FP7-212214

Title	**Database specification for classifications and conversion information (D9.1a)**
Work Package	WP9
Authors	
Dissemination Level	PU (Public)

Summary/abstract

This document summarises the harmonisation materials used by Task 9.2 and can be used as short guidance for reading the distributed files (available from: http://www.cessda.org/project/deliverables.html) that form the Deliverable 9.1.

Content:

List of distributed files:

1. **D9.1 Collection of Harmonisation Materials (WP9.2).mdb**

2. **D9.1 ISCED-97.xls**

3. **D9.1 Income Data.xls**

4. **D9.1 Private Household Harmonisation.doc**

5. **D9.1 Descriptive Results.doc**

1 Introduction

Aim: This document summarizes the harmonisation materials (incl. harmonisation routines) created by WP9 (see: http://www.cessda.org/project/deliverables.html) for development of use cases and testing purposes.

Background: The goals of Task 9.2 comprised the development of the proof of concept for a database (DB) and a prototype software application to support survey data harmonisation within the future CESSDA portal. In addition to consider consultations with (external) experts in the field of comparative social sciences (realized by T9.1 and T9.3 until Month 5), a first step of Task 9.2 was to build up its own use cases by creating harmonized variables (resulted from conversions[1]) across different data sources. This was continuously undertaken until Month 18.

Testing materials: Resources involved in this work were stored and in turn the analysis of working steps resulted in a minimum of functional requirements for the desktop prototype application named CHARMCATS[2]. Therefore, these materials will be the primary option for testing purposes of this application in the future. In the second section of this document an overview on these materials stored so far is given. For structuring in a comprehensive way all the documentation used in conversions and the working process as such as well, Task 9.2 created a preliminary internal Access database (DB file name: D91.mdb). It also helped as a tool for exchanging the materials between WP9 members. Consequently, this file constitutes the core for this deliverable. It is recommended to read first Section 3 of this document before opening the Access file for having brief definitions of the data forms at had. Since this DB (first versions: Month 6) focused on the main elements required in the harmonisation work it was not designed to contain all the documentation and metadata to be included in the future online platforms on harmonization routines and question databank[3]. Thus, besides using the simple Access DB several additional materials were created for documenting the specific coding or detailed variable information (mainly in Word-spreadsheets or Excel- table format). Two of these files are outlined in Section 4 and are also included in this *Deliverable* as they are referenced in the Access DB. Based on the workflow analysis and the conceptual database of the prototype application, harmonisation materials were created in the latter phase of the project (between months 15-18), with the specific scope to test the prototype and as a result were not included in the Access DB, but summarized in a special document (see section 5)

Moreover, creating harmonized variables implied to scrutinize for their face and criterion validity. Merely descriptive and correlation analyses were used to support this and were collected in the attached document "Descriptive Results". Section 6 shortly summarizes how to read these findings.

Finally, brief conclusions were drawn and the future use of harmonisation materials within Task 9.2 is sketched in the last section.

2 Content of the Access DB: harmonisation work within Task 9.2

After internal consultations, Task 9.2 decided to focus on harmonizing measurements that meet the criteria of being based on available (archived within CESSDA) representative data of European countries and that could be applicable in interdisciplinary fields of comparative social research.

---

[1] Conversion routine = all the work involved in writing a syntax of commands in a statistical software package for transforming variables across different (country/time samples) data sets in order to construct comparative variables with harmonized codes (target variables); See also definitions in Section 3.2.

[2] Cessda HARMonisation of CATegories and Scales

[3] Databases projected within WP9: 1) A first prototype database is currently implemented by CHARMCATS (The conceptual Data Model and minimum requirements for this application are sketched in "Functional Requirements of CHARMCATS", internal draft). 2) The conceptual development employed in CHARMCATS DB was considered by Task 9.4 in specifying the future database within the CESSDA portal (DB Name: CHARMCATS/CCCDB; see Deliverable 9.2). 3) The third database was planned to make survey questions across CESSDA archives accessible/comparable within the future portal (DB Name: QDB, developed by Task 9.5; Deliverable 9.3).

As a result, following Classifications/Scale/Indexes were chosen to work on:

1. International Standard Classification of Education (ISCED-97). This comparative classification was created by OECD (OECD, 1999) as a development of the former classification known as ISCED-76. Primarily created for providing official statistics on education, this new version is more and more used also in the social sciences (e.g., Schneider, 2008) to measure education at both micro (individual) and macro (country) levels. It is also employed in policy analysis as a central indicator on European educational achievement (see for example Atkinson et al., 2003, pp.: 128-135 or UNESCO, 2007).

Two variants of this classification were created within W9.2: 1) a "simplified" version comprising only the main ISCED-levels and 2) a detailed version, containing all levels with subcategories. See also Section 4.1 for details and limitations on coding. This variable was derived for a set of 17 Countries across the ESS3, ISSP05 and ISPP06 data[4].

