Update of the underlying principles of the EurOccupations database: third reporting period: D1a1

Tijdens, K.

Project no.	FP6-028987
Project acronym	*EurOccupations*
Project title	Developing a detailed 7-country occupations database for comparative socio-economic research in the European Union
Instrument:	STREP
Thematic Priority	PRIORITY 7, Research Area 3.2.2.

Update of the underlying principles of the EurOccupations database

Update of deliverable D01a, 2007

Third Reporting Period – D1a1

Project coordinator name:	Erasmus University Rotterdam
Start date of the project:	from 01.05.2006 to 30.04.09 (36 months)
Period covered:	from 01.05.2008 to 30.04.09
Report Preparation Date:	08.06.2009
Author of this report:	Kea Tijdens (Erasmus University, University of Amsterdam)

Table of contents

Glossary

➢ <u>EurOccupations</u> is the EU-FP6 funded project whose scientific objectives aimed at (1) building a freely available web-based database with the 1,500-2,000 most frequent occupations in 8 EU member states, to be used for comparative, multi-country data-collection, using the International Standard Classification of Occupations (ISCO), and (2) testing for 150 occupations the similartity regarding job content, required skill level, and competency profiles across the 8 member states

➢ <u>EurOccupations</u> <u>key list of occupations</u> is a list of 150 occupational titles, take from the source list, and tested for the comparability of job content and skill requirements across countries

➢ <u>ILO</u> is the the International Labour Office (ILO) of the United Nations, that has developed ISCO

➢ <u>ISCO</u> is the International Standard Classification of Occupations, recently updated from ISCO-88 to ISCO-08

➢ <u>Database of Occupations</u> consist of a source list, translations, search tree and cross over tables

➢ <u>Database of Occupations source list</u> is the basic list of almost 1,600 occupational titles in use in the Database of Occupations

➢ <u>WageIndicator web-survey</u> is the multi-country, multilingual, continuous questionnaire on work and wages, posted at country-specific WageIndicator websites in national languages

➢ <u>WageIndicator Foundation</u> is the not-for-profit Foundation that owns the WageIndicator concept; the University of Amsterdam participates in the Board of Directors

Abstract

In contrast to other survey modes, web-surveys allow for using a closed response format for survey questions with a large number of categories, such as car types, camera types, countries, regions, industries, and occupations. In other survey modes, these questions require an open response format, followed by field- or office-coding. Web-surveys allow the use of drop-downs, for example for an alphabetically sorted list of 200+ countries, or for the use of search trees. A search tree is the most appropriate way for self-identification of occupation, which is a key variable in socio-economic research.

This paper summarizes the design principles underlying a Database of Occupations for the measurement of occupations in multi-country web-surveys by means of self-identification. The Database was compiled as part of the FP6 funded EurOccupations project (2006-2009, n° 028987). The paper introduces the use of the occupation variable in datasets, in particular multi-country datasets, and explores the reasons why the Database has been designed. Using a variety of sources, the Database includes almost 1,600 occupational titles, which is judged to be sufficiently detailed for valid-self-identification. The Database consists of a source list, a search tree and translations into the languages of almost 50 countries. It is explained how the source list has taken into account issues such as skill levels, corporate hierarchies, job ladders, managerial and supervisory occupations, craft versus manufacturing occupations, composite occupations and the methods used for translations. The requirements posed by web-surveys, such as readability, alphabetical sorting, coping with the category "Other" are discussed. Finally, the paper details how the Database is used in the worldwide WageIndicator web-survey. For more information see www.euroccupations.org or www.wageindicator.org.

1. Introduction

Occupation is a key variable in socio-economic research, but two problems are associated with the measurement of occupations. The first one is a measurement problem in questionnaires and the second is the lack of international comparability of occupational coding systems. Web-surveys may in part solve both problems, as will be discussed in this paper. Web-surveys are expected to become a major data-collection tool in the social sciences, promising costs-effectiveness, large-scale data, and features such as highly flexible plug-ins. For a survey question on occupation with a closed response format, web-surveys must use of a list of occupations and a search tree for navigating through the list. As part of the EU-FP6 funded EurOccupations project (2006-2009, n° 028987), a Database of Occupations has been designed, including a list of occupations, a search tree and country-specific translations. EurOccupations' scientific objectives 1 to 3 aimed at building a web-based database with the 1,500-2,000 most frequent occupations in 8 EU member states, to be used for comparative, multi-country data-collection, using the International Standard Classification of Occupations (ISCO). The Database will be publicly available, so that it can become a widely applied tool. It is free downloadable at www.EurOccupations.org, and is primarily developed for use in multi-country web-surveys, but may suit other survey-modes or other purposes as well. If the Database of Occupations will be used by multiple data-collectors, it ensures that occupational data is collected in a comparative manner within this survey mode. This will increase comparability across data-collections, be it within or across countries. Additionally, web-surveys allow for large-scale data, making it worthwhile to collect detailed occupational titles.

Thus, the Database was not primarily designed for labour intermediary purposes, for occupational guidance, or for elaborating on skills indicators or task elements. Moreover, EurOccupations did not aim for designing or revising any occupational classification. It aimed at measuring occupations at a detailed level, such that the occupational titles can be recoded into any classification, as long as all units in the classification are reflected in at least one occupational title in the choice-set. This fits into McGuckin's (1991) statement that a statistical system should provide possibilities for generating multiple groupings of data to satisfy multiple objectives, which requires that the basic data are sufficiently detailed to support user needs and are processed and maintained in a fashion that makes the use of a variety of aggregation rules possible.

Since 2007/Q2, a draft version of the Database of Occupations has been used in the continuous, worldwide WageIndicator web-survey, under supervision of the WageIndicator Foundation. Recently, this Foundation has expanded its operations to 50 countries (see www.wageindicator.org and its research lab). The Foundation, in cooperation with its partners, has funded translations for the Database of Occupations for 10 additional languages, apart from the 7 languages present in the EurOccupations project. Until 2008/Q4, more than 170,000 respondents from 33 countries have responded to the survey question about their occupation.

This paper details the design principles used in EurOccupations' Database of Occupations. It is an update of EurOccupations Deliverable D1a (Tijdens 2007c). In section 2, the dynamics in occupational structures and state-of-the art regarding the measurement of occupations are reviewed. Section 3 outlines the methodology used in the design of the source list of occupations. Section 4 details the principles underlying the choices made with regard to the wording of occupations, the translations and the search tree. Section 5 details the principles underlying the choices made with regard to the skill levels, job ladders and corporate hierarchies. Section 6 draws the main conclusions learned from the project.

2. Reviewing occupational dynamics and the measurement of occupations

2.1 The data collection and processing

Many socio-economic surveys include a question such as "What is your occupation?", "What kind of work do you do?" or similar. The wording of this question is important. It should ask about the job title and if possible, about the main tasks and duties of the job, not about occupation. The latter is a term often interpreted by respondents to mean the type of work they were trained for (Hoffmann et al 1995). Additionally, the survey question is a sensitive one, because for many individuals their occupation is a prime identifying characteristic. It identifies social position and a set of social relationships, including those to other jobholders and to the clients in the work setting. Therefore, the phrasing of this question should be checked in case of multi-country surveys.

For the answers to this question either an open or a closed response format is used. In the open response format, respondents report their job titles as they like. This format can be used in all survey modes. In the closed response format, respondents choose from a tick list of occupational titles or groups. This self-identification method can also be used in all survey modes, but the length of the list of occupations varies across the modes, ranging from at most 5 groups in telephone-surveys, at most 50 groups in paper surveys or in face-to-face surveys when using show-cards, to 8,000 titles in web-surveys. In the remainder of this section, the advantages and disadvantages of the open and closed response format will be discussed. Table 2.1 summarizes this discussion.

Table 2.1 The use of the question on occupation in different survey modes

Survey mode	Open response format	Closed response format
Paper	used	used, choice-set max 50
Telephone	used	scarcely used, choice-set max 5
Face-to-face	used	scarcely used, choice-set max 50
Web	used	used, choice-set max 8,000

The closed response format is used for self-identification. In case of a limited choice-set, the respondent is given a choice of highly aggregated occupational groups. This may result in lower data quality, because it is difficult to assure consistency in how respondents fit their own job titles into the highly aggregated categories, introducing aggregation bias (Vries and Ganzeboom 2006). Based on a comparison of three datasets, these authors show that socio-economic status indicators derived from self-identification on a 9-category occupation list are slightly better than those from recoded open response format questions, but that the joint data leads to the best results.

