



www.wageindicator.org

WOLIWEB

The socio-economic determinants of citizens' work life attitudes, preferences and perceptions, using data from the continuous web-based European Wage Indicator Survey

Classifications and the chooser-technique used in the *WageIndicator* questionnaire

Deliverable: D17a State-of-the-art Report
Report Version: Final
Report Preparation Date: 03.01.2005
Authors of this report: Kea Tijdens, AIAS, Univ. of Amsterdam
Cecile Wetzels, AIAS, Univ. of Amsterdam



**Project funded by the European
Community under the 6th Frame Work
Programme PRIORITY 7 Citizens and
Governance in a knowledge based society
(FP6-2004-506590)**

TABLE OF CONTENT

1.	Introduction	2
2.	Choosers	2
3.	The ISCED classification and the education question	3
4.	The concept of ethnic background	5
5.	The NACE classification and the industry chooser.....	6
6.	The ISCO classification and the occupation chooser	7
7.	The ISO, UN and Worldbank classification and the country chooser	9
8.	The NUTS classification and the region chooser	10
Appendix	Principal characteristics of the six levels in the isced educational classification	12

1. INTRODUCTION

The *WageIndicator* questionnaire uses international renowned classifications for variables such as education, ethnic background, industry, occupation, country and region of residence. This paper details how these classifications are used. It is part of a series of papers all related to the *WageIndicator* project. All papers are downloadable from www.wageindicator.org, section research lab. The classification tables discussed here can also be downloaded from the research lab.

2. CHOOSERS

Survey questions about these variables are most commonly asked either as an open text field or from a short aggregated list. For the *WageIndicator* questionnaire, we have developed the so-called 'choosers', facilitating the web-visitors' searching for the applicable, detailed item. In a written, web-based questionnaire, a chooser is the only workable solution for gathering detailed information without being restricted to either a huge recoding effort or a limited list of items. A chooser enables a web-visitor to choose easily from a long list of items, for example industries, occupations, or regions. It consists of a two- or three-step ticking list from an aggregated level in

the first step to a detailed level in the second or third step. For each step a list of items is presented on the screen. Once an item is ticked, the appropriate items of the next step appear. The chooser technique in *WageIndicator* questionnaire allows the web-visitor to go easily back-and-forth in the chooser.

Underlying a chooser, a database links first level items to second and third level items. The links in the choosers are identical for all country and language versions of *WageIndicator* questionnaire. The items in the database are available in the languages covered in the *WageIndicator* questionnaire. By the end of 2004, these languages are DE, DK, EN, ES, FI, FR, IT, NL, and PL.

A chooser must meet contradictory demands. User-friendliness requires a minimum number of words, and thus a minimum number of items and steps. This reduces the likelihood that visitors quit the questionnaire before the end. Yet, visitors must be able to identify their particular enterprise, and may not be able to identify broader concepts covering their industry. Thus, it should be avoided that visitors must think whether a hairdressing enterprise is part of the retail sector. Therefore, the chooser requires detailed lists of items. Finally, for reasons of data-analysis detailed codes are preferred.

User-friendliness is a key word in any chooser, because choosers are time-consuming and its use should be minimized. For reasons of user-friendliness each list in the chooser, be it in the first, second or third step, is automatically sorted by alphabet for each language. In some choosers, the second or third step, whatever is last, has an item 'other' as the last item of the list. An item cannot be longer than 100 characters, but should preferably be as short as possible. If one item consists of two objects, for example manufacturing of wood and cork, and employment in wood is larger than in cork, it is user-friendly to place the object with the largest employment first. In addition, difficult words should preferably be replaced by words that are easier to understand. Finally, words that are commonly used in a language are preferred to the official language of the classification.

3. THE ISCED CLASSIFICATION AND THE EDUCATION

QUESTION

The ISCED classification is used to recode the national education levels according to the question: 'What is the highest level of education you have attained (with

certificate)?'. This includes the instruction 'If you are currently studying, please indicate the highest level of education you reached. If you went to school abroad enter the equivalent level'. To respondents having indicated that they are a school pupil or student in full-time education, the question is: 'At what stage of education are you at the moment?'.

The answer consists of a country-specific radio-button list of educational categories, commonly used in national surveys. This list covers the major national educational categories, including previous educational categories, i.e. a 50-yr employee who has performed an education that today is replaced by another education (al category or an educational category that holds a new name) still must be able to tick the appropriate button in the question.