2. Equivalized household income (new OECD scale). Measuring income is directly linked with comparative measures of poverty and inequality (e.g., Gordon and Townsend, 2000). Task 9.2 applied the so called OECD scales for transformations of household income (Haagenars, Vois, and Zaidi, 1994) because these are the most used in research investigations and policy reports. Based on ESS3 and ISSP06 data four different harmonized variables were constructed for net and gross income, using PPPs in EUR and dollars and with slightly different weighting calculations. These variables allow for comparisons of income across 33 countries for the reference year 2005-2006. In addition to the Access DB, see also Section 4.2 for details in calculating these variables.

3. European Socio-Economic Classification (ESeC). The development and validation of ESeC was the result of a joint European collaboration (see Rose and Harrison, 2007) with the purpose to provide researchers with a standardized measurement of occupational status across European countries. Two variants of this classification with similar operationalizations were created for all participating country samples in ESS3, ISSP05 and ISSP06; they were labelled according to the original literature as "full" and "simple". As it can be read in the Access DB the so called "full" version is based on several harmonized occupational indicators, whereas the "simple" version uses only ISCO88 as an indicator.

4. Wright's Socio-Economic Classification. Because this classification has different theoretical roots as compared to ESeC, harmonisation materials for the ESS1 to 3 and ISSP05 were created following the guidelines outlined in Wright (2005) and by adapting harmonisation routines developed by Leiulfsrud et al. (2005).

5. International Socio-Economic Index (ISEI). ISEI was chosen by Wp9.2 as an alternative to the socio-economic classifications for the conceptualization of occupational status. This index was developed by Ganzeboom (Ganzeboom and Treimann, 1992) based on various mobility and stratification international data measured between 1968 and 1982. Indexes scores were assigned to all employed respondents in the ESS3, ISSP5 and ISSP06 data. The main limitations of this harmonisation as delivered here consists in not performing the bridging of ISCO88 codes into ISCO88(com) when applying the conversion syntax developed by Ganzeboom et al. (2003) for the ESS3 data.

6. National pride scales. Although supra-national entities (like the EU) become increasingly relevant in the present globalized world, the way individuals structure their identity in relation to their nation is still an unfilled gap in empirical comparative research. Based on previous findings (Smith and & Jarko, 1998) 16 Items on national pride asked across 22 countries in ISSP95 and ISSP03 were analyzed in an exploratory way and assigned to 2 or 3 possible meaningful underlying dimensions (*nationalism, patriotism* and *pride on specific national achievements*). This grouping of items is presented in the document "Descriptive Results", Section C. No active conversion had to

---

[4] Abbreviations for data sets cited in this document: ISSPXX: different ISSP modules for the respective years; ESSX: ESS Round 1(2002) and ESS Round 3 (2006).

be performed with these items. Source questions and variables for these items were stored in the Access DB.

For a complete overview of all country samples and data sets used see Table 1 in the Appendix. Table 2 in the Appendix gives and overview of all source variables used for deriving the harmonized variables

The main harmonisation materials (incl. the conversions) can be consulted in the Access DB (D9.1 mdb. Available at: http://www.cessda.org/project/deliverables.html)[5].

3 Structure of the Access DB: Forms and fields list

The Access DB consists from four tables:

1. ClassificationScales: where the theoretical basis of the measurement are given (incl. theoretical definitions and relevant references). For reading this table the form "Classifications" may be opened (see 3.1 below).

2. ConversionRoutines: where a description of the conversion process and the conversion syntaxes for creating the standardized variables was stored; for reading this table opens the corresponding form in the DB (see 3.2 below).

3. DatasetsVariables: where basic information on variables and data sets is delivered.

4. Question of Variables: where the literal text of the variable is stored.

For reading the database it is recommended to open only the two data forms (called "ClassificationScales" and "Conversions"). Both tables on source data (3 and 4) are sub- tables that can be accessed through the "Conversions" form (see 3.2 below).

In the "real" MySQL DB of the prototype some fields will be subdivided in tables and additional fields created.

There were several fields listed below that allowed for commenting on other content elements: 1.9 Conceptual Limitations, 2.8 Technical Comments, 2.12 Substantive Comments, and 2.15 Coding Comments. It was hard to define in advance which exact kind of comment will be entered here and who the legitimate author of such comments is.

