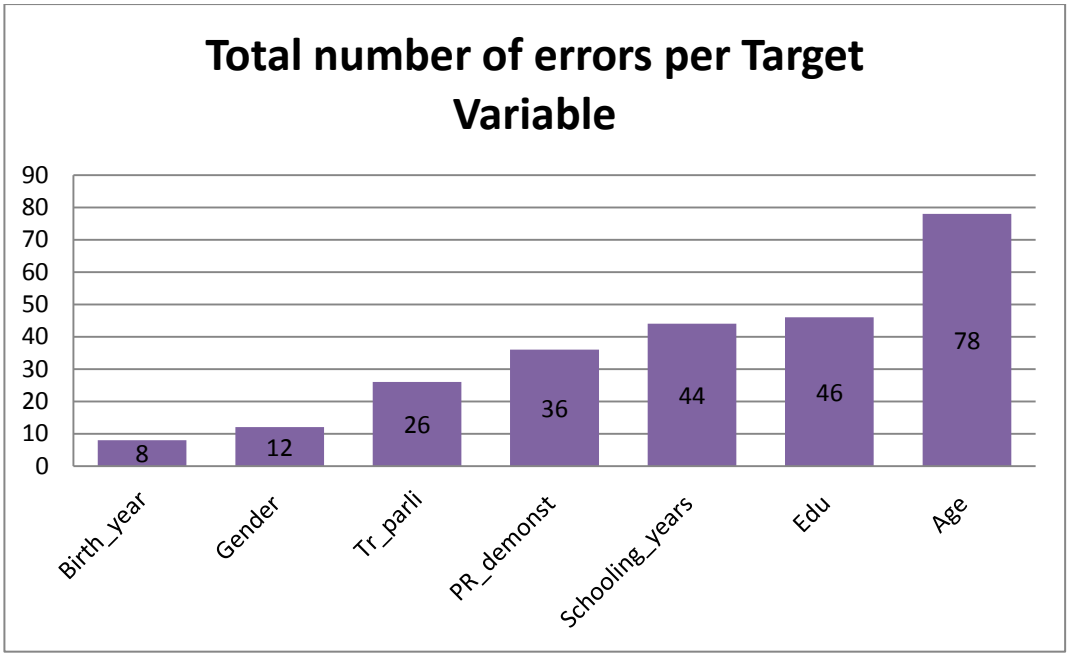

Figure 1. Total number of errors by target variables.

Looking closer at the target variables, it is important to highlight what is the exact distribution of error types in them (see Table 8 and Figure 2). As for the error types, we have used here broader categories showing whether the error occurred in value (which includes suspicious values, misleading values and variable value discrepancy), in label (meaning contradictory value labeling, lack of labels and misleading labels) and generally in information (which is insufficient information and translation issues). For example, age variables tended to have mostly errors concerning the values (suspicious or misleading), as they usually have a very high range of values. Some of the value-related errors in age

variables are due to the fact that this range is not specified, so for example, encountering a value of "123" raises suspicion. Similar value-related problems were encountered with variables matched with the target variable regarding schooling years.

On the other hand, a large number of the variables matched with the target variable "education level" exhibited a "lack of labels" error, often with missing data not being labeled properly in documentation.

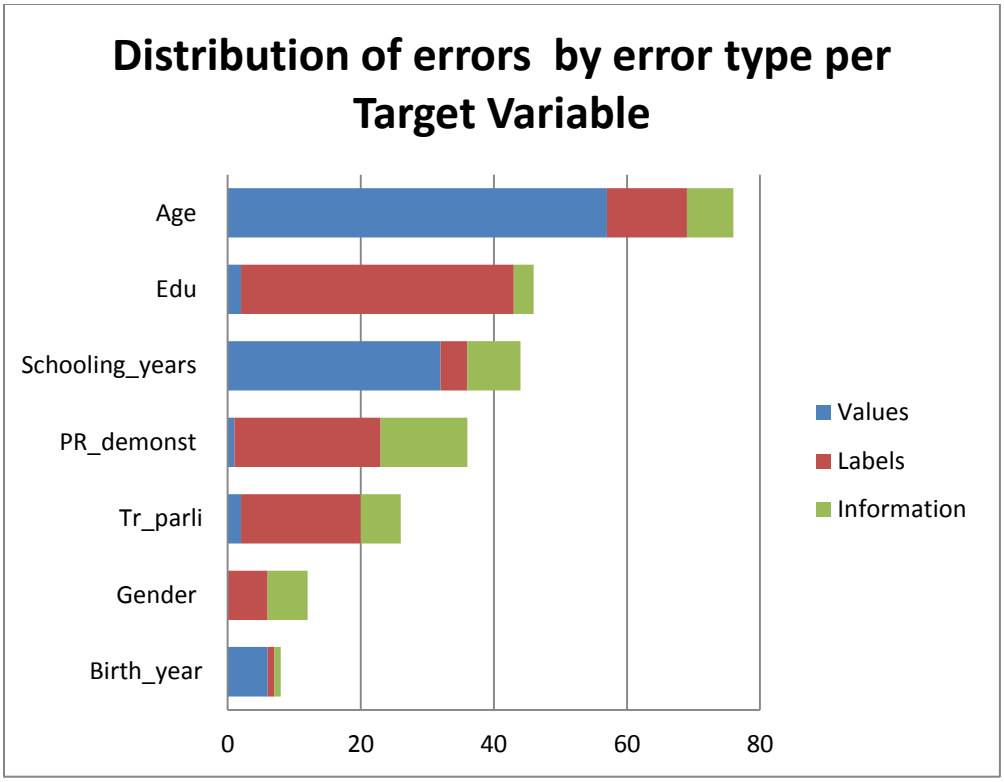

Figure 2. Distribution of error types by target variables.