For the survey question on occupation, the open response format requires a recoding effort of the data-collector, either field-coding or office-coding. Field-coding is instant recoding which is advantageous because it allows the interviewer to ask for additional information if needed. Office-coding is recoding at a later point in time and is disadvantageous in budget terms, although this may become less of a problem with advanced automatic coding tools currently being developed (Michiels and Hacking 2004; Elias, Ellison and Jones 2009). For a long time, statistical agencies in the UK, as in some other countries, use software programs for automatic recoding, such as CASCOT and its update CASCOT2000 (Halstead, Elias and Prandy 1993; Jones and Elias no year[1]). Many coun-

---

[1] See http://www2.warwick.ac.uk/fac/soc/ier/publications/software/cascot, accessed 20090401

tries have developed a coding index, which classifies the most frequently mentioned job titles into the proper occupational titles.

The open response format question is advantageous in terms of data quality, because respondents mostly tend to report a detailed job title, as they know it from their employment contract, job classification scheme, collective bargaining agreement, job advertisement, or just from a common understanding in the workplace. For classifying this response a coding index is required, allowing to classify highly disaggregated occupations, such as *Lithographic stone grinder,* or very firm-specific job titles, e.g. *Appls Prog I.* However, a minor share of respondents tends to report in highly aggregated categories, e.g. *Clerical worker* or *Teacher,* or not specific at all, e.g. *Employee of department X, Senior supervisor,* or *Dogsbody.* In case of field-coding, additional survey questions need to be asked for clarification; in case of office-coding these reported job titles have to be classified either in highly aggregated occupational groups or as unidentifiable data. In conclusion, the open response format question may lead to aggregation variation in the data.

Occupational coding is an inexact process (Elias 1997). Based on a meta-analysis of the results from occupational recoding studies, he concludes that agreement rates increase at higher levels of aggregation, that is at 1- or 2-digit levels. At the 3-digit level agreement rates in excess of 75 per cent are hard to obtain. According to the author, this implies that comparisons between countries at a 3-digit ISCO level will be exposed to the low level of reliability associated with occupational classification.

A closed response format in a web-survey, using a database with a choice-set of thousands of occupational titles, allows for obtaining 4- or 5-digit level codes within or across countries. The main advantages are the detailed level of information, the absence of unidentifiable or too aggregate data and the fact that the dataset does not include the 75% level of reliability associated with field- or office-coding, as is summarized in Table 2.2. In case of cross-country comparisons, the choice-set needs to cover multiple languages and for the sake of self-identification high quality translations of the source list are needed (see also section 4).

Table 2.2 Problems related to open versus closed response format for survey questions on occupation

Item	Open response format	Closed response format
Coding index	high quality index needed	no recoding needed
Aggregation level	mixture of detailed and aggregated job titles	no variation if an extended choice-set is used
Aggregation problems	aggregation variation: X% too aggregate answers	aggregation bias if a limited choice-set is used
Missing data	X% unidentifiable answers	n.a.
Recoding costs	expensive recoding	n.a.

2.2 **Occupational classifications and their characteristics**

To facilitate international comparisons, the International Labour Office (ILO) of the United Nations has developed the International Standard of Occupational Classification (ISCO) in 1958, with revisions in 1968, 1988, and 2008.[2] In the 1990's, ILO has undertaken efforts to implement ISCO-88 in many countries, among others by discussing the national data collection and processing procedures (Hoffmann et al 1995; Greenwood 2004). ISCO has become the standard classification in many countries, including most of the EU-27 countries (see Greenwood 2004, for an overview). However, countries with a long tradition in conducting Labour Force Surveys or Censuses, such as the United States, Austria, United Kingdom, Germany, France, and the Netherlands, have continued using their own National Occupational Classifications (NOC). Detailed occupational classifications tend to differ cross-nationally with respect to the level of detail, with respect to specific occupational titles included in the classifications and, with respect to their logic (Ganzeboom and Treiman, 1996). In the eight EurOccupations countries, only the national classification of Poland is based on ISCO-88, while other countries have developed mapping tables NOC -> ISCO, usually to the 3-digit ISCO level (Greenwood 2004; Jacobs, Michon and Tijdens, 2007).

The logic of the ISCO-88 classification and its predecessors are mostly derived from skill requirements at the expense of industry distinctions (ILO 1990; Ganzeboom and Treiman 1996; Greenwood 2004). This principle has been confirmed for ISCO-08 (ILO, 2007). ISCO-88 distinguishes four skill levels, whereby skill levels have been operationalised referring to the International Standard Classification of Education (ISCED), developed by UNESCO (UNESCO 2006, re-edition). It can easily be understood that for a worldwide classification problems arise as skill levels of occupations may vary across countries. During the preparation of the 2008 revision of ISCO, the similarity of occupations raised few discussions, but the major discussions concerned the skill levels assumed with the ISCO codes (Elias and Birch 2006). These discussions are obviously hampered by the fact that no empirical investigation of required skill levels has been undertaken, probably because on a worldwide scale this would be a too costly and time-consuming undertaking. Attempts to harmonize the occupational classifications were hampered by the fact that ISCO does not allow skill levels of occupations to vary across different national contexts (Elias 1997).

The most problematic side of an occupational classification however is the definition of an occupation. An occupation is defined as a set of similar tasks and duties undertaken in the job. However, this concept remains vague, as can be illustrated by the wide variety of occupational units at the lowest level of the classification (see Table 2.3). At the 4-digit level, ISCO-88 distinguishes 390 units and ISCO-08 433 units. Statistics Netherlands counts 1,211 occupational units, Statistics Spain 482 units, Statistics France 484 units, and Statistics Poland 389 units (Jacobs, Michon and Tijdens 2007). The occupational classification in the USA distinguishes 506 units at the 3-digit level.[3] Statistics Belgium identifies 3,367 occupations at the 5-digit level and Statistics Germany does so for 2,287 occupations at the most detailed level. Many statistical agencies usually employ a job title coding index in the national language for recoding reported job titles into the

---

[2]   ISCO-08 was adopted through a resolution of a Tripartite Meeting of Experts on Labour Statistics held in December 2007. This resolution was subsequently endorsed by the Governing Body of the ILO in March 2008.

[3]   See www.nlsinfo.org/nlsy97/docs/97HTML00/97guide/matt1.htm, and www.nlsinfo.org/nlsy97/docs/97HTML00/97guide/matt1codes2002.htm accessed 20070130. Please refer to http://www.bls.census.gov/CPS_OCCUPATION_CLASSIFICATION.htm for document

classification. The ISCO-88 coding index comprises some 5,000 job titles (excluding re-verse titles). The one used by Statistics Netherlands has more than 30,000 job titles, the UK Office of National Statistics has one with almost 25,000 titles, and the 2002 US Stan-dard Occupational Classification (SOC_US) describes 46 characteristics of 12,099 job ti-tles (all not excluding reverse titles).

Table 2.3 The number of occupational groups at the four measurement levels for four occupational classifications.

Digits	Logic	ISCO-08	SOC2002_USA	SBC2002_NL	SOC2000_UK
1-digit	Skill category*	10	24	5	9
2-digit	Job similarity	42	26	43	25
3-digit	Job similarity	131	506	121	81
4-digit	Unit group	433		1,211	353
Database of Occupations	Occupational title	1,500			
-	Coding index	5,000	12,099	33,000	25,000

2.3 **The occupation variable in major datasets**

Multi-country datasets are typically surveyed by national survey agencies with the data merged afterwards (Hoffmann no year). In these cases, the survey operations, the ques-tion formulations or the coding procedures most likely are not harmonized. Therefore the comparability of the resulting statistics may only be partly solved by aggregation of the national data, most likely into a 1-digit, a 2-digit or at best a 3-digit ISCO code. For ex-ample, for most but not all EU member states, Eurostat is able to provide comparative Labour Force Survey (LFS) data with 3-digit ISCO information. The European Community Household Panel holds ISCO 2-digit occupation data. The 1990 and the 1995 European working conditions survey (EWCS) has 1-digit ISCO occupation data only, but the 2000 and 2005 EWCS survey has 2-digit ISCO data. The World Values Survey has occupation data on ISCO 1-, 2-, or 3-digit level, varying across countries.

Elias and McKnight (2001), although noticing improvements in the comparability rating of national occupational distributions over the 1990's, identify several problems in multi-country datasets as composed by Eurostat. They call for harmonization of survey ques-tions, for the adoption of common coding procedures and for a common understanding of the conceptual basis of ISCO, in particular its skill concept.