Since some of the fields to be filled in the Access DB were strongly required and others were only optional, they are marked accordingly in the list below:

N = necessary

O = optional

3.1 Form "ClassificationScales"

---

[5] The conversion routines written for the harmonisation of private households can be found in Korré and Linardis, 2009 (*Harmonization of private household data across EU countries, 2009*).
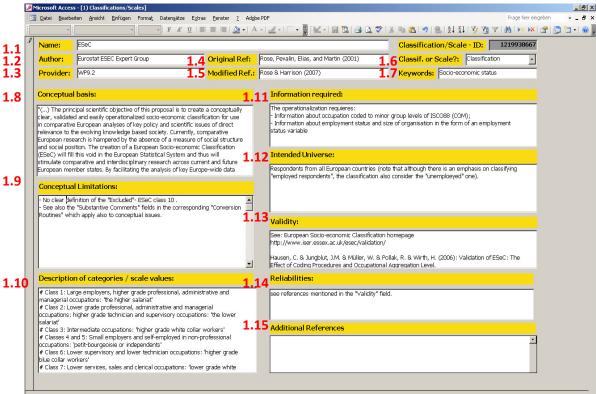
Figure 1: *Screenshot of the Access DB form "ClasssifcationScales"*

1.1. Name (N): Name of the Classification; common abbreviations first

1.2. Author (N): intellectual originator of the classification, i.e. the person, team, or organisation who defined the concept of the classification or scale and made the substantial category/value definitions.

1.3. Provider (N): Person/team/organisation providing/technically entering the scale into the system (GEODE uses the term 'supplier' here)

1.4. Reference of Original: (N) Exemplary, representative publication where the Classification/Scale has been presented for the first time.

1.5. Reference of Modified: (O) Presentation of a new / modified version of a previous one (e.g. the EGPs; ISCED-97; new versions of a social/attitude scale).

1.6. Classification/Scales: (N) Drop list for specifying if a classification OR a scale is documented

1.7. Keywords: (N) Free defined keywords were used (controlled vocabularies will be used in the later DB).

1.8. Conceptual Basis (new field) (N): definition, aim, what does the classification measure?

1.9. Conceptual Limitations: (O) conceptual limitations of the scale that are either already postulated by the author and / or added by any other person

1.10. Description of Categories / Scale Values: (N) brief explanations for the categories, or scale-values; e.g., descriptions in few lines of the EGP-categories; in case of ISEI, of the range of index-values; short description of the ISCED- levels; rating scales: meaning of anchor points, etc. For long classifications (e.g. ISCO88), this may point to external documents for the moment (e.g. ICSO 88 documentation by ILO).

1.11. Intended Universe: (N) The universe at which the classification conceptually aims; e.g., ESeC - only European countries, ISCED - originally only OECD-countries; for psychological instruments - countries/groups/samples for which the scale was designed.

1.12. Information Required: (N) All (or the minimum) information required for building the classification / harmonized target variable (i.e., which variables in the source data set would be needed). Classifications can have several conversion/derivation methods, depending on the information/variables available in the data – these differing derivation methods should also be listed here.

1.13. Validity: (O) References or results supporting validity

1.14. Reliabilities: (O) References or results supporting reliability

1.15. Additional Reference: (O) Here all relevant publications to a specific classification were added.
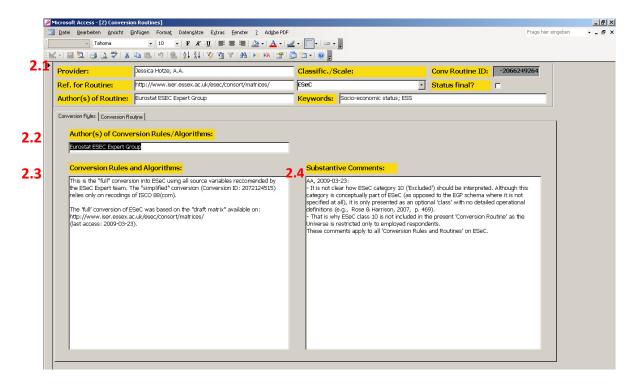
3.2 Form "Conversions"

Figure 2a: *Screenshot of the Access DB form "Conversions"*

2.1. Provider: (N) Person/team/organisation providing/technically entering the routine into the system. Later: automatically the logged-in User.

2.2. Author(s) of Conversion Rules / Algorithms: (O) intellectual originator of the conversions rules and algorithms, i.e. the person or team who made the substantial category assignments or designed the transformation algorithms.

2.3. Conversion Rules and Algorithms: (O) Operationalization or prescription for implementation of the underlying classification. A conversion/derivation method building on ideal types or ideal cases, assuming that all required information (field 1.10) is available in the data. This field either refers to a description of such codings - e.g., ISCED-Maps, an algorithm and parameter set for removing national item bias etc., or it lists the description in full.

2.4. Substantive Comments: (O) Comments on the validity of the operationalisation (2.3). Any doubt was mentioned here; comments on the substantial conversion decisions, use and application of the target variables. The source of comments (provider / author / author of routine) was provided in the comment field.
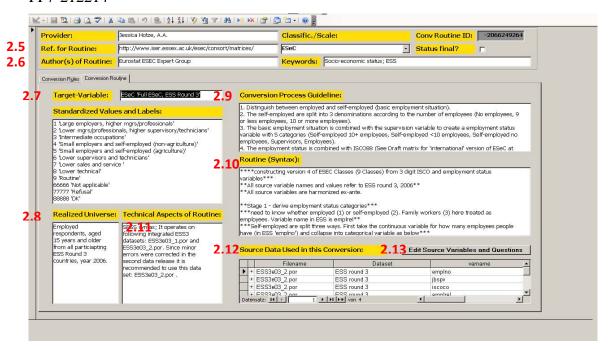
Figure 2b: *Screenshot of the Access DB form "Conversions"*

2.5. Reference: (O) Full bibliographical information, URL, etc. on where routine or work with that routine has been published for the first time (if ever). For newly created routines, this field remains blank.

