Leibniz-Gemeinschaft

Final report

Title of the project: Computer-assisted measurement of education in surveys (CAMCES)

Stand: 28.02.2018

gesis
Leibniz Institute
for the Social Sciences

Computer-assisted measurement of education in surveys (CAMCES)

Final project report

Grant number SAW-2013-GESIS-5

Silke Schneider,
Roberto Briceno-Rosas,
Jessica Herzing,
Verena Ortmanns

Table of contents

Executive summary

The project "Computer-assisted measurement of education in surveys" (CAMCES) had the aim to develop tools to facilitate the measurement and cross-nationally comparable coding of educational attainment in computer-assisted surveys. The CAMCES tools intend to support both cross-nationally comparative surveys that need to harmonize educational attainment data across survey countries, as well as (national) surveys of migrants, whose educational attainment is difficult to measure given their various educational backgrounds.

To this aim, firstly, a database was developed, identifying educational systems, educational qualifications and their international classification codes, mostly based on UNESCO's International Standard Classification of Education (ISCED). Secondly, a short questionnaire module was developed, to identify the relevant educational system (and language, for multilingual countries) and then ask for the respondents' highest educational qualification. This questionnaire module was translated into French, English and Spanish so that the CAMCES tools can be impelemented in a range of countries without the need for further translation. Thirdly, two interfaces for accessing the database from the questionnaire during the survey were designed: the combination box, searching the database using text string matching, and the search tree, allowing respondents to 'browse' the educational qualifications available in the country they were educated in. The CAMCES tool was developed for both computer-assisted personal interviewing (CAPI) and computer-assisted web interviewing (CAWI) surveys.

We extensively tested successive versions of the CAMCES tool by various methods, using both general population samples as well as migrant samples, both personal interviewing as well as web interviewing techniques, and using quantitative (survey) as well as qualitative methods (such as two cognitive pretesting studies). We also conducted comprehensive desk testing of each version before using it in the empirical studies reported here. This mix of methods was intended to ensure that we capture most bugs and problems and could optimize the tools as much as possible before publication.

As one example, the CAMCES tool has been successfully implemented on a pilot basis in the IAB-SOEP Migration Sample in 2015 and 2016 (run by the German Institute for Economic Research, DIW, in Berlin), funded by the CAMCES project. An optimized version of the software was implemented in 2017, and for the 2018 data collection, the database was also updated. The CAMCES project has thus produced more adequate data on the educational attainment of migrants in Germany already.

The CAMCES tool and its various components are distributed free of charge at https://www.surveycodings.org/education. It is also open source, to be maximally accessible and useful to the scientific community. For surveys not using the CAMCES tools for data collection, the underlying database, available online at the above URL, will be a rich resource supporting more consistency in ex-post coding of educational qualifications (e.g. as a result of open responses by respondents mentioning 'other' educational qualifications) across countries.

Since the end of the CAMCES funding period, more countries are continuously being added to the CAMCES database within the context of the EU project "Synergies for Europe's Research Infrastructures in the Social Sciences" (SERISS), Work Package 8 "A coding module for socio-economic survey questions" which runs until mid-2019. This project also for the time being ensures the technical maintenance of the online database and the www.surveycodings.org web portal, and will support future dissemination activities. Further countries were also added by integrating CAMCES in the BAMF-IAB-SOEP Survey of Refugees 2017, and the project "ReGES – Refugees in the German Educational System" conducted at the Leibniz Institute for Educational Trajectories (LifBi).

1 Project motivation and aims

Education is an important cause, mediator and outcome of social inequality. The individual's level of education to a considerable degree influences his/her labor market and wider life chances, e.g. income and health, but also his/her children's education. Educational attainment is also an influential factor of individuals' behavior, values and social attitudes. One reason for those relationships is that education increases the individual's human and cultural capital (specific knowledge, skills, and competences) that have positive outcomes in later life. Another reason is the socializing effect of education with respect to norms and values. Finally, educational qualifications assume symbolic meaning and signaling power on their own. Therefore, education is a concept of extensive academic and policy interest at the national and international levels. There are numerous comparative studies involving educational attainment, e.g. on educational inequality (e.g. Breen, Luijkx, Müller, & Pollak, 2009; Shavit & Blossfeld, 1993), immigrant integration (e.g. Kristen et al., 2011; Kristen & Granato, 2007), returns to education (e.g. Bol & Van De Werfhorst, 2011), school-to-work transitions (e.g. Kogan, Noelke, & Gebel, 2011; Müller & Gangl, 2003), and education/work mismatch (e.g. Allen & van der Velden, 2001; Levels, van der Velden, & Di Stasio, 2014).

The individual's educational attainment is therefore a core social background variable and captured in virtually all surveys of individuals (Smith, 1995). It is usually measured by the highest educational qualification the individual has obtained. When surveys ask for the highest educational qualification obtained, they usually use a closed-ended question format that offers respondents a list of country-specific educational categories for self-classification. The advantage of this response design is that it is not very costly for researchers in terms of time and labor, and not very burdensome for respondents anc can be applied across survey modes.

However, this approach also brings along some challenges, especially in cross-national surveys (Braun & Mohler, 2003; Braun & Müller, 1997; Schneider, 2009). Each educational system has its unique idiosyncratic institutions, and their certificates often have proper names which rule out translation (Schneider, Joye, & Wolf, 2016). The state-of-the-art for cross-national surveys is thus to use country-specific questionnaire items for education, which are coded into a cross-nationally comparative coding scheme after data collection. Ensuring that this is done consistently across countries is a major challenge and requires a major logistical effort. Research has shown that this is actually often rather unsuccessful, and that the education distributions substantially differ, within countries and years, across surveys (Ortmanns & Schneider, 2016a, 2016b).

Increasing differentiation of educational systems and migration complicate the measurement of educational attainment in surveys considerably. The introduction of new qualifications via educational reforms means that the list of response options to the survey question on education either gets longer or less accurate over time. The display of categories, restricts the respondents to a limited choice-set (Schneider, 2008), they cause confusion if respondents read unknown terms and might lead to response order effects (Couper, Tourangeau, Conrad, & Crawford, 2004). This method makes it difficult to present complex education systems, which then cause troubles in recoding the answers.The response categories are in fact often incomplete – for example, they do not cover qualifications that were recently introduced, that are rare, or they lack outdated qualifications that are held by older respondents. Furthermore, a qualification may be known by different names, e.g. in different regions, only one of which is mentioned in the response categories.

The standard survey practice for individuals with foreign qualifications is to ignore the fact that the respondent holds a foreign qualification and ask them to choose the domestic response category that is the closest perceived equivalent to their foreign qualification. This practice leads to data losses because foreign qualifications are not recorded. It also leads to intentional or accidental misreporting because this practice assumes unrealistically detailed knowledge on the side of respondents and interviewers about the equivalence of foreign and domestic educational qualifications.

In times of growing mobility on the part of students and workers, who both import foreign educational qualifications into the survey country, foreign educational qualifications are increasingly common and important. Depending on their age at migration, migrants have often obtained their educational qualifications in their country of origin. One measurement approach, especially in general population samples, is to ask respondents with foreign educational qualifications to indicate the "equivalent" qualification in the educational system in which the survey takes place. This is likely burdensome for migrant respondents, especially for those who did not (yet) have much contact with the educational system or labor market in their host society. Another approach is to offer a list of categories that are generic descriptions of educational levels that are assumed to be understood universally by respondents from various backgrounds. This is, however, a strong assumption: Language ability and cultural differences in the understanding of constructs underlying a survey question may introduce measurement error (Kleiner, Lipps, & Ferrez, 2015). For example, generic terms such as "primary" or "mandatory education" correspond to schooling of different durations in different countries, and "secondary" education includes vocational training in some countries but not in others. Neither measurement approach – using survey country education categories or using generic categories – takes into account institutional differences between educational systems across the world, and large measurement errors are likely. Without adequate data on migrants' educational attainment however, it is very difficult to investigate the integration of migrants e.g. in the domestic labor market, or the "portability" and skill mismatches regarding migrants' educational qualifications (Tijdens & van Klaveren, 2012). In order to achieve acceptable measurement quality and thereby comparability across respondent groups from different origins, we argue that it is desirable to measure educational attainment using country-specific response categories in migrant surveys, following the example of cross-national surveys.