Regardless comparability is improving, the national occupational data-collection may be-come a problem in the years to come as data collection methods are changing. National statistical offices in EU-member states increasingly rely on administrative sources such as personnel records for their socio-economic data collections, among others attempting to get rid off expensive survey data. Unfortunately, administrative data hardly register oc-cupation data.

3. The methodology

3.1 Introducing the Database of Occupations

With the increased use of the Internet and their relative cost-effective way of surveying the population, web-surveys have become a major data-collection method in the social sciences. Web-surveys can of course cope with an open response format question "What is your occupation?". Yet, unique to web-surveys, they can equally present a long list of items, e.g. car types or occupational titles, for self-identification. A systematic way of navigating through the list is required by either an alphabetically sorted list or a search tree. In case of a list of some 200 countries, it has become common to present an alphabetically sorted list. In case of 1,500 occupational titles, an alphabetically sorted list is not the best solution. Here, a search tree is required. Job-vacancy sites for example employ search trees using a concept of job families. In computer-assisted face-to-face interviews, a list of occupational titles and a search tree can also be used for self-identification, when the interviewer turns the screen to the respondent.

Compared to the open response format question or to the self-identification with highly aggregated occupational groups, the use of a long choice-set and a search tree is advantageous for valid data-collection on occupations. First, the choice-set can be rather long. Because respondents prefer to report their job titles and disaggregated occupational titles are closer to job titles, a long list most likely improves the validity of occupation data. Second, the list of occupations can be designed such that the aggregation level is similar across all occupations, thereby avoiding aggregation variation. Third, compared to recoded NOC-data in merged national datasets, data collected with a multi-country choice-set of occupations plus search tree is expected to provide better comparability of disaggregated units because the same survey methodology is used across countries. Fourth, compared to an open response format question, costly recoding and unidentified data are avoided.

The Database of Occupations has four components, to be discussed later:

- A choice-set of more than 1,500 occupational titles in English, called the source list[4].
- Country-specific translations of these occupational titles.
- A search tree to facilitate respondents' self-identification of their occupations.
- Numerical codes, relating all occupations to the ISCO-08 classification.

3.2 The methodology

In May 2006, when the EurOccupations project started, the initial plan for the database departed from the idea that national statistical agencies would be able to provide labour force survey frequency tables for the most detailed occupational units, including the distribution of gender, education and age groups by occupation. These tables could have been merged, using mapping tables NOC into ISCO-88_4-digit, and from this detailed list the most common occupational units in the member states could be identified. Yet, the

---

[4]   In the EurOccupations project the source list was called the extended list of occupations, in contrast to key list of 150 occupations which were tested with respect to their task-descriptions and skill requirements (see De Ruijter, De Ruijter and Jacobs, 2008).

statistical agencies of most member states involved in EurOccupations could not supply us with tables for detailed 4-digit occupational units (Jacobs, Michon and Tijdens 2007). Even when data had been supplied, the tables would reflect ISCO-88, while ILO meanwhile had taken steps towards the revision into ISCO-08. Moreover, the use of the Database of Occupations in the WageIndicator web-survey opened possibilities for worldwide use. Therefore, comparability with the Standard Occupational Classification of the United States of America became an extra objective. Taken all these reasons into account, the initial strategy was adapted. EurOccupations undertook a stepwise procedure on the drafting of the source list of Database of Occupations:

1. Draft 3 of ISCO-08, published in September 2006 (ILO, 2006) with its 447 occupational units at 4-digit level was taken as the point of departure for the source list of occupations. Next, the occupations additionally specified in this draft were added to the list.

2. The 564 occupations at 3-digit level of the 1991 revised fourth edition of the Standard Occupational Classification of the U.S. Department of Labor were mapped into the ISCO-08 classification, taking into account that SOC_US is not easy recodable into ISCO.[5] Occupations, specified in SOC_US but not in ISCO-08_draft 3, were added to the source list. In January 2009, the 2010 Standard Occupational Classification was published. The classification had been revised substantially. As much as possible at 4-digit level, a cross over table from the source list to the 2010 SOC_US was made, though many Database occupations could only be mapped at the 3-digit level. One of the major differences is the degree of break down within occupational groups. Whereas the 2010 SOC_US breaks down the teacher occupations into very detailed teaching fields, the source list does so to a lesser extent. Whereas the source list breaks down the shop keepers and sales assistants into approx. ten fields, the 2010 SOC_US does not provide a break down for these occupations. For this reason, the cross over table does not include a full match.

3. We went carefully through the detailed Alphabetical index of occupational titles for ISCO-88(COM). Occupations that could be expected to have large numbers of jobholders were added to the source list.

4. Frequently reported occupations in the German and Netherlands data of the WageIndicator web-survey were added to the source list.[6] This step allowed for including new-arising occupations and for identifying occupations that had to be detailed.

5. National occupational classifications, such as the SBC classification from the Netherlands, the SOC (UK), the Belgian VDAB-list, the US-based O*Net occupations database, and the Canadian occupations website www.workfutures.bc.ca were reviewed and occupational titles that could be expected to have large numbers of jobholders were added to the source list.

6. Early 2007, the source list held 1433 occupational titles, in April 2007 published as Deliverable D01b (see Jacobs, Michon and Tijdens 2007). This list was translated into German, French, Spanish, Italian, Polish and Netherlands. Translations were made by professional translators, but all translations were carefully checked by the EurOccupations partners, who all are national labour market experts. Their comments mostly related to occupational boundaries, e.g. two distinct occupa-

---

[5] See www.census.gov/hhes/www/ioindex/occ_a.html for the Coding Index of SOC in the US.

[6] This dataset holds almost 80,000 observations for Germany (Sep/2004-Sep/2006, and more than 160,000 observations for the Netherlands (Sep/2000-Sep/2006). Some 10% of the respondents use the open-ended question for specification after having ticked an occupation in the occupation search tree. These Netherlands data have been checked in September 2006 as part of the Netherlands occupations search tree (Tijdens 2007a). The German dataset has been checked in 2007 and has resulted in additional occupations.

tional titles in the source list were not considered distinct in the country at stake. Consequently, occupational titles were removed from the source list.

7. Since 2007/Q2, the translations of the source list have been implemented in the continuous, worldwide WageIndicator web-survey (Tijdens 2007c). For this purpose the list was translated into the languages of Brazil, China, Denmark, Finland, Hungary, Indonesia, Russia, Sweden, and South Korea. For each language two professional translators performed the translations, which in turn were checked by an editor, and were subsequently checked by national labour market experts. For countries for which translations were available, such as Argentina, India, Mexico, South Africa, and the United States, the translations were checked by national labour market experts with regard to its applicability to the country at stake. In 2008, Czech Republic, Slovakia and Turkey joined, followed by Bulgaria in 2009. Translations for Greece, Romania and other countries, as well as Hindi for India, are foreseen in 2009 or 2010. The translations have been used as a check on our understanding of occupational boundaries and similarities. It occurred that two occupations in the source list had been translated similarly one particularly language. In these cases, the source list was checked if these two occupations were really different from each other. If not, one of them was removed from the source list.

8. In Spring 2008, ILO had published the final ISCO-2008 classification with 433 occupational titles at 4-digit level (ILO 2007). Among others, in the final version a number of occupations was assigned a skill level, different from that in Draft 3. The source list of occupations was adapted to ISCO-08, and was critically reviewed with regard to internal consistency and suitability within the search tree.

9. In Spring 2009, the Database of Occupations was updated taken into account the results of the EurOccupations similarity test of skill requirements and job content of 150 key occupations, selected from the initial list of 1,500 occupations (Ruijter, Ruijter and Veldhoen 2007a; Ruijter et al 2007b).

10. On June, 1, 2009, ILO distributed the ISCO-08 draft definitions, indicating detailed occupational titles for many of the 4-digit ISCO-08 occupations units (Hunter 2009). The Database of Occupations was checked against this list and updated where needed.

11. In Autumn 2009, the updated Database of Occupations will be implemented in the WageIndicator web-survey, which by then will cover more than 50 countries. Before, the WageIndictor team will do its best to have the translations checked by national labour market experts.

For all decisions made during the construction of the source list, the design principles will be detailed in this section. The principles concerning wording, translations and search tree are so in section 4. The principles concerning skill levels, job ladders, blurred occupations and corporate hierarchies are detailed in section 5.