2.6. Author(s) of Routine: (N) the person or team that utilized the Conversion Rules and Algorithms on a specific dataset / set of variables (later on: often same person as provider).

2.7. Target Variable: (N) In this field only the name of the final variable of interest (like EGP) was recorded. In some conversions, several interim harmonized variables were required in order to create the harmonized target variable. Such variables were listed in the conversion process guideline field (2.9).

2.8. Realized Universe (new field): (N) implicitly defined by the source data used, however, this field makes this explicit by listing the samples (e.g. geographical units; cultural groups) and points in time to which the routine could be generalized.

2.9. Conversion Process Guideline: (N) This field comprises a short description of the practical steps used in deriving the Target Variable with the data at hand, e.g., EGP: Step 1: Code occupation into ISCO 88; Step 2: Derive respondent's employment status/organization size… Step 5: With all variables produced in step 1 to 4 assign EGP categories- assignment is produced using the coding rules- field 2.3

2.10. Routine: (N) the full routine in e.g. SPSS syntax, ready for execution. Full variable, file names and brief instructions for the execution were here included.

2.11. Technical Comments: (O) Comments on technical aspects of the routine.

2.12. Source data used in this conversion(s): (N) source data sets (with file name and with version information, if available) and source variables used (name and label) are included here (subtable DatasetsVariables).

2.13. Edit source variables and questions: (N) when opening this dialog, source variables and question's literal texts in different languages for each variable can be read/edited (Figure 3 below). Since different variants classifications/scales used the same source data, the corresponding variables and questions were fully stored only once.
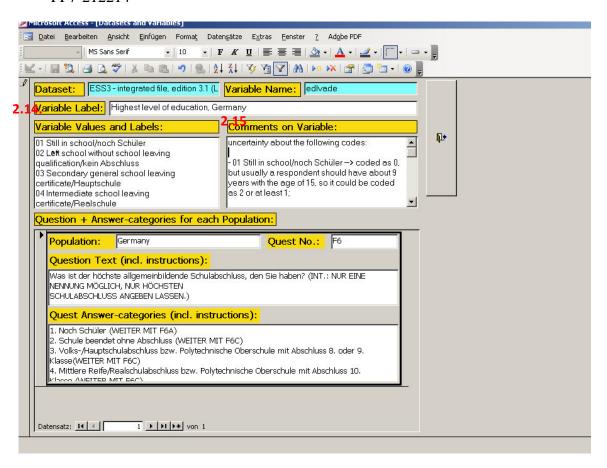
Figure 3: *Screenshot Access DB form "Edit source variables and questions"*

2.14.   Values and Labels of the Target Variable: (N) Technical definition and description of the target variable as created in the modified, final data set.

2.15.   Comments on variable: (O) these are comments made on coding. This field is also accessible via the separate dialog named "Edit source variables and questions" (2.13). Comments in this field are notes made on the correspondence lists of source variables to target variable; limitations of the source data compared to requirements for the ideal coding (field 1.12., 2.3) and / or other difficulties in applying the operationalization prescriptions (field 2.3). These comments are at the source variable level and the corresponding field was assigned in the sub-table DataSetsVariables.

4   Additional files (to D9.1.mdb) on data coding

As previously mentioned, the Access DB fulfilled the function of storing the main files for a quick understanding and application of the conversions and as an exercise of the involved researcher ("Providers") in designing the logic of the actual DB. But what CHARMCATS and the other two CESSDA DBs are supposed to support in this methodological work can be also exemplified by the additional working "instruments" used for organizing the relevant information on data coding procedures that are not included in D91.mdb. Here, only two of these additional created files are distributed that are also referenced in the Access DB:

1. ISCED-97 codings.xls (4.1): This file documents the ISCED-97 coding employed by W9.2 and could also give an insight on how and why the decisions on coding were undertaken in detail. It also shows the limits from a practical point of view on using such kind of files.

2. Income data.xls (4.2.): This file stores the main macro data on Purchasing Power Parities and the data set variable information that were used for harmonising income for cross-country comparisons.