The aim of the CAMCES project was therefore to provide new survey measurement instruments for measuring individual's educational attainment that will enable the accurate and cross-nationally comparable coding of educational qualifications, especially facilitating the measurement and coding of educational qualifications across countries. Because of increasingly differentiated educational systems, even national surveys may benefit from an innovative re-design of the education question. To achieve these goals, we designed a short questionnaire module asking respondents to firstly indicate where they were educated and then to choose their highest qualification from a list of educational qualifications specific to the respective education system. The country-specific response options are generated via innovative interfaces to an underlying database of educational qualifications. After data collection, the detailed code of every educational qualification can be recoded to education classifications via syntax files provided with the CAMCES tool. The three-digit coding system of the International Standard Classification of Education (ISCED) 2011, provided by the UNESCO (2012), serves as the cross-national coding scheme. This is because it is available for virtually all countries in the world.

The instruments and tools were developed for survey modes which are computer assisted and where content can be shown visually to respondents. This is necessary because interviewers will, at least in the case of migrant respondents, often not be able to read the language of the country in which the respondent was educated. This applies to computer assisted personal interviewing (CAPI) and computer assisted web interviewing (CAWI).

It was hoped that usage of such a measurement strategy will significantly improve the surveying process and data quality, especially for cross-cultural (including migrant) surveys.

2   Project implementation and challenges

The project aims were planned to be achieved through five work packages: 1) develop and test an international database of educational qualifications and corresponding ISCED codes; 2) develop and test optimal survey questions, questionnaire routing and interviewer/respondent instructions assessing educational attainment and ways to link responses with

the database in an automated way; 3) develop and test a technical interface to employ the database, survey questions, conditional routing and instructions for use in computer-assisted surveys; 4) test the integrated service in CAWI and CAPI pilot studies; and 5) project coordination, expert meetings and activities ensuring the long-term sustainability of the service.

Database development, being a huge task in itself, is an ongoing task. The strategy for filling the database has evolved over time and did not always follow the initial plans. Firstly, information from a wide range of existing surveys was used to fill the database, cross-checking and complementing the information using the official ISCED mappings available at http://uis.unesco.org/en/isced-mappings. Initially, database development focused on Germany, the origin countries of migrants in Germany, the Netherlands and Venezuela, to satisfy the needs of the diverse test studies (see section 3.2). Not all of Europe is yet available and including the countries of origin of migrants in all European countries turned out not to be realistic. Nevertheless, as of October 2017, the CAMCES educational database contains 60 educational systems, including country/language combinations (corresponding to 50 countries). Secondly, developing the technical infrastructure for the database held up the process considerably, with various solutions being tested until the final solution was implemented. Both costs and time required for establishing the technical infrastructure for the project were seriously underestimated in the proposal. The process now involves developing the qualifications data per country in Excel and importing the resulting Excel file into an online database (where editing is possible, but much less convenient than in Excel). Issues such as versioning of database entries could not be tackled during the CAMCES project but will be resolved in the SERISS project (see section 4), which also funds the addition of many more countries to the database and further technical database development, so that the consequences of not sufficiently planning for this in the CAMCES proposal could be successfully mitigated.

Work packages 2 and 3 were combined early in the project because the relevant tasks were often closely related. The tasks (e.g. questionnaire translation) were also typically closely driven by the pretesting and pilot studies in work package 4. Interface development was a highly iterative process that required a lot of very detailed discussions with two different programmers (one being an expert in Windows application programming, working on the CAPI tool, the other being an expert in web survey programming, working on the web survey tool and web portal, see section 6.2.4) on the exact functionality of the tool. The level of detail of the specifications required for programming is very high and programmers often came up with unexpected questions, which were usually solved in a very satisfactory manner but required more time than anticipated. The project staff, including the project leader, not coming from a technical background, has learned a lot through these conversations, but it took time to establish a common terminology and a clear vision of the actual project products. Fortunately, and thanks to the strong ambition and commitment of the project team and programmers, no major changes to the intended functionality had to be made eventually. The only difference – and a positive one – is that coding during data collection turned out not to be advisable. The resulting data can be used more flexibly if coded in detail (i.e. at the qualification level) during data collection and recoded into a more aggregate international coding scheme only afterwards. This means that the coding scheme does not need to be determined before data collection but can be changed ex post by applying a different recoding syntax file to the data. The disadvantage of this approach is that the respective syntax has to be provided, which can in principle be automatically derived from the database. This still needs to be implemented though (again, this is currently prepared in the SERISS project) and so far, the recoding syntax was produced manually.

Regarding work package 4, we added two studies that had not been planned, because opportunities came up that we did not want to let pass. One junior researcher on the project applied for an experimental study in the SOEP Innovation Sample 2014 and won it. Another junior researcher had contacts at a sociological research lab in Venezuela so that a cognitive pretesting study outside of Germany could be run at little cost, adding to the cross-cultural evidence base of the project. In hindsight though, too many empirical studies were conducted by the CAMCES project, leaving too little time to digest the results. The studies them-

selves were very successful though (see section 3.2 for details and references to the reports).

As part of Work Package 5 (project management), the project was accompanied by three expert workshops held at GESIS in Mannheim in 2013, 2014 and 2016 (for details, see section 6.1 in the appendix). These workshops proved highly useful for getting an outsiders' view on the project work, generate solutions for problems we faced, and making the project known in the survey community. Furthermore, all team mebers regularly presented results of their work at international conferences.

It was not possible to reuse elements of the CASCOT software (Elias, 2012), developed for the ex-post coding of occupations, to CAMCES, as was hoped initially. The requirements for a survey instrument are too different, as are the conceptual and cross-cultural equivalence issues regarding occupation and education (e.g. in the domain of occupations, translation is possible, while it is not possible to translate the names of educational qualifications). Altogether, the CAMCES tool was conceptualized in too complex a way than to be feasible to fully develop in three years' time, which is why we were grateful for the cost-neutral extensions of the project runtime, and for getting the opportunity to extend the work via the SERISS project (see section 4). Some more advanced elements of the tools, such as database versioning, will only be available in 2019. The tool can nevertheless be used in the form in which it is available today.

Organizationally, the biggest challenge of the project was that one junior researcher left the project in late 2014 after successfully applying for a university PhD position with a longer running contract. A very suitable replacement could be found, and the handover eventually was relatively smooth. On the positive side, it meant that the project benefited from the different perspectives and skills of three rather than just two junior researchers over its lifetime. Still, replacing key personnel comes at a cost of introducing a new project member to the project background, aims and methods.

3   Project results

3.1   The CAMCES tool as a new measurement instrument

The CAMCES tool aims to alleviate some of the challenges associated with the measurement of educational attainment, especially in surveys of migrants and other cross-cultural surveys, by considering different educational systems and referring to specific educational qualifications. The tool combines (1) a short questionnaire module, (2) an international database of educational qualifications and their classification, and (3) two survey interfaces that dynamically show relevant response options from the database in the questionnaire (combination box and search tree). The tool thereby offers response categories relevant for the specific educational systems in which respondents completed their education.

Two versions of the CAMCES interfaces were developed to support different survey modes: a CAPI version for computer-assisted personal interviews and a CAWI version for web surveys. With regard to the CAPI version, the interviewer needs to hand over the device to respondents educated abroad who are then asked to report their educational qualification by themselves. This is necessary, firstly, because the interviewer may not speak the respective language. Secondly, the interfaces are, due to their dynamic nature, more suitable for self-completion.