Variable	Values	Labels	Information
Birth_year	6	1	1
Gender	0	6	6
Tr_parli	2	18	6
PR_demonst	1	22	13
Schooling_years	32	4	8
Edu	2	41	3
Age	57	12	7

Table 8. Distribution of error types by target variables.

Variation across error types

Similarly, the total number of identified errors was spread out very unevenly throughout the categories we worked with (see Figure 3). More than a half of all errors (135 out of the total of 250) belong to the "lack of labels" and "suspicious value" category. We found a significant amount of misleading values and contradictory value labeling, which may cause major misunderstanding in using the data. There were also many issues with translation for some variables. It is interesting that the inconsistencies of translation to English were only from Spanish.

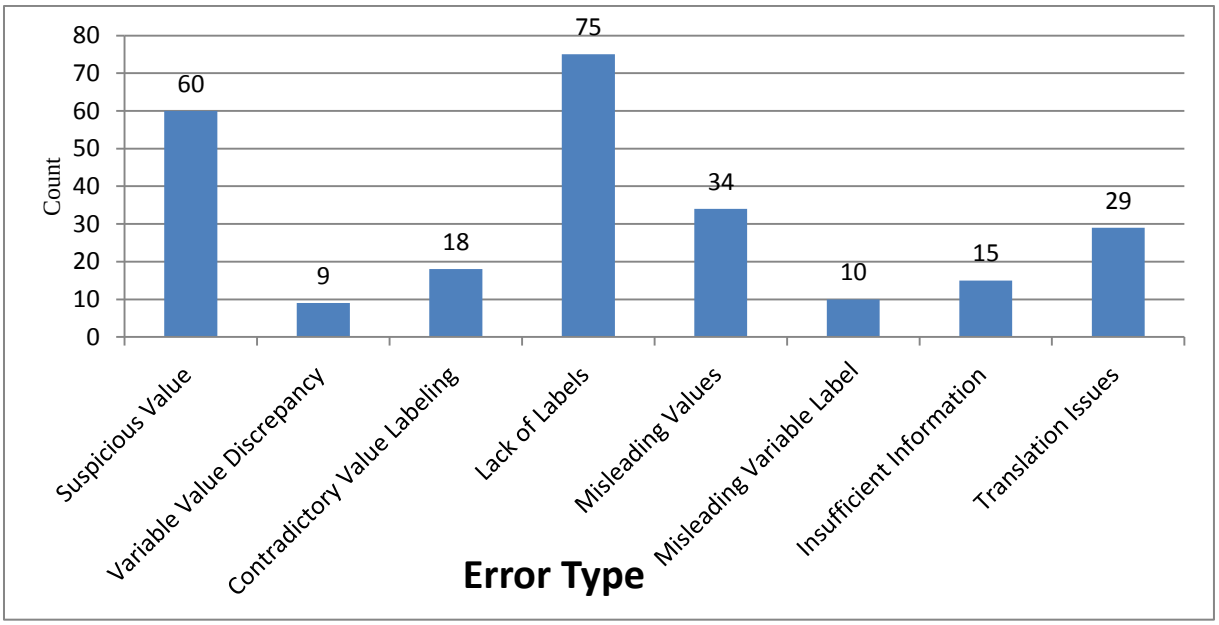

Figure 3. Distribution of errors by error type.

The above mentioned errors can be grouped in three major categories: issues dealing with values, problems with labels and insufficient information together with translation issues (see Illustration 1). On the illustration 1 there are errors grouped by categories, together with the frequency of occurrence of the error type. From this figure we can see that generally we have found similar number of inconsistencies with labeling and with values themselves. The arrows on the graph show error type which occurred most often ("lack of labels") and two error types which were the most rare ("variable value discrepancy" and "insufficient information").