4.1   ISCED-97 coding with ESS3, ISSP05 and ISSP06

This file groups several Excel spreadsheets that summarize following working steps:
1. <u>Overview ISCED countries</u>: Here, the selected country samples are listed for which ISCED-97 re-classifications were performed on the basis of ESS3, ISSP05 and ISSP06 data. A total of 17 selected European countries were used and for 11 countries re-classifications could be made across all 3 data sets.
2. <u>ESS Round 3 (2006)</u>: ISCED coding for the ESS3 data was conducted by two trained coders. As this was an exercise that involved one less experienced and a more experienced researcher from GESIS, a complete blind coding couldn't be established: cases where the coding lead to different interpretations between coders were discussed and mutual decisions were taken afterwards. These final decisions on coding were stored accordingly in the column labelled "Wp9.2 final coding". Cases were coding was difficult are clearly highlighted within cells whereas brief explanatory notes on coding where stored within the table columns labelled with "Comments". In a second phase, these re-classifications were also discussed within Task 9.2 and compared to the ISCED-97 variable constructed by the ESS team (stored in the column on the ESS variable "EDUVL"). Cases where there was a discrepancy between the ESS and WP9.2 variant of ISCED are highlighted with red.
3. <u>ISSP 2005</u>: Like in the case of ESS3, ISCED-97 re-coding was performed for the ISSP 2005 data by two researchers. However, in this file only the final coding of WP9.2 is stored. Mappings of the country specific educational variables into the harmonized education variable provided by the ISSP (DEGREE) are also stored in a separate column.
4. <u>ISSP 2006</u>: This table on ISCED-97 coding with ISSP 2006 data has the same structure as the previous table with ISSP 2005.
5. <u>Recoding matrix Germany</u>: This table documents only for the case of Germany how the reliable recoding into ISCED-97 could have been performed with the ISSP data if additional source data would have been available from the German General Social Survey (ALLBUS).

Difficulties in coding the ISCED-97 could be summarized as a consequence of:
a. The insufficiencies provided by the description in the literature of the ISCED-97 underlying dimensions and their operationalizations;
b. Limited available information on educational qualification provided by the country specific variables. Obviously, only the later aspect is more visible in this file.

What did this work also showed? First, it is evident that the recoding of national educational categories into international ones will reflect only partially the criteria as set out in the international typological framework, particularly if it is performed by different researcher in different countries and with varying levels of expertise. This aspect is also repeatedly stressed in the literature on comparative measurement of education (Schneider, 2008). The use case with ISCED-97 presented here revealed that for about 30% of ISCED-97 mappings no reliable harmonisation could be established, regardless of the used data set (in the sense of "perfect" agreement between coding performed by different researcher). Given the differences in the available data it would be very demanding to have a platform where all these coding procedures can be made transparent with stored available documentation (metadata) on variables in "one place" (for example, about 80 documents provided by the archives were consulted for every conversion routine on ISCED-97 in order to have the full literal text of questions and their coding into variables). Second, having stored in the same system conceptual and operational description of the ISCED-97 classifications would support

drawing a clearer line between interpretational bias in coding originated by the researcher who conducted the harmonisation and bias which is per se given by limitations in the available data.

4.2   *Harmonizing household income data*

Not only data on the individual level is required for harmonisation, but also data on the macro-context level the survey was conducted. This is certainly the case for transforming household income for comparative purposes. As it can be read in the Access DB, equivalizing the net household income for the ESS3 and ISSP06 data involved in a first step transformation of net and gross household income with Purchasing Power Parities (PPP) calculated in EUR and international dollar.

The PPP data was provided by the EUROSATAT (2008) and the World Bank (2009) online databases. Because income was measured on interval income categories with some country specific variants it was decided to use only mid points of this income categories after transformations with the PPP. The excel table documents all this "mid points" of income categories that were further used for weighting income with harmonized variables on household size and household members (household head, adults and children).

5   Harmonization of private household with data from EU-SILC and HBS

Harmonisation materials on household are stored in a separate document (Korré and Linardis, 2009). This document provides an example to illustrate the structure for the harmonization database application (presented in section 1.4 of D9.2). It focuses on the measurement of household across the EU countries for the year 2005. The construction of the household typology is presented as measured in the EU-SILC harmonized survey (Eurostat 2004, and Eurostat 2008a).Furthermore, the recommendations for the harmonization of household with the HBS data (European Commission 2003b) are showed. In general, the structure and content of the harmonisation documentation follows the three basic elements structure of the CHARMCATS DB model, i.e. the Conceptual Step, the Operational Step and the Data Coding Step. In addition, the EU-SILC household typology is operationalized with national specific indicators and variables; in this case, the harmonization project has the same conceptual derivation but varying country and time specific indicators and variables.

6   Descriptive analyses tables

The document titled "Descriptive Results" stores distributions of the harmonized codes of classifications (section B), analyses on the comparability of scales (Section C), distributions of Scales/Indexes and country within correlational associations between variables measuring similar concepts (Section D). In the first section (A) table 1 and 2 from the appendix (with an overview of the harmonized data) are reproduced.

All tables are preceded by a brief explanatory text whereas results that could be relevant for face and criterion validity are also highlighted within tables.