The development of the CAMCES tool was accompanied by empirical research and subsequent optimizations. A first prototype of the CAPI version was discussed with a focus group consisting of interviewers and then tested in a quantitative pre-test in 2014 with a migrant quota sample (Schmidt, 2014, 2014; Schmidt & Maier, 2014). In 2015 and 2016, an updated version of the CAPI tool was implemented in the IAB-SOEP Migration Samples (see section 3.2.4 for details). Cognitive interviews using the CAWI tool were run in Germany in 2015 and

Venezuela in 2016 (Briceno-Rosas, del Busto, & Taoli, 2016; Lenzner et al., 2015); and in 2016 the CAWI tool was piloted in the Dutch LISS Panel.

3.1.1   The questionnaire module

The CAMCES questionnaire module has, as its core, the question asking for respondents' educational attainment but it also includes questions to identify the relevant educational system and some optional questions. The whole questionnaire module was developed in German (for Germany) and has been translated into English (for the UK), French (for France), and Spanish (separate versions for Spain and Venezuela) using the team approach (Harkness, 2003; Harkness, van de Vijver, & Johnson, 2003; Harkness, Villar, & Edwards, 2010).[1] The questionnaire module consists of the following three main elements:

a) A short introductory text that defines the concept of formal education. It aims at generating a consistent interpretation of the subsequent questions by respondents. (This intro can be skipped if only the search tree interface, see section 3.1.3, is implemented.)

b) A question that aims to identify the educational system in which the respondents obtained their highest educational qualification. The educational systems are defined in the database (see section 3.1.2).[2]

c) The actual question on the highest educational qualification obtained, including a number of instructions. The response options are, depending on the educational system mentioned previously, dynamically fed in from the database (see section 3.1.2) via the interfaces (see section 3.1.3).

Additionally, the module offers some optional questions asking for:

d) The year in which the highest educational qualification was obtained.

e) The number of years the respondent has spent in formal education in total to obtain a direct measure of the actual years of education.

f) The questionnaire module accounts for the possibility that surveys may want to ask respondents for more than their highest educational qualification: For respondents, it is not always easy to tell which one of their qualifications is the highest, or their perception may not match the logic of the classification in which the data will eventually be coded. Therefore, the question module contains optional questions and routing instructions for repeating (or "looping") questions b and c. The wording of these questions is then adjusted by not referring to the highest qualification any more from the second loop onwards.

For CAPI surveys, the CAMCES questionnaire module includes specific instructions for interviewers: They have to turn around the computer to allow respondents to report their educational qualification themselves by using the CAMCES interfaces (meaning a switch to CASI mode).

The question module was translated into Dutch,[3] English, French and Spanish using the team approach standards (TRAPD – Translation, Review, Adjudication, Pretesting, and Documentation) as described in for example the CSDI Cross-Cultural Survey Guidelines (Survey Research Center, 2016). The full questionnaire was translated independently by two professional translators certified in both the source language and the target language (their mother

---

[1]   The final CAMCES questionnaire module can be downloaded in these languages at
https://www.surveycodings.org/education/question-module-measuring-educational-attainment.

[2]   A major challenge that is difficult to solve is to have this question referring to the highest qualification preceding the actual question about the highest qualification, because it assumes that respondents already have an idea of their highest qualification in mind.

[3]   The final changes after the CAWI pretest described in section 3.2.6 were not yet implemented in the Dutch version, which is thus not yet available for download.

tongue). The translation process was documented using a Translation and Verification Follow-Up Form (TVFF). Both translators and at least one reviewer/adjudicator were involved in the review of the draft translations where all parties discussed the source text and agreed on a pre-final translation. This version was then verified by the translators and reviewers in order to produce another pre-final translation of the questionnaire. The whole translation process was closely supervised by at least one substantive expert who was involved in all reviews and provided input on the measurement goals.

We recommend following the team approach standards (TRAPD) for the translation and adaptation to any other language by future users of the CAMCES tool. We also suggest pre-testing the questionnaire for any educational context not previously tested in the CAMCES project.

3.1.2   The database

The second element of the CAMCES tool is the underlying database. The database consists of three major parts:

(a) It lists the educational systems, usually corresponding to countries. In multi-lingual countries like Belgium or Switzerland, the list further differentiates between the languages in which the educational qualification could have been obtained. In countries where educational systems differ markedly by region, such as Scotland compared to the rest of the United Kingdom, the list is further broken down by region. This ensures that only relevant response options are shown to respondents.

(b) The central part of the database is a table listing educational qualifications in the relevant languages, with detailed identifying codes. If information is available on outdated qualifications, they are covered in the database, too. In addition, alternative expressions are included when several terms – e.g. official and colloquial – are used for the same qualification. The structure of the search tree (see section 3.1.3) is also defined in this part of the database. As of October 2017, the CAMCES database covers nearly all European educational systems, some neighboring countries, and the countries of origin of the largest migrant and current refugee groups in Germany. It contains nearly 2100 unique educational qualifications and more than 1000 alternative expressions.

(c) Each educational qualification is linked with detailed codes for the International Standard Classification of Education (ISCED). The following versions of ISCED are covered: ISCED 1997 (OECD, 1999; UNESCO-UIS, 1997 [2006]), ISCED 2011 (UNESCO-UIS, 2012; UNESCO-UIS, European Union, & OECD, 2015) and an alternative version of ISCED 2011, developed by the CAMCES team[4].

For countries that are not (yet) covered in the database and where low numbers of immigrant groups are expected, we suggest using an open-ended question for respondents to name their qualification and using the optional question on years of education as a second indicator of educational attainment.

---

[4] The differences compared to the official ISCED codes are that, first, we developed valid codes for educational qualifications not documented in the official ISCED mappings (e.g. for outdated qualifications). Second, when we had reasons to doubt the official ISCED code, e.g. after consultations with country experts (see also chapters in Schneider, 2008), we determined an unofficial code. The alternative ISCED codes are intended to improve cross-national comparability compared with official ISCED codes. Most of the codes are, however, identical to the official ISCED codes.

3.1.3   The interfaces

The final "ingredient" of the CAMCES tool are the survey interfaces that make the database accessible to respondents while completing the questionnaire.[5] One can search the database either by typing (text-string-matching) or by using a structured list. The respective user interfaces for these tasks are the "combination box" and the "search tree". By using these interfaces, more response categories can be offered than with a simple showcard.

The combination box allows respondents to dynamically search their educational qualification in the database by typing and then selecting the best-matching result (see Figure 1).

Figure 1: Interface "combination box" of the CAMCES tool version 1.0 (example shows qualifications for Croatia in an English-language survey environment)

If they do not find an adequate match, they can submit the text they have typed. It is thus a combination of a simple text box and a drop-down box as tested by Couper and Zhang (2016), which is why we call it a combination box. The entry part looks like an empty text field as if it was an open-ended question. However, in order to signal to respondents that the text field actually includes a search box, it is marked with a magnifying glass and contains a watermark reading "search" in the language in which implied by the previously selected educational system. Respondents are asked to enter the response into the text box. With each letter they type, the response is matched with the educational qualifications and alternative expressions in the database (text-string-matching). The results retrieved from the database are presented to the respondent below the text field. The number of results narrows further down as the respondent continues typing. Respondents can then select the best match from the resulting list of educational qualifications. If no matching educational qualification is found or if the respondent does not make a selection among the list of results, the typed text is saved as a text string, as it would be for an open-ended question.