3.3 The Database of Occupations in WageIndicator web-survey

Since 2007/Q2, the list has been implemented in the WageIndicator web-survey. Until 2008/Q4, more than 170,000 respondents from 33 countries have responded to the survey question about their occupation. Respondents' feed-back is important for testing the user-friendliness of the search tree. Over the years, the WageIndicator web-sites have received thousands of emails from web-visitors, but less than one percent concerned feed-back from respondents who could not find their occupation. In those cases, both the codeset and the search path was checked and adapted if necessary. These results indicate that the use of the Database of Occupations is feasible and that it indeed leads to detailed occupational data, comparable across countries.

In the WageIndicator web-survey question on occupation the respondents are asked to tick the title that comes closest, if they cannot find the appropriate occupational title and to use the open response format for specification of the occupational title.[7] The results have been used in step 4, mentioned in the previous section. The results indicate that respondents primarily add specifications to their occupational title, but sometimes occupations were mentioned that were not present in the database.

3.4 **Defining an occupation**

The database should allow for valid self-identification in surveys, in such a way that the vast majority of respondents are able to identify easily their occupational title, taken into account that respondents prefer to tick occupational titles that come closest to their job titles. For the purpose of the Database of Occupations, the definition is

> *"an occupation is an aggregated job title, that respondents still in a valid way will recognize as at their job title"*

Thus, occupational titles should be precise, understandable, and close to the wording used in job titles. A respondent should not have to guess which occupational title fits best. The choice-set should not include synonym occupational titles, as these will confuse respondents. The choice-set should not be too long, because the higher the average reading time for respondents, the higher the likelihood of dropout during survey completion. Thus,

> *"the source list of occupations has to optimise between the demand to include as many occupational titles as possible to facilitate recognition, and the demand to be as brief as possible to reduce reading time"*

A search tree usually consists of a two- or three-tier tick list, detailing broad categories in the first tier to lists of detailed items in the second or third tier. To prevent visitors from scrolling, a standard search tree on a computer screen can cope with up to 20 first tier items, up to 20*20=400 items in the second tier, and up to 20*20*20=8,000 items in the third tier. As 400 items definitely are too few, the search tree had to be a three-tier search tree with a limit of 8,000 occupation titles. It will be clear from this definition that a coding index does not suit the purpose of a source list of occupations, because of the synonyms and because most indexes have too many items.

A major additional design principle was an evenly distribution over the occupations in the Database. Here, the distributions across occupational groups have been taken into account, derived from Eurostat's database for 3-digit ISCO1988 occupations for many EU member states. In addition, the country-specific distributions across occupations according to the data of the WageIndicator web-survey were taken into account (Tijdens 2009). Nevertheless, in some cases the size of an occupation was based on an estimated guess of the researchers only, not on survey results. The principle implies that units with large numbers of jobholders have been detailed, whereas a breakdown was not perceived desirable for occupational units with a small share in the labour force. Therefore, the Database of Occupations does not hold an occupational title *Clerk*, but has many distinct occupational titles for clerks. Similar breakdowns of occupational units have been made for teachers, nurses, attendants, marketing staff, IT staff, social workers, sales assistants, sales representatives and a few other large occupations.

---

[7]   Note that the character of the answers in this open response format will differ from the text strings collected in open-ended questions without a preceding search tree.

3.5 No additional information needed on industry, firm size or employment status

For good quality recoding of occupations, when derived from an open response format survey question, statistical offices mostly use three additional variables, notably industry, firm size and employment status. For the Database of Occupations it was decided that occupational titles should not require additional information for its coding system. The main reason is that surveys without these three additional questions should also be able to use the Database of Occupations.

Consequently, some occupational titles include a reference to industry or firm size. For example, the Database differentiates between the *Dairy-products process controller* and *Chemical products process controller*, because for both occupations the skill requirements and the competencies are distinct. Concerning firm size, the Database differentiates between three occupational titles for company directors, notably *Company director, chief executive 10-50 employees*, *Company director, chief executive 50-500 employees*, and *Company director, chief executive >500 employees*. In addition, three 2nd tier items in the search tree refer to firm size by making a distinction between management for organisations less than 50 employees, 50-500 employees and more than 500 employees.

The Database does not differentiate occupations according to employment status. The main reason is that employment status is not a feature of ISCO, in contrast for example to the National Occupational Classification of Statistics Netherlands. Additionally, worldwide countries are assumed to vary largely in shares of self-employed within occupations, implying that the Database would almost be double the size when taking employment status into account. In case the Database is used for constructing a Socio-Economic Classification, the survey needs to include an additional survey question on employment status. Alternatively, percentages self-employed per occupation can also be provided for a large number of countries using WageIndicator web-survey data.

3.6 Avoiding gender bias

According to the EurOccupations objectives, gender bias should be avoided. In most classifications, male-dominated occupations are often defined at more narrow boundaries than female-dominated occupations (Tijdens 1990; Ruijter 2002; Ruijter 2005). Analyses on Labour Force Survey data for 8 countries indicate that in the initial EurOccupations lists male-dominated occupations indeed appear to be specified on a more disaggregate level than female-dominated occupations (Ruijter, Ruijter and Jacobs 2009).

Consequently, the average number of workers in female-dominated occupations is higher than in male-dominated occupations. In the 1980's in the Netherlands, the average size of a female-dominated occupation was 3.1 times larger than that of a male-dominated occupation. Three explanations have been given (Tijdens 1990). First, the occupational classification was dated back to 1968, when traditional male-dominated occupations in manufacturing and agriculture were more numerous than in the 1980's. Men tend to work more often in shrinking occupations and women more often in growing occupations. In addition, Taylorism was predominantly found in manufacturing industry, which may have led to more specified jobs in this male-dominated industry compared to for example jobs in clerical work. Second, compared to men are women less equally distributed over job levels, which may cause a less refined occupational structure. Third, female respondents in surveys may be less aware of their professional identity, leading to an underestimation of the number of occupations and a larger proportion of unidentified occupations. The first and second issue can be solved by breaking down large female-dominated occupational units into detailed occupational titles. In the Database of Occupations, this has been applied to the nurse occupation, which was broken down into 11 distinct occupational titles.

3.7 **Coping with composite occupations**

Small organisations tend to employ workers in composite jobs. In most EU member states approximately one out of five employees is employed in a company with less than 10 employees. Spain is an exception with a percentage at least twice as high. For the survey question on occupation, respondents in these organisations may want to classify themselves in more than one occupation. There are two solutions to this problem. The first one is allowing respondents to tick more than one occupation. However, in the WageIndicator web-survey technical constraints have limited this possibility, but this problem may well be solved in some years' time.

The second solution proposes an instruction after the survey question "What kind of work do you do?". Here, respondents should be instructed that in case of composite jobs, they should look for the occupational title that requires the highest skill level, and if there is no measurable difference in skill levels, they should look for the occupation they spend the most time on. In addition, the respondent may be offered an open response format question to specify their composite occupational title.

3.8 **Coping with n.e.c. occupations**

Commonly, the coding instructions for open response format questions state that if a reported occupation cannot be identified as a distinct detailed occupational title in the classification, it should be classified in the appropriate residual occupational category.

For the 436 occupations at the 4-digit level of ISCO-08, 23 are residual units with a "*not elsewhere classified*" addition, e.g. "*Legal professionals not elsewhere classified*". In order to capture all ISCO-08 4 digit codes, these residual occupations had to be included in the Database. Yet, for self-identification, the concept of residual occupations doesn't make sense, because respondents will not read the whole choice-set and then determine that their occupational title is not present. Two strategies have been followed to solve this problem. First, for reasons of readability, all 23 residual occupations have been re-phrased as "*all other*". Second, these residual occupations have been located only at one place at the 3rd tier of the search tree and they have always been sorted at the bottom of the tier, assuming that respondents have red all occupational titles in that particular 3rd tier list before deciding to tick the residual occupation (see also section 4.6).

3.9 **Coping with country-specific occupations**

The Database of Occupations does not aim to be a classification, but a list of occupations that facilitates the vast majority of jobholders in any country to identify their occupation. Therefore, the Database of Occupations has to cope with country-specific occupations.