For example, in section B tables 3 to 14 may answer the question "What is the highest level of education attained across 17 countries?" using as indicator the ISCED-97 classification. But, these tables point also to inconsistencies between distributions that could be only interpreted if the limitations and ambiguities aroused by data recoding (see Access DB and Excel files) are known.

These tables will be continuously updated in the remaining project months.

7   Concluding remarks

- Harmonisation work involves working with a set of very diverse documentation (Section 2 and 3) ranging from journal papers or books to metadata on variables available online (on data archive WebPages or NESSTAR based systems) and .pdf documents with survey questionnaires.
- The usage of these materials is in practice made ad hoc (see Section 4) and is difficult to understand for "externals".
- Analyzing the harmonised data within one system for creating for example descriptive statistics (Section 6) would help users of the CESSDA portal to have a first documentation on the quality of their data coding and on others harmonisation work.

Besides the harmonisation work stored in the Access DB, following other harmonisations are under progress and could be used as testing materials for CHARMCATS by publishing them as so called "harmonisation projects":

- Left right scales: reclassifying respondents preferences for their national parties into a comparable left-right scale across the ISSP (Modules: 2005 and 2006) and ESS3 data were analyzed and a first version of this harmonized variable was created;
- Three items on the concept of "Interpersonal Trust" asked in a similar way in different surveys (ISSP modules, World Values Survey and ESS; time span: 1980-2006, across 103 countries/cultural groups) were compared and could be included as an example of harmonisation in CHARMCATS when construction of scales is involved;
- Education: Literature review on the CASMIN classification (Braun & Müller, 1997) was undertaken, and at least its conceptual basis could be used for contrasting it to the ISCED-97 classification within the future prototype. However, contradictory coding procedures were presented in the literature (e.g., Kerckhoff et al., 2002) and thus this classification has to be further carefully examined before developing a harmonisation routine;

Literature

Atkinson, T., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Social indicators*. Oxford: Oxford Univ. Press.

Bourmpos, Michael; Linardis, Tolis (with Alexandru Agache, Martin Friedrichs, and Markus Quandt) (2009): *D9.2 Functional and Technical Specifications of 3CDB.*

Braun, M., & Müller, W. (1997). Measurement of education in comparative research. *Comparative social research, 16*, 163-201.

EUROSTAT (2008). *EUROSTAT Database. Purchasing power parities (PPPs), price level indices and real expenditures for ESA95 aggregates*. Publication: http://nui.epp.eurostat.ec.europa.eu/nui/show.do?dataset=prc_ppp_ind&lang=en (retrieved on 19 May 2008).

Ganzeboom, H. B. G., & Treimann, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variables* (pp. 159-193). New York: Kluwer Academic/Plenum.

Ganzeboom, H. B. G., de Graaf, P. M., & Treimann, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*, 1-56.

Gordon, D. & Townsend, P. (Eds.). (2000). *Breadline Europe.* Bristol: Polity Press.

Hagenaars, A., K. de Vos and M.A. Zaidi (1994). *Poverty statistics in the late 1980s: Research based on micro-data*. Luxembourg: Office for Official Publications of the European Communities.

Kerckhoff, A. C., Ezell, E. D., & Brown, J. S. (2002). Toward an improved measure of educational attainment in social stratification research. *Social Science Research*, 31, 99-123.

Leiulfsrud, H., Bison, I., & Jensberg, H. (2005). Social class in Europe. European Social Survey 2002/3. Trondheim: NTNU Social Research Ltd., Retrieved from http://ess.nsd.uib.no/files/2003/ESS1SocialClassReport.pdf , on 2009-08-05.

OECD (1999). *Classifying educational programmes. Manual for ISCED-97 implementation in OECD countries*. Paris: OECD Publications Service.

Rose, D., & Harrison, E. (2007). The European Socio-economic Classification: A new social class schema for comparative European research. *European Societies, 9*(3), 459-490.

Schneider, S. (Ed.) (2008). *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung (MZES).

Smith, T. W., & Jarkko, L. (1998). National pride: A cross-national analysis. *GSS Crossnational Report* No. 19. Chicago: NORC.

UNESCO Institute for Statistics (2007). *Global education digest 2007: Comparing education statistics across the world*. Available from http://www.uis.unesco.org/template/pdf/ged/2007/En_web2.pdf (retrieved on 2 April 2009).

Wright, E. O. (2005). *Approaches to class analysis*. Cambridge: Cambridge University Press.