The search tree (see Figure 2) is similar to showcards or response lists in standard educational attainment questions. However, a search tree can accommodate a larger number of response options because of its nested structure. For measuring occupations, Tijdens (2014, 2015) uses a three-level search tree. In order to be as manageable for respondents as possible, and because the number of educational qualifications per education system is not as high as the number of occupations in modern societies, we opted for a two-level search tree. The response categories are structured in a way that optimizes the representation of the selected educational system for respondents, rather than following an international education classification structure, which would potentially lead to a non-intuitive representation of educational qualifications for respondents. The first level offers summary terms such as "sec-

---

[5]   For more research on such interfaces, see Couper and Zhang (2016) who focus on lists of prescription drugs and Tijdens (2014, 2015) who focuses on lists of occupations.

ondary education" or "university degree". When respondents click on a first-level entry, the search tree expands and reveals the different qualifications within this group of qualifications (for example Bachelor's, Master's and PhD degrees within the university degrees section).

The search tree is used as a fallback option for respondents who do not make a valid selection using the combination box. This can happen, firstly, if respondents leave the combination box empty (item nonresponse); secondly, if the text entered does not generate any matches with the database, or thirdly, if respondents enter text that generates matches, but do not select any of the matches offered.

If the script of the language of the selected educational system does not correspond to the script of the keyboard used during the survey, for instance, when interviewing Russian or Greek migrants in a CAPI survey in Germany, the search tree interface is used. For this purpose, the database includes, in the table identifying educational systems, also the scripts used in the respective countries and languages. This ensures that only the search tree interface is shown if the script used in the educational system deviates from the one used in the questionnaire.

Figure 2: Interface "search tree" of the CAMCES tool version 1.0 (example shows qualifications for the Russian Federation in an English-language survey environment)

When respondents select a qualification from the combination box or from the search tree, the respective detailed qualification code is saved in the survey data set. After data collection, this detailed qualification code can be recoded to education classifications such as ISCED (see section 3.1.2) via syntax files provided with the CAMCES tool.[6]

For the CAPI version, the interfaces were programmed as a separate piece of software. This piece of software can be loaded by the CAPI software and the resulting code can be fed back to the CAPI software. For this purpose, the identification codes of the relevant educational system resulting from the respective questionnaire item (see element b) in section 3.1.1) is transmitted to the software. The detailed qualification code as well as information about which interface has generated the code is transmitted back to the CAPI software. The CAMCES tool also allows survey agencies to specify the question text, instruction text, colors etc., and generates log files with paradata such as time stamps for each respondent. The relevant files are available for download at https://www.surveycodings.org/education/implementation-capi-surveys

For the web survey version, the interfaces were programmed using JavaScript. Information on implementation in web surveys is available at https://www.surveycodings.org/survey-implementation/implementation-web-surveys-cawi-mode.

---

[6] See https://www.surveycodings.org/education/classification-educational-qualifications.

3.2 Pretesting and pilot studies

We extensively tested successive versions of the CAMCES tool by various methods, using both general population samples as well as migrant samples, both personal interviewing as well as web interviewing techniques, and using quantitative as well as qualitative methods. We also conducted comprehensive desk testing of each version before using it in the empirical studies reported here. This mix of methods was intended to ensure that we capture most bugs and problems and could optimize the tools as much as possible before publication. The reports on these tests are available for download at https://www.surveycodings.org/education/pretesting-and-piloting.

3.2.1 Exploratory survey experiment in GESIS Panel Pilot

At the very start of the project, we ran two survey experiments in the GESIS Panel Pilot survey to investigate effects of different instructions and response option designs (without database or programmed interfaces).

3.2.2 CAPI interface experiment in the SOEP Innovation Sample

One of the junior researchers on the project applied for a module in the SOEP Innovation Sample 2014 and was successful in the competition. This was a very useful extension of the empirical studies initially planned for the project, since it involved a survey experiment comparing the combination box, search tree and a simple long list item. However, it also substantially increased the workload.

This study revealed that in terms of response times, the combination box is the preferred interface. However, the combination box also involves a higher amount of not automatically coded data, thus requiring post coding. If the main aim in a survey is to avoid post-coding, the search tree is better than the long list question because it is faster to navigate. The design and results are reported in detail in a PhD thesis chapter (Herzing, 2018).

3.2.3 Focus groups and quantitative pretest with TNS Infratest

In order to prepare the CAPI pilot study (see section 3.2.4), we conducted a quantitative pretest and two focus groups with experienced SOEP interviewers (one before and one after the pretest) with our cooperation partner TNS Infratest[7] in 2014 (Schmidt, 2014, 2014; Schmidt & Maier, 2014). The quantitative pretest interviewed a quota sample of 617 migrants from Turkey, Russia/the former Soviet Union, Italy, Greece, Croatia, Kosovo, Poland and Romania who arrived in Germany after 1995, mimicking the sample M1 of the IAB-SOEP Migration Sample. The pretest consisted in face-to-face interviews, again trying to anticipate the IAB-SOEP Migration Sample survey design. The focus groups were highly supportive of the CAMCES approach to measuring education, and the quantitative pretest showed the general technical feasibility of the questionnaire module, software and database, which were improved using the feedback from both focus groups and quantitative pretest.

3.2.4 CAPI pilot study: The IAB-SOEP Migration Sample

The IAB-SOEP Migration Sample is a special sample of the German Socio-Economic Panel (SOEP) initiated by the German Institute for Economic Research (DIW) and the Institute for Employment Research (IAB) (Brücker et al., 2014). The IAB-SOEP Migration Sample measures educational attainment of migrants using generic response options that are not adapted to the origin of the respondent. This approach to measuring education may be problematic because the meaning of education-related terms is highly contextual. The IAB-SOEP Migration Sample was thus a prime candidate survey to benefit from the CAMCES tool, and GESIS and the DIW (hosting the SOEP) thus entered into cooperation for the duration of the

---

[7] TNS Infratest was renamed to Kantar public in 2016.

CAMCES project. In a forthcoming paper (Schneider, Briceno-Rosas, Ortmanns, & Herzing, 2018), we report the results from this study in detail, which are briefly summarized here. The resulting data were published by the DIW in November 2017 (doi:10.5684/soep.v33).

The results show that the CAMCES database (2015 version) covered most of the countries in which the sampled migrants had been educated, even though the sample possessed a highly diverse educational background. However, the 2016 sample included quite a number of origin countries different from the ones the database was initially prepared for. The results also show that looping the questionnaire module on educational qualifications at least once is important because – in the absence of a showcard with hierarchically ordered response options – respondents do not always report their *highest* qualification when asked to do so. The main issue that we have identified in the implementation of the CAMCES tool in the IAB-SOEP Migration Sample is the higher item nonresponse and the higher proportion of non-coded entries, both especially affecting the combination box, when compared with the standard measure. We presented further analyses of respondents' entries in the combination box and concluded that the language switch was a major obstacle to responses being automatically coded. It may thus be advisable to only use the search tree interface in surveys of migrants. We also suspect that interviewers, trying to speed up the interview, affected the implementation of the combination box, especially with regard to the high item nonresponse.

The analysis of the consistency between the standard measure and the CAMCES measure shows a substantial overlap, but also allowed us to identify cases of under- and overestimation of educational attainment in either measure. The substantial degree of ambiguity in generic response options such as "graduated from higher-level secondary school" or "extended apprenticeship at a company" and their translations into various languages may lead to substantial inconsistencies between both measures. The combination box of the CAMCES tool in turn may lead to underreporting if respondents have only schooling, but not other types and levels of education, in mind. Of course, if both instruments arrived at the same result, there would be no reason to invest into a complex tool like CAMCES, and we think that the results reported here give more credibility to the CAMCES approach than the generic approach, especially as regards the search tree. We expect the CAMCES measure to be more precise, and classification of specific foreign qualifications in ISCED 2011 to lead to more valid ISCED codes than "guessing" at the ISCED code that would best represent what respondents have in mind when reacting to the IAB-SOEP response options (in various languages!). Generally, the CAMCES measure then appears to produce more cross-nationally comparable data.

