Coding of Occupations (Version 2.0)

Züll, Cornelia

Erstveröffentlichung / Primary Publication
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Züll, C. (2016). *Coding of Occupations (Version 2.0).* (GESIS Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/gesis-sg_en_019

gesis Leibniz-Institut für Sozialwissenschaften

The Coding of Occupations

Cornelia Züll

December 2016, Version 2.0

Abstract

Occupational data are collected in almost all surveys in the context of the demographic variables. The aim of the present contribution is to provide help with coding these data.

The two main standard categorisation schemes available to social scientists are the German Classification of Occupations 2010 (KldB 2010) and the International Standard Classification of Occupations 2008 (ISCO-08). While the KldB 2010 classifies specific job titles, ISCO classifies occupations. The present contribution begins by outlining the structure of the two schemes. It then describes the additional information that should be taken into account when coding occupations. This is followed by a presentation of the individual coding procedures – manual, semi-automatic, and automatic coding – and the respective software programs. The use of the coded occupation data in analyses is then addressed. Coded occupation data are often used to construct indices of occupational prestige, socio-economic status, or class affiliation, which constitute the basis for further analyses. The International Socio-Economic Index of occupational status (ISEI-08), the Standard International Occupational Prestige Scale (SIOPS), and the International Socio-Economic Classes (ISEC) scale are briefly described. As a general principle, the occupational information that is collected and the way it is coded strongly depends on the respective research question. The contribution concludes by discussing alternatives to the labour-intensive coding of occupations according to the KldB or ISCO.

Citation

Züll, C. (2016). The Coding of Occupations. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_019

1. What possibilities of classifying occupations are there?

In almost all surveys, respondents are asked for occupational details in the context of the demographic variables. Frequently, they are also asked to give details of the occupation of their partner and/or their father and mother. The aim of the present contribution is to offer help with coding these data.

The two main standard categorisation schemes available to social scientists for coding occupational data will be briefly described in what follows.

The first scheme, the German Classification of Occupations 2010 (KldB 2010), was developed by the Federal Employment Agency (Bundesagentur für Arbeit – BA) and has been in force since 1 January 2011 (Bundesagentur für Arbeit, 2011a, 2011b). Developed for the purpose of coding occupations in Germany, this classification scheme currently comprises some 27,000 job titles. However, no occupational prestige or occupational status values can be calculated on the basis of KldB-coded data.

The second scheme, the International Standard Classification of Occupations 2008 (ISCO-08) was developed for application in cross-cultural surveys in order to enable comparative analyses to be carried out. It was published by the International Labour Organization (ILO).[1] In contrast to the KldB 2010, which classifies job titles, ISCO-08 classifies occupations, although the boundaries between the two are blurred.

Both categorisation schemes have predecessor versions (e.g., KldB 1988 and 1992, ISCO-68 and ISCO-88), which will not be addressed here. However, tables showing the correspondence between the older and the newer versions are available in each case.[2]

2. How are occupational data collected in surveys?

Details of the respondent's occupation or job are collected with open-ended questions because the scope of the possible responses is too broad to capture with a closed-ended question and a list of predefined response categories (Züll, 2016). For example, ISCO-08 comprises some 440 categories. It is recommended that details of the respondent's occupation or job be asked for in several steps (Statistisches Bundesamt, 2010). In the German General Social Survey (ALLBUS) 2012, for example, occupational data were collected with the following questions: "What work do you do in your main job?"; "Please describe your work precisely."; and "Does this job, or this work, have a special name?" (Wasmer, 2014). In the European Social Survey (ESS), Round 7, the questions were: "What is/was the name or title of your main job?"; "In your main job, what kind of work do/did you do most of the time?"; and "What training or qualifications are/were needed for the job?".[3]

Collecting details of the respondent's occupation in several steps calls for good interviewer training because, in case of doubt, the interviewer should enquire more closely if the information provided is not clear. For example, information such as *civil servant*, *employee*, or *railway worker* is not sufficient for coding into ISCO as it is not possible to determine what work the respondent actually does. Here, the interviewer must intervene in order to obtain better and clearer results.