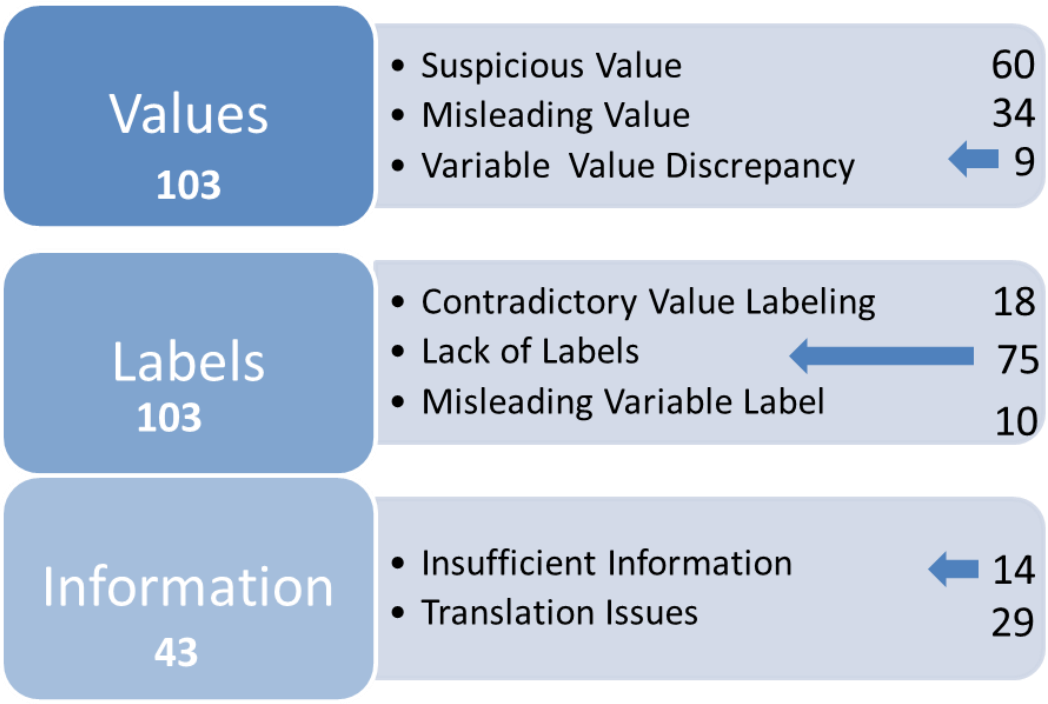

Illustration 1. Distribution of error types by error categories.

To get a better perspective, Figure 4 shows us how error types varied across different survey waves and Figure 5 depicts the differences on the level of survey projects. It is indicative, that many surveys that contain more than one error also contain more than one type of error, however in the example of ASES and ARB/2 we can see a high number of errors, all of them belonging to the same category, which speaks of a general pattern of these surveys (for a smaller number of errors, this is also seen in the cases of ESS – see Figure 11). The distribution of errors is presented here on the level of items of observation (meaning level of questions corresponding to our target variables), which shows a full picture of all analyzed units.

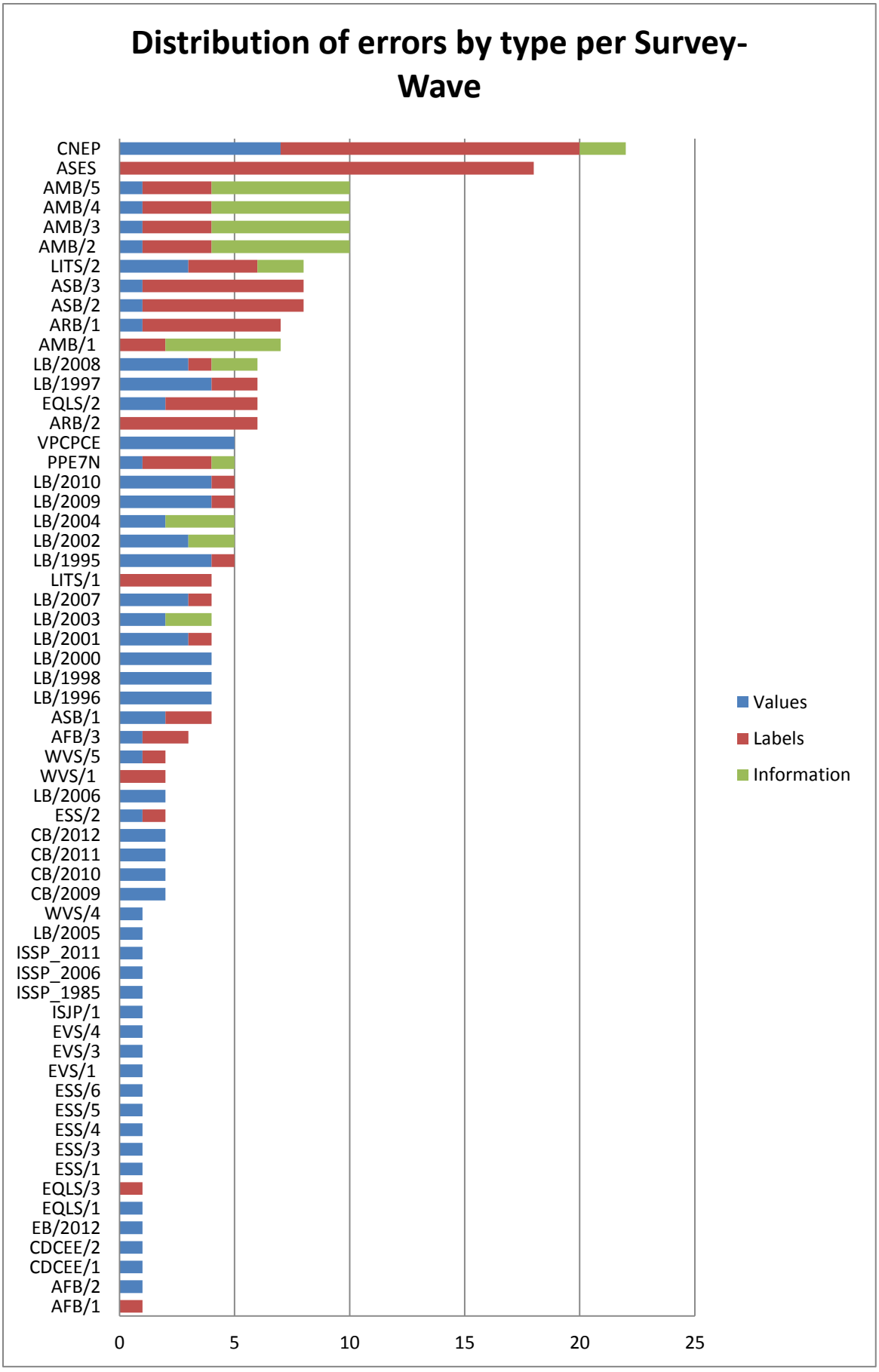

Figure 4. Distribution of error types by survey-wave.