Appendix

Table 1: Overview of harmonized variables sorted by countries

								Harmonized variables[2]			
Country	**Year**	**Dataset**	**Wright**	**ESeC**	**ISCED-97**	**ISEI**	**Scales of national pride**	**Income PPP Eur (Net)**	**Income PPP $ (Net)**	**Income PPP $ (Gross)**	**Income Ratio**
Australia	1995	ISSP95					X				
	2003	ISSP03					X				
	2005	ISSP05		X		X					
	2006	ISSP06		X		X				X	X
Austria	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2006	ESS3	X	X		X			X		
Belgium	2002	ESS1	X	X		X					
	2005	ISSP05				X					
	2004	ESS2	X								
	2006	ESS3	X	X		X			X		
Bulgaria	2005	ISSP05		X		X					
	2006	ESS3	X	X		X			X		
Canada	1995	ISSP95					X				
	2003	ISSP03					X				
	2005	ISSP05		X		X					
	2006	ISSP06		X		X				X	X
Chile	2006	ISSP06		X		X				X	X
Croatia	2006	ISSP06		X		X				X	X
Cyprus	2005	ISSP05		X		X					
	2006	ESS3	X	X		X			X		
Czech Re-public	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X		X					
	2006	ISSP06		X		X				X	
Denmark	1995	ISSP95									
	2002	ESS1	X	X		X					
	2003	ISSP03									
	2004	ESS2	X								
	2005	ISSP05		X		X					
	2006	ESS3	X	X		X			X		
	2006	ISSP06		X		X				X	X
Dominican Republic	2005	ISSP05		X		X					
	2006	ISSP06		X		X				X	X
Estonia	2004	ESS2	X								
	2006	ESS3	X	X		X					
Finland	2002	ESS1	X	X		X					
	2004	ESS2	X								
	2005	ISSP05			X						
	2006	ESS3	X	X		X		X	X		
	2006	ISSP06			X					X	X
France	2002	ESS1	X	X		X					
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X	X	X				X	X
Germany	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X	X	X				X	X

Table 1: Overview of harmonized variables sorted by countries, years and data set (Part 2)

Country	Year	Dataset	Harmonized variables[2]								
			Wright	ESeC	ISCED	ISEI	Scales of national pride	Income PPP Eur (Net)	Income PPP $ (Net)	Income PPP $ (Gross)	Income Ratio
Great Britain	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X		X			X		
	2006	ISSP06		X	X	X			X		X
Greece	2004	ESS2	X								
	2002	ESS1	X	X		X					
Hungary	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X	X	X					
Ireland	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X		X			
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X	X	X		X		X	X
Israel	2002	ESS1	X	X		X					
	2005	ISSP05		X		X					
	2006	ESS3	X	X		X					
	2006	ISSP06		X		X				X	X
Italy	2002	ESS1	X								
	2004	ESS2	X								
Japan	1995	ISSP95					X				
	2003	ISSP03					X				
	2005	ISSP05		X		X					
	2006	ISSP06		X		X					X
Latvia	1995	ISSP95					X				
	2003	ISSP03					X				
	2005	ISSP05			X						
	2006	ESS3	X	X		X		X	X		
	2006	ISSP06		X	X	X				X	X
Luxembourg	2002	ESS1	X	X		X					
	2004	ESS2	X								
Mexico	2005	ISSP05		X		X					
Netherlands	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X		X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X		X			X		X
New Zealand	1995	ISSP95		X		X	X				
	2003	ISSP03					X				
	2005	ISSP05		X		X					
	2006	ISSP06		X		X			X		X
Norway	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3,	X	X	X	X		X	X		
	2006	ISSP06		X	X	X				X	X
Philippines	1995	ISSP95					X				
	2003	ISSP03					X				

16

FP7-212214

Year	Survey				
2005	ISSP05	X	X		
2006	ISSP06	X	X		X

Table 1: Overview of harmonised variables sorted by countries, years and data set (Part 3)

Country	Year	Dataset	Wright	ESeC	ISCED	ISEI	Scales of national pride	Income PPP Eur (Net)	Income PPP $ (Net)	Income PPP $ (Gross)	Income Ratio
Poland	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X		X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP		X		X				X	X
Portugal	2002	ESS1	X	X		X					
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X	X	X		X	X		
	2006	ISSP06		X	X	X			X		X
Romania	2006	ESS3	X	X		X					
Russia	1995	ISSP05					X				
	2003	ISSP03					X				
	2005	ISSP05		X		X					
	2006	ESS3	X						X		
	2006	ISSP06		X		X					
Slovenia	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ESS3	X	X		X					
	2006	ISSP06		X	X	X					X
Slovakia	1995	ISSP95					X				
	2003	ISSP03					X				
	2004	ESS2	X								
	2006	ESS3	X	X		X			X		
South Africa	2005	ISSP05		X		X					
	2006	ISSP06		X		X			X		X
South Korea	2005	ISSP05		X		X					
	2006	ISSP06		X		X			X		X
Spain	1995	ISSP95					X				
	2002	ESS1	X	X		X					
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X		X					
	2006	ISSP06		X		X			X		
	2006	ESS3	X	X	X	X		X	X		X
Switzerland	2002	ESS1	X	X		X					
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ISSP06		X	X	X			X		
	2006	ESS3	X	X	X	X		X	X		X
Sweden	1995	ISSP95					X				
	2003	ISSP03					X				
	2004	ESS2	X								
	2005	ISSP05		X	X	X					
	2006	ISSP06		X	X	X			X		
	2006	ESS3	X	X	X	X		X	X		X
Taiwan	2005	ISSP05		X		X					
	2006	ISSP06		X		X					X
Turkey	2004	ESS2	X								
Ukraine	2004	ESS2	X								
	2006	ESS3	X	X		X					
United States	1995	ISSP95					X				
	2003	ISSP03					X			X		X
Urugay	2006	ISSP06		X		X				X	X
Venezuela	2006	ISSP06		X		X					X