Countries may add occupational titles specific to their country. The current source list holds almost 1,600 occupational titles, and already an additional 300 country specific occupational titles. Most likely, one occupation in the source list falls apart into two or more occupations. In South Korea the *Restaurant cook* falls apart into *European food cook, Japanese food cook*, and *Korean food cook*. In Germany the *Archivist* falls apart into the *Archivar/in, Diplom (FH)* and the *Archivar/in, Diplom (Uni)*. These two occupations relate to distinct ISCED levels, but within one ISCO skill level. Currently, the Database cannot cope with country-specific deviations of the ISCO-skill levels. This would require a good operationalisation of the four ISCO skill levels for multi-country comparisons and it would require an empirical testing of both required skill levels of occupations and attained skill levels of job holders for all 4-digit ISCO occupational units, assuming sufficient observations per unit per country. The WageIndictor database in part allows for such an empirical investigation. It has the jobholder's attained educational level, it has years of experience and it has a self-assessed skill level in relation to the required skill

level. However, such an analysis was no part of the EurOccupations project. A design for the measurement of country specific skill levels of the ISCO 4-digit unit groups can be made, but it needs a separate project to undertake such an investigation. The result would be country-specific columns, indicating the occupations' skill levels per country, for example ISCOlvl_UK, ISCOlvl_DE, etceteras.

Countries may not translate an occupational title from the source list, because it may not exist, e.g. the winegrower in Finland, or the regional police officer in the Netherlands. In that case, this occupational title is not translated and is thus not part of the national choice-set. However, the number of not-translated occupations should be limited, because any occupation that is not shown to the respondent will not be measured, and will therefore hamper cross-country comparisons.

3.10 **The principles underlying the codes**

The numerical codes in the Database of Occupations correspond with the codes used in ISCO-08. All occupational titles are assigned a 13-digit code (see Table 3.1).

Table 3.1 Break-down of the 13-digit code in the Database of Occupations

Digit	Explanatory note	Range	Explanatory note
1 - 4	ISCO-08 4-digit	100 - 9999	-
5-6	Follow-up code within ISCO-08 4-digit	0 - 99	-
7-8	Country-specific Follow-up code	0 - 99	If within a 4-digit occupation country-specific occupations are added.
9-11	Country code 4 (Albania) -894 (Zambia)	4 -894	ISO Country code s
12-13	Version co	0 - 99	Set to 0

4. Wording, translations and search tree

The wording of the occupational titles in the source list and its translations had to meet several demands. First, respondents should understand the occupational titles quickly, titles should be unambiguous, and the search tree should be designed such that these occupational titles could easily be compared to related occupational titles. Second, translators and national labour market specialists had to be able to understand the occupational titles from the source list. This section details these principles.

4.1 The readability

Self-identification assumes respondents' reading skills and therefore poses high demands on readability. The wording of an occupational title should be simple, brief, easy to understand, and unambiguous. So, *Bee-keeper* has been preferred over *Apiarist.* The singular has been preferred over the plural. Synonymous occupational titles have been avoided, as these are assumed to confuse respondents. As little as possible two or more related occupational titles have been placed in one item. The occupational titles *Bricklayer* and *Stonemason* have been preferred over *Bricklayer or stonemason*. The use of feminine occupational titles is restricted to a minimum, unless the feminine title is common or distinct from the masculine form. So, *Actor* is in the database and *Actress* is not. If possible, occupational titles including the word 'man' are reworded as 'person', such as *Driver-salesperson*.

These readability recommendations may not hold for all countries. In some countries for example both male and female occupational titles have to be used. For Germany, the *DTP operator* is translated into *DTP Operator/in*. In this case, the disadvantage of increasing respondents' average reading time and thus the likelihood of breaking-off survey completion is considered smaller than the advantage of adhering to the national understanding of gender equality.

4.2 The translations

Apart from the readability for respondents, the wording in the source list must also be easy to understand and unambiguous for translators. So, *Chef cook* is preferred over *Chef.* Similarly, the word *Helper* is preferred over *Labourer*. Translations by national labour market experts are preferred over translations by professional translators. In case of the latter, a check by a national expert is strongly advised. For the translations nationally used occupational titles are preferred over literal translations.

Finally, the extent to which job titles are well-defined and widely known varies across countries. Job titles may be known from employment contracts, job classification schemes, collective bargaining agreements, job advertisements, or from a common understanding in workplaces. However, in some countries the institutions that shape job titles might hardly exist, thus if classification schemes are not applied, if collective agreements do not exists or do not include references to job titles, if job recruitment is for unspecified work only, or if the division of labour within the workplace is little. In these cases, national labour market experts are advised to translate occupational titles by means of a brief description rather than a title.

4.3 Occupational upgrading

The survey question about occupational titles may be subject to occupational upgrading, that is that respondents provide an occupational title at a higher skill level than their actual occupation. This problem however can happen in web-surveys and in other survey modes alike. For example, respondents ticking engineer rather than technician or ticking professsional nurse rather than associate professional nurse. The Database of Occupations has no straightforward solutions to this problem, apart from employing unambiguous occupational titles, and a uniform wording, such as the word "technician" for occupations in major group 3 (skill level 3) and "engineer" for occupations in major group 2 (skill level 4).

4.4 Changing occupational titles

When surveying occupational titles, the problem of imago-related changes in occupational titles has to be solved. Some years ago, a *Receptionist in a hotel* all of a sudden was called a *Hotel front desk officer*, and because many hotel chains are operating worldwide, the new occupational title became in use in many countries. The Database of Occupations was adapted accordingly because this new occupational title is quite likely recognisable for jobholders whose job title is *Receptionist in a hotel*.

Recently, the *Call centre agent* is growingly called *Customer care service agent*. Here, we decided to keep the *Call centre agent*, because the *Customer care service agent* was assumed to be too ambiguous to be recognized instantly by jobholders and therefore might not lead to good translations.

4.5 **The search tree**

Three techniques facilitate choices from a long choice-set, notably a search tree, an alphabetical list or a word recognition facility. Here, the first method is addressed, because only this method has been used in the WageIndicator web-survey. Neither an alphabetical list nor word recognition have been used, although they might be in the future. A search tree provides a grouping of occupations in job families. The clustering must facilitate quick and easy understandable search paths to the 3rd tier occupation, i.e. the occupations with the largest numbers of jobholders must have the most obvious search paths.

In the Database of Occupations, the 1st tier uses a mixture of broad occupational groups and industry groups, as is shown in the Table 4.1. Since 2004, this 1st tier has been in use in the WageIndictor web-survey and has proven to be a good entry. Only few minor adaptations have been made over the past years. Other web-sites with search trees use similar job families. Note that this clustering is very different from the major groups (1-digit) in ISCO-88 and ISCO-08. The ISCO hierarchy is designed for taxonomy reasons. The search tree in the Database of Occupations is not designed for taxonomy reasons, but only for facilitating self-identification.

Table 4.1 The 1st tier of the search tree in the Database of Occupations

1st tier list of the search tree
Agriculture, nature, animals, environment
Care, children, welfare, social work
Cars, mechanics, technicians, engineers
Cleaning, housekeeping, garbage, waste
Clerks, secretaries, post, telephone
Commercial, shop, buy and sale
Construction, fittings, housing
Education, research, training
Finance, banking, insurance
Food manufacturing
Guards, army, police
Health care, paramedics, laboratory
Hospitality, tourism, leisure, sports
HRM, labour intermediary, organisation
Industrial production, manufacture, metal
IT, automation, telecommunication
Language, library, archive, museum
Legal, administration, inspection, policy adviser
Management
Marketing, PR, advertising
Media, graphic, printing, culture, design
Oil, gas, mining, utilities
Transport, logistics, port, airport

4.6 **Alphabetical sorting**

Alphabetical sorting is a major principle for a search tree. In each language and all scripts, all items are sorted alphabetically within each tier. The obvious reason is that alphabetical searching from a long list of items is the most common way people tend to select an item from a long list. It increases the search tree's user-friendliness, and thus improves data-quality. For an example, see Tale 4.2.

Alphabetical sorting assumes people reading a list from a to z. This sorting is disadvantageous because respondents might be tempted to tick an item at the top. As the sorting varies across countries, this will affect cross-country comparisons. A randomly sorted list is preferred over an alphabetical sorted one, but this is disadvantageous because it is counter-intuitive and the average reading time is assumed to be longer. Compared to the first method, the latter is judged more disadvantageous.

Alphabetical sorting requires that related occupational titles are preferably sorted one after another, allowing respondents the best choice from the choice-set. This may require the rephrasing of occupational titles, but these possibilities vary across languages. Therefore, rephrasing has not been used in the source list, but it is applied in the Netherlands translation, as the author is familiar with the occupational titles in the Netherlands.

Finally, in case a choice-set within a tier included a residual or "not-elsewhere-classified" occupation, this addition has been rephrased as "all other". Such an occupation was located at the bottom of the tier, assuming that it was the final choice from the list in the tier, see Table 4.2.