[1] http://www.ilo.org/public/english/bureau/stat/isco/index.htm

[2] For KldB, see https://statistik.arbeitsagentur.de/nn_237808/Statischer-Content/Grundlagen/Klassifikation-der-Berufe/KldB2010/Arbeitshilfen/Umsteigeschluessel/Umsteigeschluessel.html; for ISCO, see http://www.harryganzeboom.nl/ISCO08/index.htm)

[3] http://www.europeansocialsurvey.org/docs/round7/fieldwork/united_kingdom/ESS7_questionnaires_GB.pdf

3. How are the KldB 2010 and ISCO-08 structured?

3.1 KldB 2010

The KldB is assigned as a five-digit, hierarchically structured code. The first digit of the code denotes the occupational area (*Berufsbereich*; see Table 1); the second digit denotes the occupational main group (*Berufshauptgruppe*); the third digit denotes the occupational group (*Berufsgruppe*), the fourth digit denotes the occupational sub-group (*Berufsuntergruppe*), and the fifth digit, occupational type (*Berufsgattung*), denotes the requirement level (*Anforderungsniveau*; see Table 2), or degree of complexity, of the occupational activity.

Table 1. Occupational areas in the KldB 2010[4]

1	Occupations in agriculture, forestry, farming, and gardening
2	Occupations in production of raw materials and goods, and manufacturing
3	Occupations in construction, architecture, surveying, and technical building services
4	Occupations in natural sciences, geography, and informatics
5	Occupations in traffic,, logistics, safety and security
6	Occupations in commercial services, trading, sales, the hotel business and tourism
7	Occupations in business organisation, accounting, law, and administration
8	Occupations in health care, the social sector, teaching and education
9	Occupations in philology, literature, humanities, social sciences and economics, media, art, culture and design
0	Armed forces personnel

By way of an example of the codes in the KldB, an excerpt from the description of occupational group 1 is presented in Table 2 below. According to this scheme, *hops picker* would be coded as 11101, *farmer* would be coded as 11102, and *agricultural technician* as 11113.

Table 2: Job titles KldB 2010, occupational group 1 (excerpt)[5]

1	Occupations in agriculture, forestry, farming, and gardening
11	Occupations in agriculture, forestry, and farming
111	Occupations in farming
1110	Occupations in farming (without specialisation)
11101	Occupations in farming (without specialisation) – unskilled/semi-skilled tasks
11102	Occupations in farming (without specialisation) – skilled tasks
11103	Occupations in farming (without specialisation) – complex tasks
11104	Occupations in agriculture (without specialisation) – highly complex tasks
1111	Technical occupations in farming
11113	Technical occupations in farming – complex tasks
11114	Technical occupations in farming – highly complex tasks

[4] See Bundesagentur für Arbeit, 2011a; Bundesagentur für Arbeit, 2011b; https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikation-der-Berufe/KldB2010/Arbeitshilfen/Englische-KldB2010/Generische-Publikationen/Kldb2010-Englisch.xls

[5] See Bundesagentur für Arbeit, 2011b; https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikation-der-Berufe/KldB2010/Arbeitshilfen/Englische-KldB2010/Generische-Publikationen/Kldb2010-Englisch.xls

3.2 ISCO-08

ISCO-08 is assigned as a four-digit code. The first digit classifies the major group (see Table 3). This is followed by the sub-major group (second digit), the minor group (third digit), and the unit group (fourth digit).

Table 3. Major groups in ISCO-08

1	Managers
2	Professionals
3	Technicians and associate professionals
4	Clerical support workers
5	Service and sales workers
6	Skilled agricultural, forestry and fishery workers
7	Craft and related trades workers
8	Plant and machine operators, and assemblers
9	Elementary occupations
0	Armed forces occupations

Job groups are classified under *sub-major group*, comparable jobs in related areas are grouped under *minor group*, and the type of work done – the key aspect of an ISCO classification – is captured under *unit group*. The excerpt from the categorisation scheme in Table 4 below provides an example of coding into ISCO-08.