3.2.5 Cognitive pretesting in Germany and Venezuela

Cognitive interviewing enable researchers to test the questionnaire by identifying issues during the process of answering the questionnaire, for example, the comprehension of the questions or the answer process. This type of study was highly relevant for the testing the tool because it allowed to understand the underlying process of answer the question and to identify the issues that case measurement errors. The characteristics of the CAMCES tool, especially its interfaces and its international scope, require to have tailored cognitive interviews that would focus specifically on the topic that affect the quality of measure in this type of questions. The sample required as much diversity as possible in the respondents, including citizens, migrants, and different levels of education, in order to cover the diversity of backgrounds of respondents in social surveys. The German and a Spanish version of the CAMCES tool were thus tested in cognitive interviews in 2015 and 2016. In contrast to the quantitative pretest and focus groups with interviewers, these studies focused on respondents' experiences when using the CAMCES questionnaire and especially the interfaces.

In Germany, the sample consisted of 31 individuals, 15 of which had German educational qualifications and 16 foreign qualifications. This test was thus designed specifically for assessing the usability of the tools for surveys of migrants. The German subsample used quotas for representing different age groups and education levels. In addition to standard cognitive interviewing methods such as conditional probing and paraphrasing, we employed eye

tracking to monitor respondents' interaction with the questionnaire items and coding tools. A report is available (Lenzner et al., 2015), and one dissertation chapter (Herzing & Neuert, 2018) also focuses on this study.

The questions were generally understood as intended, e.g. continuing training and incomplete studies were not regarded as qualifications from formal education by respondents. The pretest revealed some difficulties of forming a judgement on the highest qualification on the side of a minority of respondents when using the combination box (e.g. in case a respondent holds two qualifications at the same level, or whether the Ph.D. counts), which is due to the lacking guidance by a fixed response list. It also showed that the visual design of the combination box interface needed to make clearer to respondents that they would in fact search a database when providing their response. This was subsequently implemented by adding a magnifying glass icon and the watermark reading "Search" (in the required language) to the text field. Furthermore, an instruction was added to not report the occupation or field of study, which many test persons tended to do, and to select the best match from the list of results (rather than providing a more detailed response). An existing instruction that was perceived as too complex was dropped. The search tree interface was easy to use for almost all respondents.

In Venezuela (Briceno-Rosas, del Busto, & Taoli, 2016), the sample consisted of 33 individuals without migration background (i.e. all were educated in Venezuela). While no eye-tracking was used, screens, audio and video were captured for later analysis, and test persons were asked to think aloud while filling in the questionnaire. This pretest had the goal to check whether the developed questions and interfaces also work in a cultural context very different from the German one, where the tool was initially developed. Also, some changes to the tool made as a result of the pretest in Germany (improved instructions and visual cues) could be cognitively tested in this study.

The translation of the questionnaire worked very well and test persons understood the questions as intended. A new introduction helped respondents distinguishing what was relevant and what not. Only the term 'vocational training' was, due to its lesser importance in Venezuela, problematic, and the translation was subsequently improved. A new routing aiming at shortening interview duration in general population (as opposed to migrant) samples however posed the risk of overreporting 'no qualification', which also occurred in the web survey pilot study (see section 3.2.6) that ran in parallel using the same routing. The improvement of instructions after the cognitive pretest in Germany obviously was 'too much' and they were cut again after this pretest, focusing on the most important ones. The lower experience with test persons in Venezuela with computer-assisted interviews however also highlighted difficulties in dealing with new checkbox elements implemented in the questionnaire version used for this study. Again, the search tree appeared more user-friendly than the combination box, also because the combination box motivates respondents to report more detail than required for a survey, and the search tree helps respondents to retrieve information from their memory. A disadvantage of the search tree lies in potential ordering effects, and some test persons indeed selected the first qualification on the list that they had, rather than the highest one.

3.2.6 Web survey pilot study

Finally, the CAMCES tools were implemented in the Dutch LISS (Longitudinal Internet Studies for the Social sciences) Panel, run by CentERdata, Tilburg University in 2016. The LISS panel is a general population sample, thus including few migrants. There was a random selection of members of the LISS panel of 16 years of age or older, with a total case number of n=818, of which 82% responded to the survey. We included a split ballot experiment to compare two routings, one optimized for general population samples and the one previously used for the migrant sample. While the optimized routing worked much faster, it leads to overreporting of 'no qualification' (see results from cognitive pretesting in Venezuela above, using the same routing). This may be a specific problem for web panel surveys, where respondents might be motivated to use shortcuts to faster complete the questionnaire.

The combination box was again more affected by item nonresponse than the standard LISS-Panel education items and the search tree, even though considerably less so in this study than the IAB-SOEP Migration Sample. Also, a substantial proportion of respondents did not receive a valid qualification code in the combination box. Almost all of these respondents then generated a valid code using the search tree, which was like before implemented as a fallback in this study. In terms of consistency of measurement, when disregarding the misreported 'no qualification' responses, the CAMCES and standard LISS-Panel measure agree in 80% of cases, which is close to what Herzing (2018) reported for the SOEP Innovation sample, and what is realistic in terms of reliability (see Porst & Zeifang, 1987).

As a result of this study and the pretest in Venezuela, the question routing was changed again, so that the main routing question was simplified substantially to shorten interview duration, and only respondents with foreign qualifications get to select from the full list of countries. For studies of migrants, it is recommended to use the search tree only, given the complexity of switching languages during the interview.

Table 1: Questionnaire, tool and database versions used in empirical studies

Version	Language	Date	Purpose
0.1	de	09/2013	Survey experiment on instruction wording in GESIS Panel Pilot
0.2	de	11/2013	Survey experiment on response options in GESIS Panel Pilot
0.3	de	2014	Focus group with CAPI interviewers
0.4	de	2014	Quantitative pretest using quota immigrant sample
0.5	de	2014	Survey experiment in SOEP Innovation Sample
0.6	de	2015, 2016	CAPI Pilot in IAB/SOEP Migration Sample (M1/M2)
0.7	de	2015	Cognitive pretesting, Germany
0.8	es	2016	Cognitive pretesting, Venezuela
0.9	nl	2016	CAWI Pilot in CentERdata's LISS Panel

4   Dissemination and future development

The CAMCES tool and its various components are distributed free of charge at https://www.surveycodings.org/education. It is also open source, to be maximally accessible and useful to the scientific community.[8] The education section of the surveycodings web portal provides

- information on the CAMCES questionnaire module and the questionnaire in French, English, German and Spanish as Microsoft Word documents for download,
- a demo questionnaire in the same four languages running on the web showcasing the integration of the interfaces in a web survey,
- a live database search providing both interfaces with access to the most recent database, including access to classification codes,
- information on the empirical studies reported in section 3.2 with respective reports available for downloading,
- information on the classification of education and the classification codes included in the CAMCES database, and
- information on the implementation in CAPI surveys and the 2015 version, which was implemented in the IAB-SOEP Migration Samples 2015 and 2016, for download. (Information on web survey implementation is available at https://www.surveycodings.org/survey-implementation/implementation-web-surveys-cawi-mode, since it applies to other instruments provided by the surveycodings web portal, too.)

---

[8]   We are investigating where to best archive and publish the source code (e.g. https://git.gesis.org).

For surveys not (yet) using the CAMCES tool, the underlying database will be a rich resource supporting more consistency in ex-post coding of educational qualifications (e.g. as a result of open responses by respondents mentioning 'other' educational qualifications) across countries.