While looking on the distribution of error types on the level of survey project (Figure 5) we see that the projects, which have the smallest number of errors in the sample, have usually the problems of the same kind, which is on the level of values (ISSP, EVS, EB, CDCEE, ISJP). However, the projects which have the biggest number of errors with labeling and values, usually are also the ones that have issues with translation and insufficient information (LB, AMB).

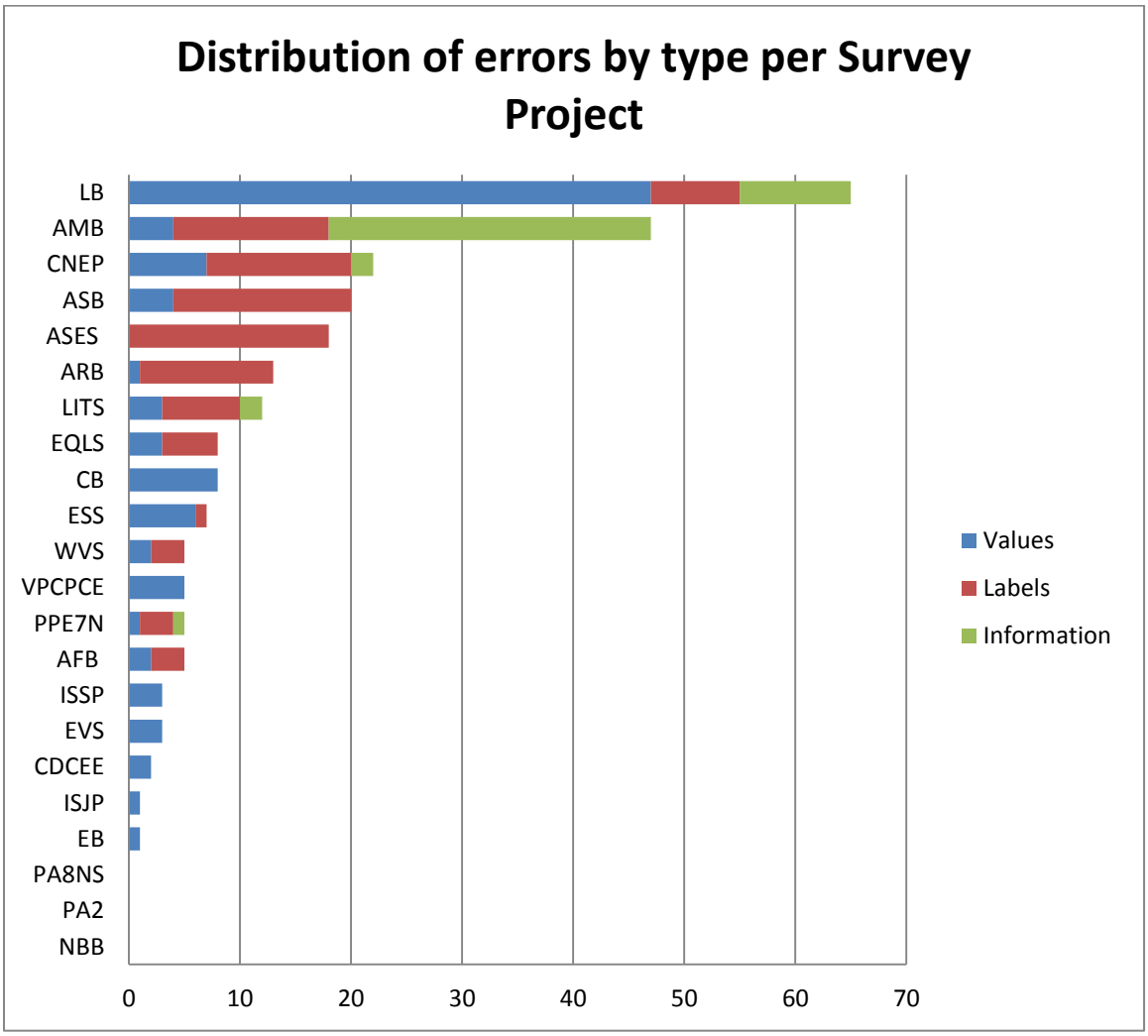

Figure 5. Distribution of error types by survey project.

Although it is important to pay attention to the diversity of surveys in terms of error types, the judgment on the quality of one survey compared to another based on this figure can be error prone, while the graph presented does not take into account the number of waves that were taken into account. To remind the peculiarities of the sample, it should be mentioned that although Latino Barometer, Americas Barometer and Comparative Nation Elections Project are on the top of the graph, they all have very different number of waves: LB has 15

waves in the sample, AMB has five and CNEP has one wave (but country specific questions). To account for these differences we have looked at the average amount of errors per wave in each project, the results see in Table 9. Looking from the perspective of the average number of error per wave we see that two leading positions take CNEP (22 errors in one wave) and ASES (18 errors per wave). The case of both of these survey projects is specific, while we have the country level items of observation in both of them, which makes those projects the outliers in terms of average number of errors per wave. Despite the fact that there were country specific questions in CNEP and ASES, the number of errors found in them still is alarming. Starting from Americas Barometer project the average number of errors shows a comparable picture. Interesting to mention, Latino Barometer survey project although having overall 65 errors, after dividing them per waves analyzed has an average of four errors per wave, which located it in the middle of the sample.