Table 4. Example of coding into ISCO-08 (excerpt)

2	Professionals
21	Science and engineering professionals
211	Physical and earth science professionals
2111	Physicists and astronomers
2112	Meteorologists
2113	Chemists
2114	Geologists and geophysicists
212	Mathematicians, actuaries and statisticians
213	Life science professionals
2131	Biologists, botanists, zoologists and related professionals
2132	Farming, forestry and fisheries advisers
2133	Environmental protection professionals

3.3 What additional information is needed for coding occupations?

When coding occupations, it is helpful to be able to draw on additional information that enables the occupational activity to be classified as precisely as possible. Without this supplementary information, some details cannot be coded completely, or at all. When coding occupations, it is recommended that use be made of the following variables:

- Status in employment
- Sector
- Self-employment (yes/no)

In Germany, status in employment can be collected by using questions 12–12F of the *Demographische Standards* (Statistisches Bundesamt, 2010). The list of response categories in Question 12 is based on that of the International Classification by Status in Employment. Information collected with this battery of questions includes, for example (a) the respondent's position in the occupational hierarchy, (b) the civil service grade if the respondent is a civil servant (c) the respondent's job autonomy, (d) whether the respondent is self-employed and, if he or she is an employer, the number of employees. For many occupational categories, these attributes are decisive for correctly coding the occupational data into ISCO. A self-employed heating engineer with no employees is coded differently than one who has ten employees (Hoffmeyer-Zlotnik, 2003). For social science research purposes, it may be of interest to distinguish between blue-collar and white-collar workers. An example of an instrument for the measurement of status in employment can be found in the 2008 German General Social Survey (ALLBUS).[6]

A further variable that can be used as additional information for coding occupational data is the sector in which the respondent is employed. For example, in many cases it is important to know whether the respondent is employed in the private sector or in the public service. When coding according to the KldB 2010, many categories are directly classified on the basis of sector (see Tables 1 and 2 above).

The question about self-employment can be used, for example, to distinguish between the owner and the manager of a business.

Variables such as sex, age, and income should not play a role in occupational coding. The use of the educational attainment variable is controversial. In the international context, it is recommended that this variable not be used. However, when an unequivocal classification is not otherwise possible, educational attainment can furnish important information when coding occupations in Germany because – at present, at least – career paths in this country depend greatly on educational qualifications. However, caution is warranted because, depending on the labour market situation, over-qualification is not unusual. Moreover, there is a danger of distorting the correlation between education and occupation.

4. How can occupational data collected with open-ended questions be coded?

4.1 Manual coding

Coding occupations is a demanding and time-intensive activity because in each case the coder must decide anew which category from the long list is the correct one. The coding method used here is content analysis – the method generally employed in the case of open-ended questions (Früh, 2015; Züll, 2016).

In contrast to most other open-ended questions, when coding occupations into ISCO-08 or the KldB 2010 the categorisation scheme has already been defined and coding can begin straight away. Nonetheless, it is essential to lay down some rules in advance (e.g., the definitions and category descriptions to be used; the supplementary information that may be used; the way in which ambiguous or incomplete responses are to be handled; and the way in which multiple responses are to be coded). The exact procedure to be followed when coding occupations was described in detail by Geis and Hoffmeyer-Zlotnik (2000) taking ISCO-88 as an example. Moreover, comprehensive information on the practical implementation of occupational coding can be found in Geis (2011). In a conference paper, Ganzeboom (2010) proposed 20 practical rules for occupational coding that provide further information on the procedure.

[6] http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/allbus/Fragebogen/Listen_2008.pdf, Listenheft

Coding into the KldB 2010 is often easier because, instead of the description of the occupational activity, a complete list of job titles is available that can be used as a basis for coding.

As with every content analysis, the reliability of the coding should be tested after it has been completed. To this end, a sample of the material is independently coded by a second coder and a measure of reliability is computed that can be used as an indicator of the quality of the coding. Several measures of reliability are available, for example simple percent agreement, Cohen's kappa, Scott's pi, and Krippendorff's alpha (Freelon, 2010).