More countries will be added to the CAMCES database within the context of the project "Synergies for Europe's Research Infrastructures in the Social Sciences" (SERISS, see section 6.2.1 in the appendix and http://www.seriss.eu/), which runs until mid-2019, Work Package 8 "A coding module for socio-economic survey questions". This development will focus on the search tree. The surveycodings web portal will also be further developed, e.g. by adding an overview table to show which instruments are available for which countries, and by improving documentation for users.

One PhD thesis partially based on CAMCES has recently been submitted, one project is expected to be completed in late 2018, and another one is in the early stages (see section 6.3 in the appendix). The CAMCES project and results from the various empirical studies conducted for the project were presented at numerous international conferences and published in peer-reviewed journals (for a list, see sections 6.4, 6.5 and 6.6 in the appendix). We hope to continue working with the data generated by the CAMCES project, e.g. by exploiting them for ongoing work on dissertations and further publications in peer-reviewed journals.

The possibilities offered by the CAMCES tool are promising for the IAB-SOEP Migration Sample and other cross-cultural surveys. An updated version of the tool has been implemented in the IAB-SOEP Migration Sample 2017 (and will be implemented again in 2018) and in the BAMF-IAB-SOEP Survey of Refugees (Brücker, Rother, & Schupp, 2017)[9]. Moreover, the CAMCES database is used in a survey of the project "ReGES – Refugees in the German Educational System" conducted at the Leibniz Institute for Educational Trajectories (LifBi)[10]. The 2017 IAB-SOEP Migration Sample data will show whether the improvements of the combination box interface reduced the number of responses given in German to a substantially lower level – if not, the combination box likely has to be regarded as unsuitable for migrant surveys in which respondents are not interviewed in their language of origin.

We argue that the CAMCES tool, especially the search tree interface, represents a promising addition or even an alternative for the measurement of educational attainment as it improves the precision, validity and cross-national comparability of the measurement. It is most effective and efficient to use in cross-national surveys using one survey delivery platform, be in on the web or offline. To date, it was impossible to test the tools in a major cross-national survey, but we have received expressions of interest from a few such projects.

In the case of migrant surveys, while a generic measurement instrument is easier to implement than an instrument providing context-sensitive education response options relative to the country in which the respondent was educated, supposedly universal but abstract response options are open to interpretation. The CAMCES tool, in contrast, measures foreign qualifications directly. By making country-specific response options available for the education question in cross-cultural surveys, the measurements generated by the CAMCES tool are far more detailed and specific than measurement using generic response categories that can be interpreted differently. However, this increased precision comes at the price of more complex preparation of the CAPI system and taking slightly more interview time, which is probably only justifiable for studies of migrants rather than general population samples.

By means of more detailed measurement, it is possible to conduct a more precise mapping to cross-nationally comparative education classifications. Coding detailed education data to such classifications using a freely available database allows the comparison of qualifications across educational systems without the efforts and costs of ex-ante harmonization or further post-coding. It also avoids post-coding errors when processing the data, and reduces the

---

[9] BAMF is the Federal Office of Migration and Refugees, whose Research Center on Migration, Integration, and Asylum has entered into a cooperation with IAB and SOEP to establish a new refugee sample.

[10] For further details, see https://www.lifbi.de/en-us/furtherstudies/reges.aspx.

efforts related to documentation of post-coding. It can be used by many surveys, which then do not need to develop their own showcards any more for the education question, which is especially difficult when surveying migrants educated in various educational systems.

5 References

Allen, J., & van der Velden, R. (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, *53*(3), 434–452. https://doi.org/10.1093/oep/53.3.434

Bol, T., & Van De Werfhorst, H. G. (2011). Signals and closure by degrees: The education effect across 15 European countries. *Research in Social Stratification and Mobility*, *29*(1), 119–132. https://doi.org/10.1016/j.rssm.2010.12.002

Breen, R., Luijkx, R., Müller, W., & Pollak, R. (2009). Nonpersistent inequality in educational attainment: Evidence from eight European countries. *American Journal of Sociology*, *114*(5), 1475–1521.

Briceno-Rosas, R., del Busto, K., & Taoli, A. (2016). *Report on Cognitive Interview Study in Venezuela for the Project "Computer-assisted Measurement and Coding of Educational Qualifications in Surveys (CAMCES)" at GESIS – Leibniz Institute for the Social Sciences*. Caracas: LACSO – Laboratorio de Ciencias Sociales. Retrieved from https://www.surveycodings.org/sites/default/files/files/Report CAMCES Cognitive Interviews Venezuela_final.pdf

Couper, M. P., Tourangeau, R., Conrad, F. G., & Crawford, S. D. (2004). What they see is what we get: Response options for web surveys. *Social Science Computer Review*, *22*(1), 111–127. https://doi.org/10.1177/0894439303256555

Elias, P. (2012). Cascot: Computer Assisted Structured Coding Tool. Retrieved from https://warwick.ac.uk/fac/soc/ier/software/cascot/

Herzing, J. M. E. (2018). Investigating alternative interface designs for long-list questions. In *The impact of technological change on survey nonresponse and measurement. Inauguraldissertation zur Erlangung des akademischen Grades einer Doktorin der Sozialwissenschaften der Universität Mannheim* (pp. 77–105). Mannheim.

Herzing, J. M. E., & Neuert, C. (2018). How respondents use combo boxes? An evaluation. In *The impact of technological change on survey nonresponse and measurement. Inauguraldissertation zur Erlangung des akademischen Grades einer Doktorin der Sozialwissenschaften der Universität Mannheim.* (pp. 107–136). Mannheim: University of Mannheim.

Kogan, I., Noelke, C., & Gebel, M. (2011). *Making the transition: Education and labor market entry in Central and Eastern Europe*. Stanford University Press.

Kristen, C., Edele, A., Kalter, F., Kogan, I., Schulz, B., Stanat, P., & Will, G. (2011). The education of migrants and their children across the life course. *Zeitschrift Für Erziehungswissenschaft*, *14*(0), 121–137. https://doi.org/10.1007/s11618-011-0194-3

Kristen, C., & Granato, N. (2007). The educational attainment of the second generation in Germany: Social origins and ethnic inequality. *Ethnicities*, *7*(3), 343–366. https://doi.org/10.1177/1468796807080233

Lenzner, T., Neuert, C., Otto, W., Landrock, U., Adams, F., Disch, K., … Menold, N. (2015). *Computer-Assisted Measurement and Coding of Educational Qualifications in Surveys (CAMCES): Kognitiver Pretest*. Mannheim. Retrieved from http://pretest.gesis.org/pdf/ProjektBericht/Pretest-Projektbericht-15-05.pdf

Levels, M., van der Velden, R. K. W., & Di Stasio, V. (2014). From school to fitting work: How education-to-job matching of European school leavers is related to educational system characteristics. *Acta Sociologica*, *57*(4), 341–361. https://doi.org/10.1177/0001699314552807

Müller, W., & Gangl, M. (2003). *Transitions from education to work in {Europe}: The integration of youth into {EU} labour markets. Transitions from education to work in {europe}: The integration of youth into {eu} labour markets*. Oxford: Oxford University

Press.