Table 4.2 Part of the search tree place, including a "not-elsewhere-classified" occupation in the 3rd tier of the search tree

1st tier in search tree	2nd tier in search tree	3rd tier in search tree
Agriculture, nature, animals, environment	Fishery	Aquaculture fishery worker
		Fisheries advisor
		Fisheries manager
		Fishery or aquaculture helper
		Fishery technician
		Inland waters fisherman
		Seafood farm worker
		Seafood farmer
		Fishery worker, all other
	Forestry, nature	Charcoal burner or related worker
		Countryside or park ranger
		First line supervisor forestry workers
		Forest worker, lumberjack
		Forestry advisor
		Forestry helper
		Forestry manager
		Forestry planter
		Forestry technician

5. Skill levels, job ladders and hierarchies

5.1 Occupational titles and skill levels

As discussed in Section 2, in ISCO-88 skill levels were taken as the basis of the hierarchical structure of the occupational classification and this principle has been confirmed for ISCO-08 (ILO, 1990; ILO 2007). Skill levels have been operationalised referring to the International Standard Classification of Education (ISCED), developed by UNESCO (UNESCO, 2006, re-edition). Table 5.1 shows how the ten ISCO-88 major groups map to four ISCED skill categories.

Table 5.1 ISCO-88 major groups and skill level

Code	Major group - label	Skill level	ISCED category
1	Legislators, senior officials and managers	-	
2	Professionals	4	6 -7
3	Technicians and associate professionals	3	5
4	Clerks	2	2 - 3
5	Service workers and shop and market sales workers	2	2 - 3
6	Skill agricultural and fishery workers	2	2 - 3
7	Craft and related workers	2	2 - 3
8	Plant and machine operators and assemblers	2	2 - 3
9	Elementary occupations	1	1
0	Armed forces	-	

Source: Elias and Birch, 1994

The EurOccupations investigations about the comparability of 150 occupations across 8 EU member states showed that the vast majority of these occupations did not vary largely within and across countries with regard to their task descriptions, but that they did vary with regard to their skill levels (see Deliverables D07-D22). These across-country differences might be understood from differences in educational systems. Within-country differences might be understood from measurement problems. Skill levels have been operationalised as educational entry requirements, following the tables mapping national educational categories into ISCED (OECD, 1999). However, in the EurOccupations project entry requirements were sometimes understood as minimal and sometimes as average requirements. Additionally, a focus on educational entry requirements left little room for years of experience as entry requirement, and this turned out to be particularly problematic for the managerial occupations. A final measurement problem relates to difficulties for empirical investigations of skill level requirements.

Following the discussion on the empirical underpinnings of the skill levels of ISCO occupations, Elias and McKnight (2001) conclude from a comparison of two UK occupational classifications that analysis of detailed occupational categories should be undertaken with care. At the aggregate level, however, where perhaps only three or four skill categories may be defined, occupational classifications appear to provide a robust method for the measurement and analysis of skill. Dumont (2006) assumes that the ISCO skill levels in principle are to be preferred instead of the proxy data for skills that are more often used. Yet, based on simple wage regressions for 4 EU member states Dumont concludes that the ISCO skill levels are not very reliable.

Although it has been argued that the ISCO-skill categories are so broad that they are not vulnerable to continuously changing entry requirements, empirical underpinnings of the ISCO skill levels might strengthen the value of the classification. This will however be a major and difficult undertaking. The EurOccupations project undertook empirical skill level investigations for 150 occupations for 8 countries. It can easily be understood that for ILO's ISCO-08 an empirical investigation of skill levels for 436 occupational units in more than 100 countries must have been perceived a too huge and uncertain undertaking, let alone keeping the skill level measurement updated. The lack of empirical underpinnings may explain why skill levels for quite a number of occupations have been changed from Draft 3 to the final ISCO-08 classification.

5.2 Occupational titles within the corporate hierarchy

In Section 3.7 it has been argued that small organisations tend to employ workers in composite jobs. In most EU member states approximately one out of five employees is employed in a company with less than 10 employees. In contrast, large and medium-sized organisations usually have a well-developed division of labour, shaping clear demarcation lines between occupations. Because the majority of the Labour Force is employed in these organisations, we assumed clear demarcation lines across occupations as well as a stylized, six-layer corporate hierarchy, serving the design of the Database. As Table 5.2 shows, the six layers are CEO's and board members, managers, heads of department or branch, first line supervisors, occupations, helpers. In the following, these six layers and their implications for the database will be discussed.

Table 5.2 The stylized, six-layer corporate hierarchy applied in the Database of Occupations

Hierarchy	Description	ISCO major group
OCC+4	CEO, board members and area managers of large firms or organisations (50 or more employees)	Major group 1
OCC+3	Managers of institutions, centres, branches and alike Company director, chief executive 10-50 employees	Major group 1
OCC+2	Departmental managers, using a stylized setting of 14 departments	Major group 2-3
OCC+1	First line supervisors	In same group as occupation
OCC	Occupation	-
OCC-1	Helpers	Major group 9

CEO's, directors, managers

Both ISCO and SOC_US include occupational titles for managers. Over the years, the rapidly increasing number of managers as well as the ambiguity in the occupational title of manager has sometimes been problematic for occupational classifications (see for the latter, Elias and Birch, 1994). In quite some languages, no clear differences exist between the hierarchical manager and the person who is responsible for a product or a service within the organisation, mostly also called manager. Since the late 1990s, some European countries have for example noticed a rapid increase in the number of account managers. To distinguish between types of managers, the source list has followed the ISCO-08 principles and classified all hierarchical managers in major group 1, and all other managers, such as *Periodical or newspaper manager* or *Brand manager*, in major group 2 or 3.

The Database of Occupations distinguishes three categories of managers, in Table 5.2 abbreviated as OCC+4, OCC+3, and OOC+2. As for OCC+4, Table 5.3 shows the occupations in the source list. As for OCC+3, a long list of managers of all types of businesses in included in the source list, such as *Call centre manager*, or *Energy plant manager*. The *Company director, chief executive 10-50 employees* is also classified as OCC+3.

Table 5.3 CEO's, board members and area managers

OCC+4: CEO's, board members and area managers in 50+ organisations	
Company director, chief executive >500 employees	Company director, chief executive 50-500 employees
Administrative services manager	Logistics manager
Advertising or public relations manager	Manufacturing plant manager
Commercial, sales or marketing manager	Policy or planning manager
Finance manager	Production or operations manager
HR manager	R&D manager
IT manager	Manager, all other services

The search tree in the Database of Occupations plays a major role for a valid self-identification of managers. Table 5.4 shows the part of the search tree, specific for managers. The search tree allows locating one occupation at several places at the 3rd tier. Therefore, all Managers of institutions, centres, branches and alike (OCC+3) are also located elsewhere in the search tree. For example the *Call centre manager* is also located in the 3rd tier list of *Call centre agent inbound*, *Call centre agent outbound*, and *First line supervisor call centre agents*.

Table 5.4 Management occupations in the search tree

1st tier in search tree	2nd tier in search tree	3rd tier in search tree
Management, direction	Company director, chief executive >500 employees	Company director, chief executive >500 employees + List of Area managers
	Company director, chief executive 50-500 employees	Company director, chief executive 50-500 employees + List of Area managers
	Company director, chief executive 10-50 employees	Company director, chief executive 10-50 employees + List of Managers of institutions, centres, branches and alike
	Department manager	List of Departmental managers

Department managers

Neither ISCO nor SOC_US identifies department/group/team/division managers. The Database of Occupations however does so (OCC+2), because the from the WageIndicator web-survey it is known that this category exists. If they are assigned the code of the occupation that they supervise, they are not distinct from the first-line supervisors. This may seem justified when coding occupations, but in case of self-identification respondents prefer to stress that they are department manager, as the WageIndicator web-survey shows. Moreover, if not clearly indicated, they just say "I'm a department manager", without reporting what kind of department. For validity reasons the Database of Occupations holds distinct occupational titles for the department managers, whereby we applied a stylized horizontal corporate structure of 14 departments (see Table 5.5).

Table 5.5 A stylized view on the 14 departments within organizations

OCC+2: Departmental managers	
Production, core activities	After sales, including installation and repairs
General management	Finance, accountancy, audit
Administration, excluding finance	Human Resources personnel management
Technical support, quality control	Research and Development
Logistics, purchasing	Engineering
Marketing	IT
Sales, commercial	Cleaning, security, housekeeping

First line supervisors

The US occupational classification employs occupational titles for first line supervisors, but neither ISCO-88 nor ISCO-08 do so, although this issue has been heavily debated in the preparation of ISCO-08. To meet the US mapping requirement the Database of Occupations includes distinct occupational titles for first line supervisors, e.g. *First line supervisor personal care workers, First line supervisor protective service workers*, or *First line supervisor retail sales workers*. First line supervisors (OCC+1) are assigned the same ISCO code as the occupation that they supervise.