Note:

1) ISSPXX: different ISSP modules for the respective years; ESSX: ESS Round 1(2002) and ESS Round 3 (2006);

Table 2: Overview of harmonized and source variables

Harmonized Variable[1]	Countries	Datasets[2]	Variables used for harmonization	Comments
ESeC	Australia, Austria, Belgium, Bulgaria, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Estonia, Finland, France, Germany, Great Britain, Greece, Hungary, Ireland, Israel, Japan, Latvia, Luxembourg, Mexico,Netherlands, New Zealand, Norway, Philippines, Poland, Portugal, Romania, Russia, Slovenia, Slovakia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Ukraine, Urugay, Venezuela	ESS1 (2002), ESS3 (2006), ISSP 2005, ISSP 2006	ESS: emplno (number of employees), jbspv (supervising other) Iscoco (ISCO88 com) emplrel (employment relation) ISSP: wrkst (employment relation), wrksup (supervising other), isco88 (ISCO), nemploy (number of employees), wrktype (worktype)	- The 4-digit ISCO88 codes were converted into 3-digit codes - ESeC for respondent's spouse only for ESS data
ISCED	ESS 2006: France, Germany, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland ISSP 2005/2006: Finland, France, Germany, Great Britain, Hungary, Ireland, Lavita, Norway, Portugal, Slovenia, Sweden, Switzerland	ESS3 (2006), ISSP 2005, ISSP 2006	ESS: edulvXX (national coding of highest education level) ISSP: XX_dgre (national coding of highest education level)	
ISEI	Australia, Austria, Belgium, Bulgaria, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Estonia, Finland, France, Germany, Great Britain, Greece, Hungary, Ireland, Israel, Japan, Latvia, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Portugal, Russia, Slovenia, Slovakia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Ukraine, United States, Uruguay, Venezuela	ESS3 (2006), ISSP 2005, ISSP 2006	ESS: iscoco (ISCO88 com) ISSP: isco88 (iSCO88)	
Income (New OECD-Scale)	Finland, France, Germany, Hungary, Ireland, Latvia, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland	ESS3 (2006)	hinctnt (householdincome), yrbrn-yrbrn15 (member of the houshold's year of birth), countries' ppp	No conversion for ISSP possible due to inconsistent measurement of the household income
Scales of national pride	Austria, Australia, Canada, Czech Republic, Germany, Great Britain, Hungnary, Ireland, Japan, Latvia, Netherlands, New Zealand, Norway, Philipines, Poland, Russia, Slovenia, Slovakia, Spain, Sweden, United States	ISSP 1995/ ISSP 2003	V22/V19 (Rather be a citizen of -Rs country), V23/V20 (Things about -Rs country- feel ashamed), V24/V21 (World better place other like -country), V25/V22 (Rs country- better country than other), V26/V23 (Support their country even is wrong), V27/V24 (Well in international - makes proud), V28/V26 (Proud of: way democracy works), V29/V27 (Proud of: political influence in world), V30/V28 (Proud of: economic achievements), V31/V29 (Proud of: social security system), V32/V30 (Proud of: scientific achievements), V33/V31 (Proud of: achievements in sports), V34/V32 (Proud of: achievements in arts), V35/V33 (Proud of: armed forces), V36/V34 (Proud of: its history), V37/V35 (Proud of: fair treatment of groups)	

Table 2: Overview of harmonized and source variables (Part 2)

Harmonized Variable[1]	Countries	Datasets[2]	Variables used for harmonization	Comments
Wright	Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom	ESS1 (2002) ESS2 (2004) ESS3 (2006)	Iscoco, Emplno, Emplrel, Jbspv, Orgwrk, Wkdscin	

Note:
1) ISSPXX: different ISSP modules for the respective years; ESSX: ESS Round 1(2002) and ESS Round 3 (2006);

2) ESeC: European Socio Economic Classification; ISCED-97: International Standard Classification of Education; Income (new OECD Scale): Netto Household Income, transformed with PPP in EUR and calculated using the weighting of the so called "new OECD scale"; Scales of national pride: 16 Likert Items, asked in ISSP Module 1995 and 2003; ISEI: International Socioeconomic Index (Ganzeboom et al., 2003).

Acknowledgments

Jessica Hotze and Vanita Matta (both student assistants at GESIS) had a great contribution in applying the conversion routines to the data, transferring the metadata into the ACCESS DB and in creating the descriptive analyses with harmonized variables. Of course for any errors, only the authors of this report are responsible.