	Number of errors (sum of errors per wave)	**Number of Waves**	**Number of errors/waves**
CNEP*	22	1	22
ASES*	18	1	18
AMB/1-5	47	5	9
ARB/1-2	13	2	7
ASB/1-2	20	3	7
LITS/1-2	12	2	6
PPE7N	5	1	5
VPCPCE	5	1	5
LB/1995-2010	65	15	4
EQLS/1-3	8	3	3
CB/2009-2012	8	4	2
AFB/1-4	5	4	1
CDCEE/1-2	2	2	1
ESS/1-6	7	6	1
EVS/1-4	3	4	1

ISJP/1-2	1	2	1
WVS/1-5	5	5	1
EB	1	7	0
ISSP	3	13	0
NBB/1-6	0	6	0
PA2	0	1	0
PA8NS	0	1	0
Total	**248**	**89**	

Table 9. Average number of error per wave of survey project.

Data and Documentation Quality Index: accounting for number of items under analysis

Comparing surveys in terms of quality is a tempting and challenging task, while there are many peculiarities and specificities of every project that should be taken into account. The figures 3 and 4, as well as table 9 show the characteristics of the projects under analysis, although do not give us the comparable ranking of the quality of surveys. The main challenges for creating such ranking lay in the fact, that the available documentation for 22 survey projects have not been done with the similar standard and include different set of questions. Taking into the consideration the specificity of the ex-post harmonization, it seems to us that the most accurate comparison in terms of quality of documentation and data resemblance is the analysis of the average number of errors per source variable in the sample. In this case we take into account the variation of the items of observation and answer the question on how many errors were found in each wave with consideration of number of items of observation in the wave. For example, in ISSP/1998 20 source variables matched our seven target variables, so we have checked 20 items for inconsistencies between data and documentation and found zero errors. In contrast, in AMB/1 we have checked five items and found total seven errors in them (see Table 10). In Figure 5 and Figure 6 you can see the index of quality of survey waves and survey projects respectively. To obtain the coefficient for quality index on the wave level we have divided the total number of errors per the total number of items in a given wave. For the survey project level, we have summed the obtained coefficients of error per items in all waves of the projects and divided by the total number of waves of the project. Both indexes do not take into account the possible weight of the error type, although it is important to highlight that not all error types are alarming to the same extent.

Survey/Wave	**Number of items (source variables included)**	**Number of errors**	**Number of errors / Number of items**
AMB/2	6	10	1,67
AMB/3	6	10	1,67
AMB/4	6	10	1,67
AMB/5	6	10	1,67
AMB/1	5	7	1,40
ASB/2	6	8	1,33
LITS/2	6	8	1,33
ARB/1	6	7	1,17
ASB/3	7	8	1,14
ARB/2	6	6	1,00
LB/1997	6	6	1,00
LITS/1	4	4	1,00
EQLS/2	7	6	0,86
LB/2004	6	5	0,83
LB/2009	6	5	0,83
LB/2010	6	5	0,83
ASES*	24	18	0,75
LB/2008	8	6	0,75
LB/2000	6	4	0,67
LB/1995	8	5	0,63
LB/2002	8	5	0,63
LB/2003	8	5	0,63
AFB/3	5	3	0,60
LB/1996	7	4	0,57
LB/1998	7	4	0,57
LB/2001	7	4	0,57
LB/2007	7	4	0,57
CNEP *	42	22	0,52
ASB/1	8	4	0,50
CB/2009	6	2	0,33

CB/2010	6	2	0,33
CB/2011	6	2	0,33
LB/2006	6	2	0,33
CB/2012	7	2	0,29
AFB/1	4	1	0,25
EQLS/1	4	1	0,25
ESS/2	8	2	0,25
ISSP_2006	4	1	0,25
WVS/1	9	2	0,22
AFB/2	5	1	0,20
CDCEE/1	5	1	0,20
EB/2012	5	1	0,20
ISJP/1	5	1	0,20
ISSP_2011	5	1	0,20
WVS/5	11	2	0,18
PPE7N	28	5	0,18
ISSP_1985	6	1	0,17
LB/2005	6	1	0,17
VPCPCE	31	5	0,16
CDCEE/2	7	1	0,14
ESS/1	7	1	0,14
EQLS/3	8	1	0,13
ESS/3	8	1	0,13
ESS/4	8	1	0,13
ESS/5	8	1	0,13
ESS/6	8	1	0,13
EVS/1	9	1	0,11
EVS/3	11	1	0,09
EVS/4	11	1	0,09
WVS/4	11	1	0,09

Table 10. Total number of items and errors per survey-wave with the average amount of error per analyzed item. Waves for which average error is zero: AFB/4, EB/1983-2010, EVS/2, ISJP/2, NBB/1-6, PA2, PA8NS, WVS/2-3, ISSP/1989-2004 & 2007-2010.