In addition to the coding of occupations by coders, a number of procedures have been developed to automate coding in order to reduce the workload involved.

4.2 Semi-automatic and automatic coding

The content-analysis program TEXTPACK enables occupational details to be coded according to ISCO. Here, coding takes place on the basis of a user-defined dictionary that contains occupations and the corresponding codes. However, the dictionary is currently available only for coding into ISCO-88 and would have to be adapted accordingly. With the help of this dictionary, some 50% of occupations can be automatically coded. TEXTPACK does not offer a manual coding option, so that the remaining occupational details must be coded by hand outside the program (Geis, 2011).

One program that facilitates manual, semi-automatic, and automatic coding is Cascot,[7] the international version of which (Cascot International) enables occupations to be coded according to ISCO-08, for example, in different countries participating in cross-national surveys. The development of the international modules took place within the framework of the EU FP7 project Data Service Infrastructures for the Social Sciences and Humanities (DASISH).[8] Cascot International is currently available in Dutch, English, Finnish, French, German, Italian, Portuguese, Slovak, and Spanish. A Norwegian version is planned. Manual coding is facilitated by displaying lists of suggestions for every occupational description, which can be used for coding purposes. Alternatively, you can select the respective codes from the hierarchically organised list of ISCO categories (first hierarchy level, then the corresponding second level, etc.).

To facilitate coding, you can also use additional information (e.g., self-employed yes/no, or status in employment). You can compile in a separate file the information that you wish to use. When coding an occupation, this information can be accessed with a simple mouse click on a case-specific basis.

In the case of semi-automatic coding, the program first codes the occupational data automatically. Irrespective of the coding mode, a score is calculated for each code. This score represents the level of certainty that the assigned code is the correct one. The score can have values between 0 und 100 (0= coding is incorrect, 100 = coding is correct). Users can themselves specify the minimum score that a code must achieve in order to be accepted. If the score given during automatic coding is below this value, the user is requested to manually code the item in question.

For automatic coding, the program uses lists of words, lists of abbreviations, and rules that can be individually supplemented by the user. Depending on the quality of the input data, some 35-45% of occupations can be automatically coded at present. Non-codable occupational data must then be manually coded.

Schierholz (2014) proposed a procedure for automating the coding of occupational data using the KldB 2010. The method presented in his paper is based on the supervised learning approach – that is, the program "learns" from examples, and after completion of the training phase it is able to generalise from the training data. It does not simply learn the examples, but rather it recognises patterns and

[7] http://www2.warwick.ac.uk/fac/soc/ier/software/cascot/
[8] http://dasish.eu/about_dasish/

regularities in the training data and is thereby also able to assess unknown data. The author tested a number of different approaches and concluded that the best results could be achieved by combining rule-based coding with supervised learning. The prerequisite for coding occupational data according to this approach is a training sample that contains already coded occupational data. The proposed approach uses this training sample as well as available dictionaries in order to assign suitable KldB 2010 codes to the respective occupational data. The quality of coding and the number of codable responses are strongly dependent on the size of this training sample.

4.3 Conclusion

The type of coding that you opt for depends greatly on the material available. However, generally speaking, one can say that the automatic approaches are not worthwhile in the case of small numbers of occupations (<1000 occupations) because the effort involved in developing a dictionary or a training sample far exceeds the yield. Hence, manual coding should definitely be used in this case.

In the case of large numbers of occupations, automatic procedures are worthwhile even if 35%–50% automatically codable occupations does not appear at first to be a lot. Nonetheless, even if only half the occupations can be automatically coded, the time saved is enormous. Moreover, the occupations that are automatically coded are mainly those that are easily codable, which considerably reduces the coders' workload. The automatic coding of easily codable occupations also prevents coders from becoming careless due to the monotony of the coding task.

However, the number of occupations to be coded, and the quality of the coded occupations, also depends greatly on the quality of the responses collected. The more complete and exact the responses are, the better they can be coded (either manually or automatically). Moreover, spelling mistakes and abbreviations, which are especially common in web surveys, lead to significantly worse automatic coding results. Therefore, it is definitely worthwhile subjecting the data to – at least rough – cleaning.