Ortmanns, V., & Schneider, S. L. (2016a). Can we assess survey representativeness of cross-national surveys using the education variable? *Survey Research Methods*, *10*(3), 189–210. https://doi.org/http://dx.doi.org/10.18148/srm/2016.v10i3.6608

Ortmanns, V., & Schneider, S. L. (2016b). Harmonization still failing? Inconsistency of education variables in cross-national public opinion surveys. *International Journal of Public Opinion Research*, *28*(4), 562–582. https://doi.org/https://doi.org/10.1093/ijpor/edv025

Porst, R., & Zeifang, K. (1987). A Description of the German General Social Survey Test-Retest Study and a Report on the Stabilities of the Sociodemographic Variables. *Sociological Methods & Research*, *15*(3), 177–218. https://doi.org/10.1177/0049124187015003002

Schmidt, S. (2014). *Pretest für das Migrantensample im Sozioökonomischen Panel 2015 für CAMCES: Bericht zum Interviewer-Workshop im August 2014*. München. Retrieved from https://www.surveycodings.org/sites/default/files/content_files/Debriefingbericht für CAMCES_final.pdf

Schmidt, S. (2014). *Pretest für das Migrantensample im Sozioökonomischen Panel 2015 für CAMCES: Feldbericht*. München. Retrieved from https://www.surveycodings.org/sites/default/files/content_files/FeldberichtTNS_CAMCES_final.pdf

Schmidt, S., & Maier, C. (2014). *Pretest für das Migrantensample im Sozioökonomischen Panel 2015 für CAMCES: Bericht zum Interviewerworkshop im April 2014 – Erste Bewertung des Fragebogens und der Datenbank vor der Feldphase im April 2014*. München. Retrieved from https://www.surveycodings.org/sites/default/files/content_files/Interviewerworkshop_Bericht_final.pdf

Schneider, S. L. (2008). Suggestions for the cross-national measurement of educational attainment: Refining the ISCED-97 and improving data collection and coding procedures. In S. L. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 311–330). Mannheim: MZES. https://doi.org/10.13140/RG.2.1.4614.9526

Schneider, S. L., Briceno-Rosas, R., Ortmanns, V., & Herzing, J. (2018). Measuring migrants' educational attainment: the CAMCES tool in the IAB-SOEP Migration Sample (forthcoming). In D. Behr (Ed.), *Surveying the migrant population: Consideration of linguistic and cultural aspects (working title)*. Mannheim: GESIS – Leibniz Institute for the Social Sciences.

Shavit, Y., & Blossfeld, H.-P. (Eds.). (1993). *Persistent inequality: Changing educational attainment in thirteen countries*. Boulder (CO), Oxford: Westview.

Survey Research Center. (2016). *Guidelines for Best Practice in Cross-Cultural Surveys*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. Retrieved from http://www.ccsg.isr.umich.edu

Tijdens, K., & van Klaveren, M. (2012). A skill mismatch for migrant workers? Evidence from WageIndicator survey data. In B. Galgóczi, J. Leschke, & A. Watt (Eds.), *EU Labour Migration in Troubled Times: Skills Mismatch, Return and Policy Responses*. Ashgate: Aldershot.

UNESCO Institute for Statistics. (2012). *International Standard Classification of Education - ISCED 2011*. Montreal: UNESCO Institute for Statistics. Retrieved from http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

6 Appendix

6.1 Expert workshops and conference sessions

Schneider, Silke L. 2013. "Computer-assisted measurement and coding of education in surveys (CAMCES) preparatory workshop." GESIS, Mannheim, 25.-26.6.2013 (funded by Web-DataNet)

Ortmanns, Verena, Silke L. Schneider, and Jessica Herzing. 2014. "2nd CAMCES Expert Workshop." GESIS, Mannheim, 30.-31.10.2014.

Briceno-Rosas, Roberto, Silke L. Schneider, Verena Ortmanns, and Jessica Herzing. 2016. "3rd CAMCES Expert Workshop." GESIS, Mannheim, 24.05.2016 - 25.05.2016.

Schneider, Silke L., Roberto Briceno-Rosas, and Verena Ortmanns. 2016. "Session: Innovative tools in computer-assisted survey measurement: opportunities and challenges." 3MC International Conference 2016, Chicago, 26.07.2016.

6.2 Cooperation agreements, contracts or follow-up projects

6.2.1 Synergies for Europe's Research Infrastructures in the Social Sciences

Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS) is a four-year project that aims to strengthen and harmonize social science research across Europe, funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 654221. In SERISS, five international research infrastructures, in which GESIS is involved via the European Social Survey, join forces, e.g. to strengthen cross-national harmonization of data and methodology across Europe, and to exploit the advances in software technology for cost-effective data collection. The CAMCES project found a new home in Work Package 8, "A coding module for socio-economic survey questions". The CAMCES team contributes to this Work Package by adding further countries to the CAMCES database and extending it conceptually by adding fields of education and training. The SERISS project runs until mid 2019. The education database is also part of a follow-up proposal to SERISS currently being prepared.

6.2.2 DIW Berlin, Socio-Economic Panel Study (SOEP)

Before the start of the project, a cooperation agreement was entered between GESIS – Leibniz-Institut für Sozialwissenschaften and the German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung e.V. (DIW Berlin), Sozio-ökonomisches Panel (SOEP), Prof. Dr. Jürgen Schupp). After the project ended, a new agreement was set up in order to implement the CAMCES tool in the new BAMF-IAB-SOEP sample of refugees. For this to work, the CAMCES database had to be extended by the educational qualifications of the countries of origin of the largest refugee groups in Germany.

6.2.3 Refugees in the German Educational System (ReGES)

The study "Refugees in the German Educational System" (ReGES), conducted at the Leibniz Institute for Educational Trajectories (LifBi) in Bamberg, Germany, has funded the extension of the CAMCES database by several countries. The CAMCES project has equipped ReGeS with response options for these countries to implement in their survey in 2017.

6.2.4 Programmers involved in the project

Web portal, online database and CAWI version of the CAMCES tool:
Maurice Martens, CentERdata (https://www.centerdata.nl/en/about-us/maurice-martens)

CAPI version of the CAMCES tool: Gerald Mahlmeister (gmahlmeister@gmx.de)

6.3 Qualifications of junior staff supported by the project

The CAMCES project supported three PhD projects.

1. The first one, by Jessica Herzing, is almost completed (dissertation submitted) and entitled "The impact of technological change on survey nonresponse and measurement." It is supervised by Prof. Dr. Annelies Blom and Prof. Dr. Frauke Kreuter at the University of Mannheim. Assessors: Prof. Dr. Frauke Kreuter and Prof. Dr. Edith de Leeuw. Therein with reference to CAMCES:

    -   Chapter 4: Investigating alternative interface designs for long-list questions
    -   Chapter 5: How respondents use combo boxes? An evaluation

2. The second PhD project, by Verena Ortmanns, formally supervised by the project leader, is very advanced and submission is planned for late 2018. Working title: Challenging harmonization – Assessing the quality of the education variable in cross-national surveys (includes Ortmanns & Schneider, 2016a, 2016b).

3. The third PhD project, by Roberto-Briceno Rosas, who joined the project only in 2015 to replace Jessica Herzing, is in the early stages. Working title: Measuring educational attainment in multicultural surveys - assessing the quality of the individual-level measurement. Supervised by Prof. Dr. Michael Braun. (In the final project year, Roberto started working for another project in parallel in order to secure a longer running contract, effectively reducing time available to work on his PhD project.)

6.4 Publications

Ortmanns, Verena, and Silke L. Schneider. 2015. "Harmonization of Educational Attainment Variables in Cross-national Surveys: The CAMCES-Project." Harmonization: Newsletter on Survey Data, Harmonization in the Social Sciences 1 (2), 5-6.

Ortmanns, Verena, and Silke L. Schneider. 2016. "Can We Assess Representativeness of Cross-National Surveys Using the Education Variable?" Survey Research Methods 10 (3): 189-210. doi: http://dx.doi.org/10.18148/srm/2016.v10i3.6608.

Ortmanns, Verena, and Silke L. Schneider. 2016. "Harmonization Still Failing? Inconsistency of Education Variables in Cross-National Public Opinion Surveys." International Journal of Public Opinion Research 28 (4): 562-582. doi: http://dx.doi.org/10.1093/ijpor/edv025.

Schneider, Silke L. 2018 (Forthcoming). "Das Bildungsniveau von Migrantinnen und Migranten: Herausforderungen in Erfassung und Vergleich." In Handbuch Diagnostische Verfahren für Migrantinnen und Migranten, edited by Débora B. Maehler, Alexandra Shajek, and Heinz Ulrich Brinkmann, Göttingen: Hogrefe.