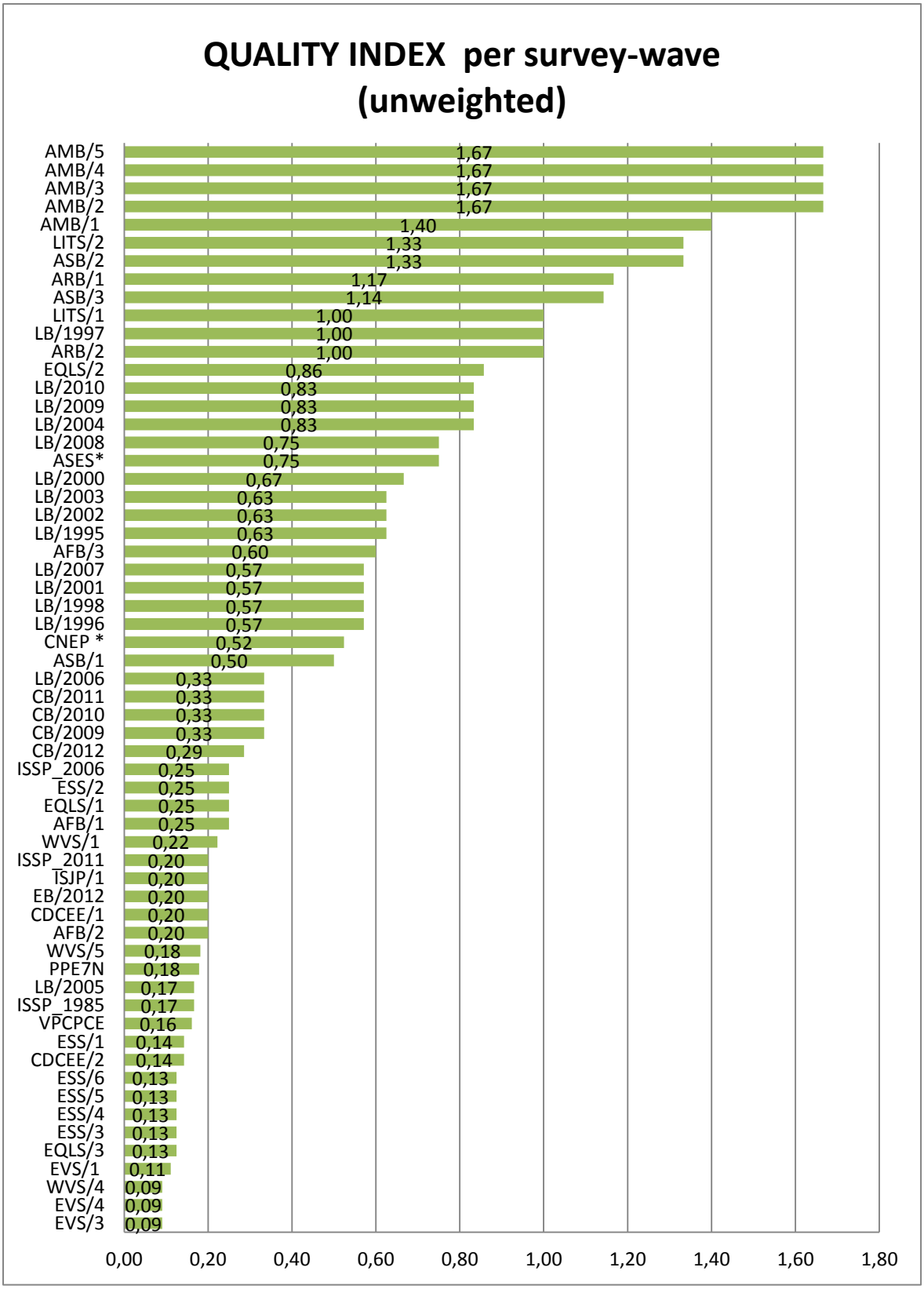

Figure 6. Unweighted quality index per survey wave with account of total number of items. Quality index equals 0 for the following survey-waves : AFB/4, EB/1983-2010, EVS/2, ISJP/2, NBB/1-6, PA2, PA8NS, WVS/2-3, ISSP/1989-2004 & 2007-2010