5. How can the coded data be used in the analysis?

Many researchers are confronted with the challenge of measuring individuals' social status. Social status plays a key role as a socio-demographic background variable, for example in the analysis of social mobility, the determinants of voting decisions or other forms of social action. As a rule, ISCO-coded data are used to construct indices of occupational prestige, socio-economic status, and class affiliation, which then constitute the basis of further analyses. These indices cannot be calculated on the basis of KldB 2010. Rather, ISCO-08-coded data are needed. The placement of a person in the social hierarchy is carried out via his or her social status. Social status is determined on the basis of the person's education, occupation, and income. The term *occupational prestige* refers to the esteem assigned to an occupational activity and an occupational position or status.

5.1 Socio-economic status (ISEI-08)

One scale with which socio-economic status can be measured in cross-national comparative surveys is the International Socio-Economic Index (ISEI) of occupational status developed by Ganzeboom (2010). The original index was constructed on the basis of information about the income, education, and occupation of some 74,000 full-time employed men (Ganzeboom, De Graaf, & Treiman, 1992). When developing this scale, it was assumed that each occupational activity called for a certain level of education and was correspondingly remunerated.

ISEI-08 was constructed and validated with the data of the International Social Survey Programme (ISSP). A description of the procedure can be found in Ganzeboom (2010). The SPSS syntax module for the assignment of ISEI scores to ISCO-08 codes can be found on Ganzeboom's website.[9]

5.2 Occupational prestige scale (SIOPS)

The Standard Occupational Prestige Scale (SIOPS) was developed as an instrument for use in cross-national comparative research (Ganzeboom & Treiman, 2003; Treiman, 1977). To construct the original scale, respondents in 55 countries were asked to rate and rank a set of occupational titles with respect to their social prestige. A scale for the measurement of prestige was then constructed from these rankings. It can be used to assign a corresponding prestige value to each ISCO-coded occupation. SIOPS can be used to estimate the occupational prestige hierarchy in every country. An SPSS syntax module for assigning SIOPS scores to ISCO-08 occupational codes can also be found on Ganzeboom's website.[10]

5.3 Nominal class categories (ISEC-08, EGP, and ESEC)

In ISEC (International Socio-Economic Classes), the occupational data coded in ISCO-08 are combined with information on status in employment and grouped into 13 social classes (Table 5). The ISEC classes are based on the original Erikson, Goldthorpe and Portocarero (EGP) scheme (Erikson, Goldthorpe, & Portocarero 1979), which has been further developed several times.

Table 5. ISEC classes following Erikson et al. (1979)

I-a	Higher level professionals	1
I-b	Higher level managers and entrepreneurs	2
II-a	Lower level professionals	3
II-b	Lower level managers	4
III-a	Clerical routine non-manual workers	5
III-b	Sales and service routine non-manual workers	6
IV-a	Small self-employed with employees	7
IV-b	Small self-employed without employers	8
IV-c	Small self-employed in agriculture	9
V	Manual supervisors	10
VI	Skilled manual workers	11
VII-a	Semi- and unskilled manual workers	12
VII-b	Agricultural labourers	13

The European Socio-Economic Classification (ESEC) is an adaptation of the EGP scheme proposed by Rose and Harrison (2007). However, the adaptation consisted mainly in reducing the number of classes from 13 to 11.

Further information on the class categories can be found on Ganzeboom's website.[11]

5.4 Other scales

Depending on the research question, further scales and indices based on occupational data are available. Examples include the general, physical, and psychosocial workload scales (Kroll, 2011). These

[9] http://www.harryganzeboom.nl/isco08/index.htm
[10] http://www.harryganzeboom.nl/isco08/index.htm
[11] http://www.harryganzeboom.nl/isco08/index.htm

scales were developed and validated on the basis of the 2006 *Erwerbstätigenbefragung* (Survey of Employed Persons) for the occupational classifications KldB 1992 and ISCO-88 with the aim of providing easily applicable scales for surveys in which it is not possible to use comprehensive instruments for the measurement of workload.