Helpers

Following the instructions of the SOC_US, helpers and aides should be classified distinctly. According to ISCO, helper occupations are mostly classified as unskilled occupations. In the ISCO-08 terminology, these occupations are primarily called labourers. In order to prevent translation mistakes and on behalf of valid self-identification, the word helper is preferred over labourer, because the latter refers to a far more general group than the former and may therefore be ambiguous. The Database of Occupations includes almost twenty distinct occupational titles for helpers, e.g. *Assembling helper, Building construction helper, Cattle station helper*, or *Construction helper.* Helpers (OCC-1) are assigned an ISCO skill level different from the related occupation.

The corporate hierarchy in agriculture and retail

In agriculture and in retail, the above sketched stylized hierarchical corporate structure might be perceived as less adequate. However, both in agriculture and in retail, large corporations can be noticed. In agriculture, in the stylized hierarchical structure the farm manager (OCC+3) has been distinguished from the farmer/breeder/grower (OCC), the first line supervisor of farm workers (OCC+1), the farm worker (OCC) and the farm helper (OCC-1). In retail, the stylized structure consists of the store manager (OCC+3), the shop keeper (OCC), the first line shop supervisor (OCC+1), and the sales assistant (OCC) / check-out operator (OCC). Apart from the shelf-stacker (OCC), no helpers have been distinguished in retail.

5.3 Job ladders: from assistant to junior and to senior job titles

The concepts of careering, job ladders and job-enlargement blur the demarcation lines across occupations, whereas clarity remains a necessary precondition for valid measures of occupations. In our study, blurred demarcation lines turned out most problematic for the assistant occupations. Is the assistant plumber part of a job ladder to become a plumber? Will the medical assistant ever become a General Practitioner through on the job training? For the former case this is likely and thus not assigned a distinct occupational title, for the latter it is not likely and therefore assigned a distinct title. The word 'assistant' is not informative as to what extent the job is part of a job ladder. For coding purposes statistical agencies mostly have instructions for coping with this problem in case of field- or office-coding. For self-identification purposes as well as for high quality translations, the problem needs to be solved too. For the Database of Occupations the word ´assistant´ was only included in the few cases that synonyms were lacking and the occupation was not a part of a job ladder.

Job ladders may include junior and senior job titles, in some countries more than in others. It refers to a common understanding that a jobholder will start as a junior and through on-the-job training will reach a senior position. For jobholders, the seniority of their job title may be an important part of their job identification, and therefore they may prefer to see the word senior added to their occupational title. However, the Database of Occupations does not include this detail in its occupational titles. Similarly, the database

neither includes distinct occupational titles for jobholders who alongside their job provide training for junior workers in the occupation, nor for jobholders who are trainees. Surveys that aim for more details about the respondent´s relative position within the occupation are advised to include an additional survey question. For Germany, with its elaborated system of vocational training, on-the-job learning, and the "Meister" title, several country-specific occupations have been added. Table 5.6 shows an example.

Table 5.6 Country-specific job ladders in occupations

Code	Source list	de_DE
7133030000000	Chimney sweep	Schornsteinfeger/in
7133030227600	DEU Chimney sweep Schornsteinfegermeister	Schornsteinfegermeister/in

In surveys, the desire of respondents to report about the job ladder hierarchy in their occupation can be channelled by a follow-up open response format question or by a follow-up survey question presenting the most likely phrases used for job ladders. Table 5.7 shows the question in the WageIndicator web-survey about the hierarchy within the respondent's occupation. In WageIndicator, country-specific response categories exist for this survey question.

Table 5.7 Survey question for measuring a hierarchy within occupations

Question	Would you like to specify your occupation?
1	No, the job title is all right
2	Assistant
3	Junior
4	Apprentice
5	Trainee
6	Trainer, training responsibilities
7	Foreman, forewoman
8	Senior
9	Team leader
10	Supervisor
11	Head
12	Master

5.4 Understanding occupational dynamics over time

Handicraft workers and machine-operators

In the cause of the 20th century, small-scale workshops have been replaced by factories and craft occupations by machine-operators due to industrialisation and technological innovations. Countries vary with respect to the degree that these processes have taken place. Handicraft occupations however have nowhere fully disappeared. Even in highly industrialised countries traditional craft occupations are supplying handicraft goods for commercial markets. In the Database of Occupations the machine operator and the handicraft workers are assigned distinct occupational titles, e.g. the *Handicraft weaver, knitter, embroiderer* is distinguished from the *Weaving machine operator*, and the *Handicraft leather worker* from the *Shoemaking machine operator*. For food manufacturing, the word handicraft worker is not applicable. For bakers and butchers, the occupational titles in most countries primarily refer to retail, and different phrasing is used for similar occupations in manufacturing. Table 5.8 shows a selection of the handicraft occupations and the paths in the search tree. Note that one occupation can be located at two places in the search tree.

Table 5.8 Selection from the Database of Occupations – handicraft versus manufacturing workers

3rd tier - Occupation	2nd tier in the search tree	1st tier in the search tree
Handicraft leather worker	Repair	Commercial, shop, buy and sale
Handicraft weaver, knitter, embroiderer	Craft	Media, graphic, printing, culture, design
Handicraft weaver, knitter, embroiderer	Textile, leather, carpeting	Industrial production, manufacture, metal
Handicraft worker in basketry or brushes	Textile, leather, carpeting	Industrial production, manufacture, metal
Handicraft worker in textile or related materials	Textile, leather, carpeting	Industrial production, manufacture, metal
Handicraft worker in wood or related materials	Wood	Industrial production, manufacture, metal

Broad titles for unskilled occupations

For unskilled occupations, broad occupational titles are preferred, because it is likely that in quite a number of countries unskilled work predominantly involves day labourers, who perform one day another occupation than the other day. So, one may distinguish a *Fruit porter* from a *Fish porter*, but as the same person might perform both jobs within a single week, the Database of Occupations preferred the occupational title *Food porter*.

Subsistence farmers, fishers, hunters and gatherers

The evidence of subsistence workers has been an issue of debate. Due to rapid urbanization worldwide, the share of subsistence workers within the labour force is likely to decrease further, although the current economic crisis may well lead to a temporal increase of this group. ISCO-88 included a sub-major group *Subsistence agricultural and fishery workers* (nᵒ 62), in contrast to the *Market-oriented skilled agricultural and fishery workers* (nᵒ 61). However, within the European Union, sub-major group 62 was not used.

ISCO-08 also includes a sub-major group for the subsistence workers *Subsistence farmers, fishers, hunters and gatherers* (nᵒ 63). Now, this group is detailed into the *Subsistence crop farmers* (6310), *Subsistence livestock farmers* (6320), *Subsistence mixed crop and livestock farmers* (6330), and *Subsistence fishers, hunters, trappers and gatherers* (6340). For this reason, the source list in the Database of Occupations also includes these occupations. For the Netherlands these occupations have been translated with the extension "hobby", e.g. the *Subsistence livestock farmer* is translated into *Veehouder (hobby),* to be distinguished from *Veehouder.*

Obsolete occupations

Should the source list include very small and almost obsolete occupations? And how to detect which occupations are obsolete? One may argue that in developing countries occupations may exist that are obsolete in advanced industrialised countries. However, we assume that this disparity is decreasing, because the supply of machinery for factories, building sites, IT services and offices is increasingly a globalised business, and therefore does not leave much room for obsolete occupations.[8] Therefore, occupations that given the technological state-of-the-art in manufacturing and services are judged obsolete, such as *Riveter* or *Lead burner*, are not included.

---

[8] This is an observation not yet underpinned with empirical results.

New occupations

A major challenge for any occupational classification is how to cope with new occupations. To be more precise: does an occupational title, not yet administered in the codig indexes of a classification, reflect a new occupation, is it a fashionable change in an existing occupational title (see also section 4.4), is it a split off of an existing occupational title, or is it just a wrongly reported occupation? Compared to ISCO-88, ISCO-08 added a number IT occupations to the classification, after an exploration of these new occupations (Hunter 2006).

The WageIndicator web-survey uses an open response format question to provide respondents with a possibility to write in comments on occupational titles alongside the choice made in the search tree. Annually, a part of the respondents uses this open response format. For the EurOccupations project, these have been checked as far as the German and Netherlands data was concerned (see Section 3.2). This method allows tracing new-arising occupations.