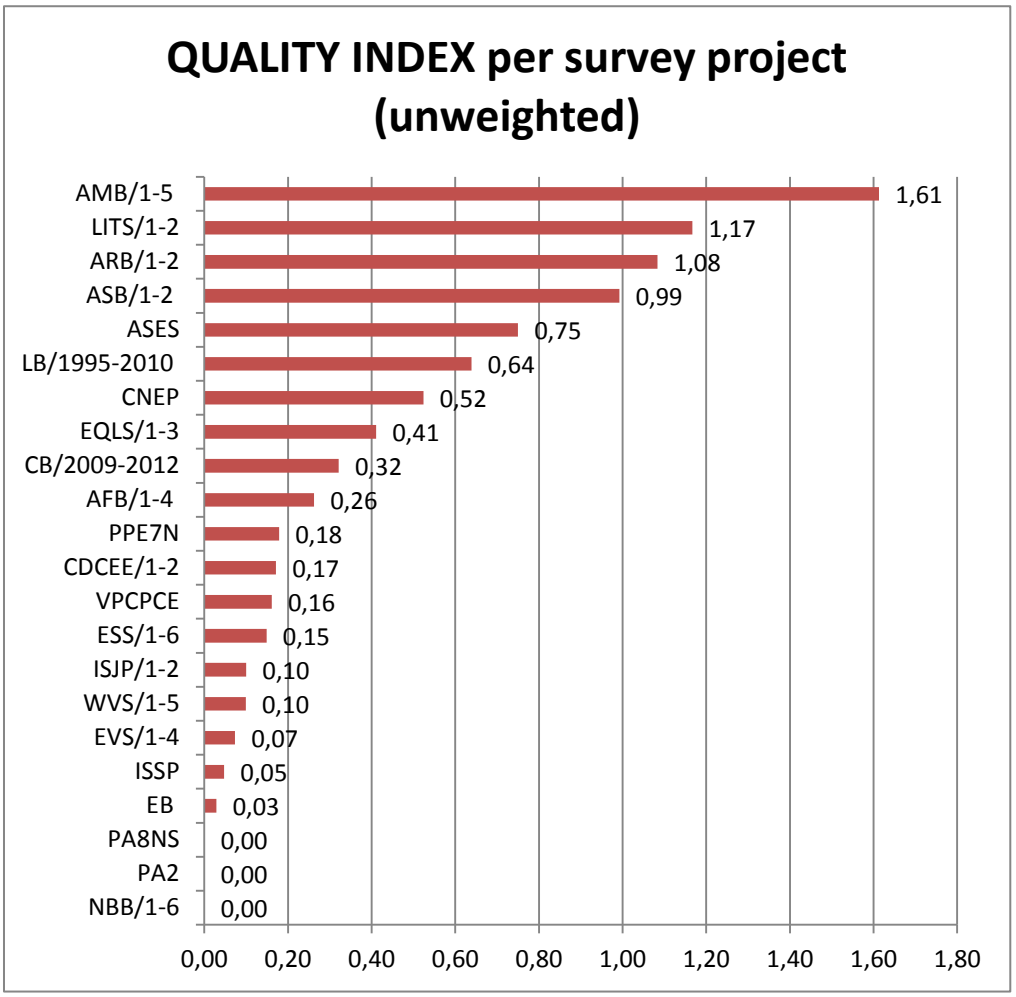

Figure 7. Unweighted quality index per survey project with account of total number of items and waves.

Both on the level of individual waves and on the level of survey project itslef, the survey project in which the greatest amoung of discrepencies between data and documentation was found is Americas Barmometer. It is followed in terms of quality by Life in Transition Survey, Arab Barometer and Asian Barometer. The three survey projects in which we have found no errors are also the oldest projects in the sample, which are Political Action - An Eight Nation Study, Political Action II as well as New Baltic Barometer. It should be highlighted that survey projects vary in terms of the quality of waves, and unfortunately there is no clear tendency to a better quality. For example the first wave of Asian Barometer contained fewer errors than the second and the third ones. Overall, out of 22 survey projects there are only seven projects which on average have more than 0,5 mistake per every wave controlling for the number of items analyzed in the seven target variables checked. Thus,

although overall situation is not dramatic, the processing errors do happen in almost every survey wave and unawareness of them may lead to dramatic outcomes.

Summary

To conclude, the critical assessment of quality and the transparency of the documentation of all stages of data gathering are the fundamental issues for the researchers that want to use the data. In the context of data harmonization, the processing errors check is crucial in order to get the awareness of the quality of data with which we are working. Thus, we see the processing quality assessment of the survey projects as the inherent part of the harmonization process. While analyzing the sample of 22 surveys 89 waves and seven chosen target variables (age, year of birth, gender, education, schooling years, trust in parliament and participation in demonstration) we have encountered a number of discrepancies between data and documentation, which were later grouped into eight error types. The most frequent type of error that occurred in our sample is the lack of labels, especially for missing data. The amount of misleading and suspicious values, which are not explained or contradictory in different sources of documentation, was also relatively high (total of 103). The biggest amount of various errors we have found in the variable 'age', which is one of the basic background variables often used in the analysis.

Generally, survey projects vary in terms of their quality in terms of resemblance of data and documentation. There is also a noticeable diversification of composition of types of processing errors in different survey waves. Based on our data gathered on the processing errors in each analyzed survey-wave, we have created the data and documentation quality index, which shows the average amount of mistakes per unit of observation. The index shows us that there are projects in which the quality of waves did not increase across time, or even declined. On the level of survey project itself, we have looked at the average amount of errors in the wave accounting for the number of items under the analysis. Based on this Data and Documentation quality index we see that Americas Barometer has the greatest need for improvement, followed by Life in Transition Survey, Arab Barometer and Asian Barometer. Which is interesting, the oldest surveys in the sample, namely Political Action - An Eight Nation Study, Political Action II, were among the ones with no discrepancies between data and documentation. The quality indexes were created without the account for the different weight of the error type, which will be implemented in the further analysis.

References

Anderson R., Judith K., and Martin F. (1979) *Total Survey Error: Applications to Improve Health Surveys*. San Francisco, CA: Jossey-Bass.

Biemer P.P., Lyberg L.E. (2003) *Introduction to Survey Quality.* Wiley-Interscience: A John Wiley & Sons Publication

Biemer, P. P., Lyberg, L. E. (2003), *Introduction to Survey Quality*, Hoboken, NJ: John Wiley & Sons

Deming, E. (1944) On Errors in Surveys, *American Sociological Review* 9:359–69.

Edward Crenshaw et al. (2014, May) *Roundtable Discussion: Past, Present and Future of the Interdisciplinary Empirical Study of Political Behavior,* OSU, United States.

Groves R. M., Lyberg L. (2010) Total survey error: past, present, and future. *Public Opinion Quarterly,* Vol. 74, No. 5, 2010, pp. 849 – 879

Groves R., Floyd F., Mick C., Eleanor S., and Roger T. (2004) *Survey Methodology*. New York: Wiley.

Groves, R. M. (1989), *Survey Errors and Survey Costs*, Source John Wiley & Sons, Inc.