6. Are there alternatives to time-intensive occupational coding?

As a general principle, the occupational information that is collected and the way it is coded strongly depends on the respective research question. Because the coding of occupations according to the KldB or ISCO is very time-intensive, one can well ask whether complete coding according to KldB 2010 or ISCO-08 is really necessary or whether there are alternatives.

6.1 Use of major groups and sub-major groups

For example, if you need only the occupational main group (KldB) or the major group (ISCO) for your analyses, it suffices to restrict coding to this group. Most aspects of the occupational world can be covered in the analyses with just two or three digits. Examples of studies that exclusively used the occupational main group/major group can be found, for example, in Klaukien et al. (2013). The authors compared respondents in the respective occupational groups in Germany with those in the OECD as a whole. The analysis focused on the proportion of employed persons in each of the occupational groups, the respondents' powers of discretion in their jobs, the possibility of exerting influence in their work environment, and on-the-job learning. Perry, Wiederhold, and Ackermann-Piek (2014) compared a number of existing measures of skill mismatch and a new measure that they developed themselves, all of which are based on the PIAAC data. In their proposed new measure, which they described as an improvement on the existing measures, the authors used ISCO-08 occupational codes – but only the major group (one-digit ISCO) and the sub-major group (two-digit ISCO).

However, it is important to bear in mind that, although this coding initially reduces the workload and the costs, you, or your fellow researchers, miss out on the opportunity of fully availing of the flexibility and possibilities of ISCO-08 coding in future.

6.2 Status in employment as a basis for determining social prestige

In order to be able to calculate the prestige values, you usually need occupational data coded using ISCO. In cases where such occupational codes are not available, Hoffmeyer-Zlotnik (2003) proposed a procedure that uses the socio-demographic variable *status in employment* to calculate social prestige values. The procedure entails recoding this variable into a variable entitled *job autonomy* (1= low to 5= high). This relatively crude placement of the respondent in the social hierarchy is precise enough for many analyses and correlates highly with the variables of social prestige generated via ISCO-88 (Hoffmeyer-Zlotnik, 2003). No results for ISCO-08 are available to date.

6.3 Simplified collection of occupation

If you wish to include socio-economic status in your analysis, Ganzeboom (2005) suggested collecting occupation in a crude way. Taking the ISSP 1987 as an example, he demonstrated that there were only minimal differences between the ISEI constructed on the basis of a crude measurement of occupation and the ISEI constructed on the basis of ISCO-68. Occupation was collected using two questions:

Question 1: Which type of job do you have now?

1 Professional and technical (for example: doctor, teacher, engineer, artist, accountant)
2 Higher administrator (for example: banker, executive in big business, high government official, union official)

3 Clerical (for example: secretary, clerk, office manager, civil servant, bookkeeper)
4 Sales (for example, sales manager, shop owner, shop assistant, insurance agent, buyer)
5 Service (for example: restaurant owner, police officer, waiter, barber, caretaker)
6 Skilled worker (for example: foreman, motor mechanic, printer, tool and die maker, electrician)
7 Semi-skilled worker (for example: bus driver, tannery worker, carpenter, sheet metal worker, baker)
8 Unskilled worker (for example: labourer, porter)
9 Farm (for example: farmer, farm labourer, tractor driver)

Question 2: In your present job, are you self-employed or do you work for someone else?

... Self-employed, own business or farm
... Work for someone else

However, Ganzeboom's proposal is based on data that were coded using ISCO-68. The adaptation of the response categories to – and the validation of the results for – ISCO-08 and ISEI are still outstanding.

6.4 ISCO field coding

Ideally, ISCO coding can be carried out in the field because, nowadays, surveys are frequently computer assisted (CAPI, CATI, or web surveys). Taking ISCO-88 as an example, Hoffmeyer-Zlotnik, Hess, and Geis (2014) made a proposal for such in-field coding. After the first tests of hierarchically organised questions to collect the variables necessary for coding into ISCO, the authors proposed a procedure in which the respondents assign themselves to one of 87 occupations. Although the initial results were promising, there was room for improvement. The instrument has yet to be adapted to ISCO-08.