Division of labour in growing businesses

In growing businesses it is likely to observe a growing division of labour. For example within the travel agency business, only one occupation used to be distinguished, notably the travel agent. With the growth in this business, the number of distinct occupations has grown too. Table 5.9 reveals the five occupational titles in the source list.

Table 5.9 Travel agency business occupations in the source list.

Code	Occupational title
1420050000000	Travel agency manager
3332030000000	Travel organiser
4221020000000	Travel agency clerk
4221040000000	Travel consultant
5113010000000	Travel guide

6. Conclusion

The EurOccupations project aimed to "build a publicly available database with the 1,500-2,000 most frequent occupations in the 7 largest EU countries to be used in multi-country data-collection, be it large-scale surveys or case-studies." In the past three years, the project team has undertaken a large effort to build and to test this database. The result is the following:

1. A source list of almost 1,600 occupational titles; these titles are thought to reflect distinct occupations; overlapping occupations have been removed. The methodology for the construction of the source list as well as the principles underlying the source list have been explained in this paper.
2. A search tree to navigate through the source list; this search tree has been tested for its usability for almost two year in the WageIndicator web-survey for a growing number of countries, currently approx 50; the search tree appears to be an appropriate tool for self-identifiction of occupation in web-surveys.
3. Translations of the source list for 7 languages, funded from the EurOccupations project, and for another 10 languages, funded from the WageIndicator Foundation; the translations have been checked extensively for the 8 EurOccuations countries and to a lesser, varying extent in the remaining 50 countries. As the WageIndicator web-survey is continuously expanding, more translations are expected to be added in the near future.
4. A test of the comparability of a selection of 150 key occupations from the source list with regard to their task descriptions and skill levels in eight countries, whereby the skill levels turned out to be less convergent than the task descriptions. The Database of Key Occupations, including all information on job content and skills will remain freely available from www.euroccupations.org.
5. The source list of almost 1,600 occupational titles, the search tree and the translations jointly make up the Database of Occupations. This Database will remain freely available from www.euroccupations.org. The WageIndicator Foundation will continue to use the Database for its web-survey.
6. Respondents as well as academic users from all over the world are invited to contribute to the Database of Occupations: suggestions for improvements of occupational titles in a particular language/country are welcomed, reflections on the search tree are welcomed, additions for new countries/languages are particularly welcomed. All messages can be send to the author and EurOccupations coordinator k.g.tijdens@uva.nl.

References

De Vries, J. and H.B.G. Ganzeboom, 2006, *Hoe meet ik beroep? Open en gesloten vragen naar beroep toegepast in statusverwervingsonderzoek.* Universiteit van Tilburg/ Vrije Universiteit Amsterdam, latest version 29 March 2006

Dumont, M., 2006, The reliability–or lack thereof–of data on skills. *Economics Letters* 93, 348–353

Elias, P., 1997, Occupational Classification (ISCO-88) Concepts, Methods, Reliability, Validity and Cross-National Comparability. OECD Labour Market and Social Policy Occasional Papers, No. 20

Elias, P., and M. Birch, 1994, Establishment of Community-Wide Occupational Statistics. *ISCO 88 (COM). A Guide for Users.* University of Warwick, Working paper Institute for Employment Research

Elias, P., and A. McKnight, 2001, Skill measurement in official statistics: recent developments in the UK and the rest of Europe. *Oxford Economic Papers*, 3, 508-540

Elias, P., and M. Birch, 2006, The Review of ISCO-88: A European perspective. University of Warwick, Working paper Institute for Employment Research

Elias, P., R. Ellison, and R. Jones, 2009, Development of associated software for EurOccupations database. EurOccupations Deliverable 26, www.eurocupations.org

Ganzeboom, H. and D. Treiman, 1996, Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research* 25**,** 201–239

Greenwood, A.M., 2004, *Updating the International Standard Classification of Occupations, ISCO-08.* Geneva, ILO Bureau of Statistics

Halstead, K., P. Elias and K. Prandy, 1993, *CASOC: Computer-assisted Standard Occupational Coding* London, HMSO

Hoffmann, E., 2000. *Structures: Problems, possibilities and the role of ISCO-88*. Geneva, International Labour Office

Hoffmann, E., P. Elias, B. Embury, and R. Thomas, 1995, *What kind of work do you do? Data collection and processing strategies when measuring "occupation" for statistical surveys and administrative records.* Geneva, ILO Bureau of Statistics, STAT working paper 95/1

Hunter, D. (2006) Occupations in Information and Communications Technology. Options for Updating the International Standard Classification of Occupations. Geneva, ILO, Discussion Paper April 2006

Hunter, D. (2009) ISCO-08 Draft definitions. Geneva, ILO, June 1, 2009

ILO, 1990, *ISCO-88: International Standard Classification of Occupations*. Geneva: International Labour Office

ILO, 2006, *Draft classification structure – ISCO-08. Including Annex 1 for updating ISCO-88 into ISCO-08.* Geneva, International Labour Office

ILO, 2007, *Resolution Concerning Updating the International Standard Classification of Occupations. Tripartite Meeting of Experts on Labour Statistics on Updating the In-*

ternational Standard Classification of Occupations (ISCO), 3 to 6 December 2007. Geneva, International Labour Office

Jacobs, J., P. Michon and K. Tijdens, 2007, EurOccupations: Developing a European occupations database. State-of-the-Art rapport. Deliverable 35 EurOccupations project, www.eurocupations.org

Jones, R., P. Elias, no year, CASCOT: Computer-Assisted Structured Coding Tool. Institute for Employment Research, Coventry, University of Warwick

McGuckin, R.H., 1991, Multiple Classification Systems For Economic Data: Can A Thousand Flowers Bloom? And Should They? US Census, Economic Studies

OECD, 1999, *Classifying Educational Programmes Manual for ISCED-97 Implementation in OECD Countries.* Paris: OECD

Osterman, P. (1995) Skill, Training, and Work Organization in American Establishments. *Industrial Relations*, 34 (2), 125-146

Ruijter, J. de (2002). *Occupational wage differences. A gender approach*. Proefschrift, Erasmus Universiteit Rotterdam

Ruijter, J. de (2005). Genderaspecten in beroepscompetentieprofielen*.* In: De Glazen Muur. *Essaybundel over beroepensegregatie*. Den Haag: Ministerie van SZW/Directie Coördinatie Emancipatiebeleid

Ruijter, J. de, E. de Ruijter, M. Veldhoen, 2007a, EurOccupations: Measuring occupational skill levels and occupational content in the EU. EurOccupations Deliverable 6a, www.eurocupations.org

Ruijter, J. de, E. de Ruijter, K. Tijdens, J. Jacobs, in close collaboration with the whole EurOccupations research team, 2007b, EurOccupations: Questionnaires data collection. EurOccupations Deliverable D6a, www.eurocupations.org

Ruijter, J. de, de Ruijter, E. de and J. Jacobs, 2008, EurOccupations: Note on the selection of key occupations and the 8 clusters of occupations. EurOccupations Deliverable D06b, www.eurocupations.org

Ruijter, J. de, de Ruijter, E. de and J. Jacobs, 2009, Paper on gender bias in the EurOccupations database. EurOccupations Deliverable D23, www.eurocupations.org

Tijdens, K.G., 1990, Beroepssegregatie en werkgelegenheid. *Tijdschrift voor Arbeidsvraagstukken,* 6 (4), 13-23

Tijdens, K.G., 2007a, De beroepen codeset en zoekboom in de Loonwijzer web-survey. WageIndicator report February 2007, www.wageindicator.org

Tijdens, K.G., 2007b, The occupations database in the WageIndicator questionnaire. WOLIWEB paper D18b, www.wageindicator.org

Tijdens, K.G., 2007c, The EurOccupations database and its underlying principles. EurOccupations Deliverable 1a, www.eurocupations.org

Tijdens, K.G. and J. Jacobs, 2007, Note on the search tree and translations used for the Extended list of occupations. EurOccupations Deliverable 1b, www.eurocupations.org

Tijdens, K.G., and J. Jacobs, 2009, Note on the revised search tree and translations of the EurOccupations extended list of occupations. EurOccupations Deliverable D01b1, www.eurocupations.org

Tijdens, K.G., 2009, Frequency tables: education distributions at aggregated levels per country. EurOccupations Deliverable D01e, www.eurocupations.org

UNESCO (2006, re-edition) *International Standard Classification of Education I S C E D 1997*, New York, UNESCO-UIS