Leeuw E. D. , Hox J.J., Dillman D.A. (2008) *International Handbook of Survey Methodology.* Taylor&Francis.

Loosveldt G., Carton A., Billiet J. (2004) Assessment of survey data quality: a pragmatic approach focused on interviewer tasks. *International Journal of Market Research* Vol. 46 Quarter 1, pp.65-82.

Lyberg, Lars E., and Diana Maria Stukel. 2010. "Quality Assurance and Quality Control in Cross-National Comparative Studies." In *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, by Janet A. Harkness, Michael Braun, Brad Edwards, Timothy P. Johnson, Lars E. Lyberg, Peter Ph Mohler, Beth-Ellen Pennell, and Tom W. Smith, 1 edition. Hoboken, N.J: Wiley.

Lynn, P. (2003) *Developing quality standards for cross-national survey research: five approaches,* International Journal of Social Research Methodology, 6(4): 323-336

Lynn, Peter. 2003. "Developing Quality Standards for Cross-National Survey Research: Five Approaches." *International Journal of Social Research Methodology* 6 (4): 323–36. doi:10.1080/13645570210132848.

Mohler, Peter Ph, Sue Ellen Hansen, Beth-Ellen Pennell, Wendy Thomas, Joachim Wackerow, and Frost Hubbard. 2010. “A Survey Process Quality Perspective on Documentation.” In *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, by Janet A. Harkness, Michael Braun, Brad Edwards, Timothy P. Johnson, Lars E. Lyberg, Peter Ph Mohler, Beth-Ellen Pennell, and Tom W. Smith, 1 edition. Hoboken, N.J: Wiley.

Morris H., Hurwitz W., and Madow W. (1953) *Sample Survey Methods and Theory* New York: Wiley.

“Product Globalization Services | Pactera.” 2014. Accessed June 16. http://www.pactera.com/outsourcing/product-globalization-services/.

Słomczyński K. (2014, May) *Garbage in Garbage out vs Recycling,* speech presented at OSU, United States.

Smith, T. (2009). *Refining the Total Survey Error Perspective*. Memo: ORC/University of Chicago.

Sztabiński F. (2011) *Ocena jakości danych w badaniach surveyowcyh*. Warszawa: Wydawnictwo IfiS PAN.

Wand Y., Wang R.Y. (1996) *Anchoring Data Quality Dimensions in Ontological Foundations*, November 1996/Vol. 39, No. 11 COMMUNICATIONS OF THE ACM

Wang R.Y., Strong D.M. (1996) *Beyond Accuracy: What Data Quality Means to Data Consumers*, Journal of Management Information Systems, Vol. 12, No. 4 (Spring, 1996), pp. 5-33

Weber M. (1949) *The Methodology of the Social Sciences,* The Free Press of Glencoe, Illinois.

Weisberg, H. (2005) *The Total Survey Error Approach*. Chicago, IL: University of Chicago Press.

Appendix 1. Survey project name and corresponding acronym

acronym	survey name	# of waves	Data files
AFB	Afrobarometer	4	AFB/1, AFB/2, AFB/3, AFB/4
AMB	Americas Barometer	5	AMB/1-5
ARB	Arab Barometer	2	ARB/1, ARB/2
ASB	Asian Barometer	3	ASB/1, ASB/2, ASB/3
ASES	Asia Europe Survey	1	ASES
CB	Caucasus Barometer	4	CB/2009, CB/2010, CB/2011, CB/2012
CDCEE	Consolidation of Democracy in Central and Eastern Europe	2	CDCEE/1-2
CNEP	Comparative National Elections Project	1	CNEP/3/ES, CNEP/3/HU, CNEP/3/MX, CNEP/3/MZ, CNEP/3/PT, CNEP/3/TW, CNEP/3/UY, CNEP/3/ZA
EB	Eurobarometer	7	EB/1983, EB/1984, EB/1989, EB/2000, EB/2004, EB/2010, EB/2012
EQLS	European Quality of Life Survey	3	EQLS/1-3
ESS	European Social Survey	6	ESS/1-5, ESS/6
EVS	European Values Study	4	IVS/1-9 (combines dataset for EVS and WVS)
ISJP	International Social Justice Project	2	ISJP/1-2
ISSP	International Social Survey Programme	13	ISSP/1985, ISSP/1989, ISSP/1990, ISSP/1991, ISSP/1996, ISSP/1998, ISSP/2004, ISSP/2006, ISSP/2007, ISSP/2008, ISSP/2009, ISSP/2010, ISSP/2011
LB	Latinobarometro	15	LB/1995, LB/1996, LB/1997, LB/1998, LB/2000, LB/2001, LB/2002, LB/2003, LB/2004, LB/2005, LB/2006, LB/2007, LB/2008, LB/2009, LB/2010
LITS	Life in Transition	2	LITS/1, LITS/2

	Survey		
NBB	New Baltic Barometer	6	NBB/1-6
PA2	Political Action II	1	PA2
PA8NS	Political Action - An Eight Nation Study	1	PA8NS
PPE7N	Political Participation and Equality in Seven Nations	1	PPE7N_AT, PPE7N_IN, PPE7N_JP, PPE7N_NG, PPE7N_NL, PPE7N_US, PPE7N_YU
VPCPCE	Values and Political Change in Postcommunist Europe	1	VPCPCE_CZ, VPCPCE_HU, VPCPCE_RU, VPCPCE_SK, VPCPCE_UA
WVS	World Values Survey	5	IVS/1-9 (combines dataset for EVS and WVS)