Schneider, Silke, Roberto Briceno-Rosas, Verena Ortmanns, and Jessica Herzing. forthcoming 2018. "Measuring Migrants' Educational Attainment: The CAMCES Tool in the IAB-SOEP Migration Sample (Forthcoming)." In Surveying the Migrant Population: Consideration of Linguistic and Cultural Aspects (Working Title), edited by Dorothée Behr. GESIS Schriftenreihe. Mannheim: GESIS – Leibniz Institute for the Social Sciences.

6.5 Conference presentations

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2018. "Computer-Assisted Measurement and Coding of Education in Surveys (CAMCES): Experiences from the LISS Panel" GOR 2018, Cologne, 28.2.-2.3.2018.

Ortmanns, Verena. 2017. "Der Einfluss von Umfrage-Charakteristika auf die Qualität der Bildungsvariablen." Gemeinsame Tagung der Arbeitsgemeinschaft Sozialwissenschaftlicher

Institute e.V. (ASI) und der Sektion Methoden der empirischen Sozialforschung der Deutschen Gesellschaft für Soziologie (DGS), 10.11.2017.

Ortmanns, Verena. 2017. "Measuring fields of education and training in cross-national surveys." ESRA 2017: 7th Conference of the European Survey Research Association, Lisbon, 19.07.2017.

Ortmanns, Verena. 2017. "The influence of survey characteristics on the quality of the education variable." 70th Annual Conference of the World Association for Public Opinion Research (WAPOR), Lisbon, 16.07.2017.

Ortmanns, Verena. 2017. "The influence of survey characteristics on the quality of the education variable." International Total Survey Error Workshop (ITSEW), 12.06.2017.

Ortmanns, Verena, Silke L. Schneider, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2017. "Die Messung von Bildung bei Migrantinnen und Migranten der ersten Generation: ein neues adaptives Instrument für Umfragen." LERN 2017, Wissenschaftliche Jahrestagung, Berlin, 26.04.2017.

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2016. "Computer Assisted Measurement and Coding of Education in Surveys (CAMCES): A new set of survey tools." Conference of European Statistics Stakeholders, Budapest, 21.10.2016.

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2016. "Computer Assisted Measurement and Coding of Educational Qualifications in Multicultural Surveys (CAMCES): a new set of survey tools." 3MC International Conference, Chicago, 26.07.2016.

Schneider, Silke L., Roberto Briceno-Rosas, Jessica M. E. Herzing, and Verena Ortmanns. 2016. "Overcoming the shortcomings of long list showcards: Measuring education with an adaptive database lookup." 9th International Conference on Social Science Methodology (RC33), Leicester, 13.09.2016.

Ortmanns, Verena. 2016. "Can survey characteristics help? - Trying to explain inconsistencies in the distributions of the education variable with survey characteristics." 9th International Conference on Social Science Methodology (RC33) , Leicester, 13.09.2016.

Herzing, Jessica M. E. (2015): Comparing Interface Designs of Database Lookups with Traditional Measurements. 6th European Survey Research Association. Reykjavik, Iceland.

Ortmanns, Verena, Silke L. Schneider, Jessica M. E. Herzing, and Roberto Briceno-Rosas. 2015. "Computer-assisted measurement and coding of educational qualifications in surveys (CAMCES)." International Workshop on "Comparative Survey Design and Implementation" (CSDI), Mannheim, 28.03.2015.

Ortmanns, Verena. 2015. "Evaluating the quality of education variables in cross-national surveys." 6th Conference of the European Survey Research Association (ESRA), 14.07.2015.

Schneider, Silke L. 2015. "Measuring education in survey research: an overview." Frühjahrstagung der Methodensektion der Deutschen Gesellschaft für Soziologie, Bremen, 27.03.2015.

Herzing, Jessica M. E., and Silke L. Schneider. 2014. "Easy question, tricky answer: measurement quality of education questions." AAPOR 2014: 69th Annual Conference "Measurement and the Role of Public Opinion in a Democracy".

Ortmanns, Verena, and Silke L. Schneider. 2014. "Quality of education data collected in surveys in Europe: Comparing LFS, SILC, ISSP, EVS and ESS." 67th Annual Conference of the World Association for Public Opinion Research (WAPOR), Nice, 06.09.2014.

Schneider, Silke L., and Verena Ortmanns. 2014. "Quality of education data collected in surveys in Europe: comparing LFS, SILC, ISSP, EVS and ESS." European Conference on Quality in Official Statistics (Q 2014), Vienna, 04.06.2014.

Herzing, Jessica M. E., and Silke L. Schneider. 2014. "Response latencies and data quality." GOR 2014: General Online Research Conference. Cologne.

Ortmanns, Verena. 2013. "Validity of measuring educational attainment: education and attitudes in the European values study 2008." 5th Conference of the European Survey Research Association (ESRA), 16.07.2013.

6.6   Other presentations

Ortmanns, Verena, and Silke L. Schneider. 2017. "Computer Assisted Measurement and Coding of Educational Qualifications in Multicultural Surveys." SERISS Survey Experts Workshop on Harmonised Coding for Socio-Economic Survey Questions, 05.09.2017.

Ortmanns, Verena. 2017. "Identification and explanation of inconsistencies of education variables in cross-national surveys." Amsterdam Institute for Advanced Labour Studies (AIAS), lunch seminar, 07.09.2017.

Schneider, Silke L., Roberto Briceno-Rosas, Jessica M. E. Herzing, and Verena Ortmanns. 2017. "Computer Assisted Measurement and Coding of Educational Qualifcations in Multicultural Surveys (CAMCES): A New Set of Survey Tools." Brown Bag Seminar, Deutsches Institut für Wirtschaftsforschung, 25.01.2017.

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2017. "Computer Assisted Measurement and Coding of Educational Qualifcations in Multicultural Surveys (CAMCES): A new set of survey tools." Symposium on „Surveying the migrant population: Consideration of linguistic and cultural aspects", 14.03.2017.

Ortmanns, Verena, Silke L. Schneider, Roberto Briceno-Rosas, and Jessica Herzing. 2017. "Die Messung von Bildung bei Migrantinnen und Migranten der ersten Generation: ein neues adaptives Instrument für Umfragen." IMIRA Workshop, RKI, Berlin, 09.11.2017.

Schneider, Silke L., and Verena Ortmanns. 2016. "Comparability of education measures across surveys revisited: Errors of observation and non-observation." Guest lecture at the Leibniz Institute for Educational Trajectories (LiFBi), 12.05.2016.

Schneider, Silke L., and Verena Ortmanns. 2016. "Comparability of education measures across surveys revisited: Errors of observation and non-observation." University of Vienna, Department of Methods in the Social Sciences, 27.04.2016.

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2016. "Computer Assisted Measurement and Coding of Educational Qualifications in Multicultural Surveys (CAMCES): a new set of survey tools." Uni Konstanz, Forschungskolloquium "Empirische Sozialforschung", 22.11.2016.

Schneider, Silke L., Verena Ortmanns, Roberto Briceno-Rosas, and Jessica M. E. Herzing. 2016. "Computer Assisted Measurement and Coding of Educational Qualifications in Multicultural Surveys (CAMCES): a new set of survey tools." Vortragsreihe „Refugees in Germany", LIfBi – Leibniz-Institut für Bildungsverläufe e.V., 15.12.2016.

Schneider, Silke L., and Jessica M. E. Herzing. 2013. "Computer-assisted measurement and coding of educational qualifications in surveys." 3rd DASISH WP3 Group Meeting, MEA.

Schneider, Silke L. 2012. "Computer-assisted measurement and coding of educational qualifications in surveys (CAMCES)." 1st DASISH Qualititative Workshop on Digital Resources in International Survey Research.