7. References

Bundesagentur für Arbeit. (2011a). *Klassifikation der Berufe 2010. Definitorischer und beschreibender Teil.* Nuremberg: Bundesagentur für Arbeit. Retrieved from https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-Kldb2010/Generische-Publikationen/KldB2010-Printversion-Band2.pdf.

Bundesagentur für Arbeit. (2011b). *Klassifikation der Berufe 2010. Systematischer und alphabetischer Teil mit Erläuterungen*. Nuremberg: Bundesagentur für Arbeit. Retrieved from https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-Kldb2010/Generische-Publikationen/KldB2010-Printversion-Band1.pdf.

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology, 30*(4), 415-441.

Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science, 5*(1), 20-33.

Früh, W. (2015). *Inhaltsanalyse - Theorie und Praxis* (Vol. 8., revised edition). Konstanz: UVK Medien.

Ganzeboom, H. B. G. (2005). On the cost of being crude: A comparison of detailed and coarse occupational coding in the ISSP 1987 data. In J. H. P. Hoffmeyer-Zlotnik & J. A. Harkness (eds.), *ZUMA Spezial* (Vol. 11). Mannheim: ZUMA.

Ganzeboom, H. B. G. (2010). *A new international socio-economic index (ISEI) of occupational status for the international standard classification of occupation 2008 (ISCO-08) constructed with data from the ISSP 2002-2007.* Paper presented at the Annual Conference of International Social Survey Programme, Lisbon.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 2*(1), 1–56.

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. 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 Publishers.

Geis, A. (2011). *Handbuch für die Berufsvercodung.* Retrieved from http://www.gesis.org/unser-angebot/daten-erheben/berufscodierung/

Geis, A., & Hoffmeyer-Zlotnik, J. H. P. (2000). Stand der Berufscodierung. *ZUMA-Nachrichten, 47*, 103-128.

Hoffmeyer-Zlotnik, J. H. P. (2003). "Stellung im Beruf" als Ersatz für eine Berufsklassifikation zur Ermittlung von sozialem Prestige. *ZUMA-Nachrichten, 53*, 114–127.

Hoffmeyer-Zlotnik, J. H. P., Hess, D., & Geis, A. (2004). Computerunterstützte Vercodung der International Standard Classification of Occupations (ISCO-88). *ZUMA-Nachrichten, 55*, 29–52.

Klaukien, A., Ackermann, D., Helmschrott, S., Rammstedt, B., Solga, H., & Wößmann, L. (2013). Grundlegende Kompetenzen auf dem Arbeitsmarkt. In B. Rammstedt (Ed.), *Grundlegende Kompetenzen Erwachsener im Internationalen Vergleich: Ergebnisse von PIAAC 2012.* Münster: Waxmann.

Kroll, L. E. (2011). Konstruktion und Validierung eines allgemeinen Index für die Arbeitsbelastung in beruflichen Tätigkeiten. *Methods, data, analyses, 5*(1), 63–90.

Perry, A., Wiederhold, S., & Ackermann-Piek, D. (2014). How to measure skill mismatch? New approaches with PIAAC. *methods, data, analyses, 8*(2).

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. doi: 10.1080/14616690701336518

Schierholz, M. (2014). Automating survey coding for occupation. *FDZ-Methodenreport, 10*(10), 1-75. Retrieved from http://fdz.iab.de/de/FDZ_Publications/publication-details.aspx/Publikation/k141027302

Statistisches Bundesamt. (2010). *Statistik und Wissenschaft. Demographische Standards.* Wiesbaden: Statistisches Bundesamt. Retrieved from https://www.destatis.de/DE/Methoden/Demografische RegionaleStandards/DemografischeStandardsInfo.html

Treiman, D. J. (1977). *Occupational prestige in comparative perspective.* New York: Academic Press.

Wasmer, M. (2014). German General Social Survey 2012. English Translation of the German "ALLBUS"-Questionnaire. GESIS Technical Reports, 2014/01.

Züll, C. (2016). Open-ended questions. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_002