The worldwide ISCED classification is used to recode the country-specific educational categories into the 6 ISCED-categories. The International Standard Classification of Education (ISCED) was designed by UNESCO in the early 1970s to serve 'as an instrument suitable for assembling, compiling and presenting statistics of education both within individual countries and internationally'. The present classification, now known as ISCED 1997, was approved by the UNESCO General Conference in November 1997. It takes into account new developments and changes in education in various regions of the world. In the appendix, we summarize the six levels of education in ISCED 1997 and the principle characteristics of each level. For details see http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm.

In the *WageIndicator* dataset, education is covered by a number of numerical variables. The variable EDUCATIO covers the country-specific educational categories, coded from 1 'primary school Netherlands' to over 200. The variable EDUCATI1 covers the same country-specific educational categories for school pupils or students. The variable EDUCATIS covers the ISCED codes of the country-specific educational categories for both individual with finished full-time education (both vocational and general education) and school pupils or students. In the variable EDUCATVO 'vocational education', each educational category is classified as either a vocational (1) or a general education (0). In the variable EDUCATY6 'years of education after age 6', each educational category is converted into the number of normal years of schooling or training after age 6.

4. THE CONCEPT OF ETHNIC BACKGROUND

There is no classification for ethnicity. According to Billiet¹ measuring country of origin, nationality, mother tongue and denomination together should make it possible to construct a highly precise indicator of ethnicity. The *WageIndicator* questionnaire covers country of origin and mother tongue in all surveys, and poses an additional question for the UK about ethnic group. No questions are posed about denomination or nationality.

The *WageIndicator* questionnaire has questions about the country of birth of the respondent self, as well as his/her mother and father. The first question is 'Were you born in COUNTRY OF SURVEY?'. If no, a question follows 'In which country were you born?' with a limited radio-button list of countries of origin, most found in the country of survey. In case the country of birth is 'other', the COUNTRY chooser pops up (see section 7). The following questions are 'Was your mother born in the same country as you?' and 'Was your father born in the same country as you?'. If ticked 'no', the question with the limited list of countries pops up again, and in case the country of birth is not listed, the COUNTRY chooser pops up again. This way of questioning will allow the large majority of respondents to easily pass these three questions. Only in rare cases respondents have to go through the COUNTRY chooser. Respondents not born in the country of survey are asked in what year they arrived and whether they came for asylum or for economic, family or other reasons.

In addition, the question 'Which language do you mostly speak at home?' is posed, followed by a limited radio-button list of most commonly spoken languages in the country of survey. The languages have been chosen on the information by each country. Some countries have ranges of languages spoken at home², but we selected only up to 10, including the native language(s) and an option 'local dialect'.

In the UK questionnaire, a question is posed to identify ethnic group, since migration history and culture towards ethnic background is different and this question make

¹ Chapter 10 Questions about National , Subnational and Ethnic Identity, Suggested by Jaak Billiet K.U. Leuven, for the European Social Survey www.europeansocialsurvey.org/

² In the UK for example Bengali / Silheti; Panjabi, Gujerati, Hindi / Urdu, Turkish, Arabic, Creoles (English-based), Yorubu (Nigeria), Somali, Cantonese, Greek, Akan (Ashanti), Portuguese, French, Spanish, Tamil (Sri Lanka), Farsi (Persian), Italian, Vietnamese, Igbo (Nigeria), Creoles (French-based), Tagalog (Filipino), Kurdish, Polish, Swahili, Lingala (Congo), Albanian, Luganda (Uganda), Ga (Ghana), Tigrinya (Sudan), Pashto (Afghanistan)

most sense to a UK audience. Here, we follow the consensus categories recommended by the UK commission for Racial Equality. For detailed information see www.cre.gov.uk/gdpract/em_cat_ew.html. However, we do consider one question for the UK as in the UK Labour Force Survey, and do not follow up to the advice of the Commission for Racial Equality to ask separate questions for England/Wales, Scotland and Northern Ireland. The *WageIndicator* question is posed as follows: 'To which ethnic group do you belong?', followed by a limited radio-button list of most common ethnic groups.

In the *WageIndicator* dataset, the country of birth (COB) with the variable names COBSEFX, COBMOTHX, COBFATHX, is coded according to the coding in the COUNTRY chooser (see section 7). The variable COBETHGR is coded 1=White, 2=Mixed, 3=Asian / Asian British, 4=Black / Black British, 5=Chinese and 6=Other. The variable COBLANGU has codes ranging from 0 to 90.

5. THE NACE CLASSIFICATION AND THE INDUSTRY CHOOSER

The NACE classification is the basis of the INDUSTRY chooser in the *WageIndicator* questionnaire. It is used for the question 'What is your current business activity'. The NACE classification of industries and branches of industries is firmly standardised by Eurostat, and it is used by almost all national statistical agencies in Europe. We used the Statistical Classification of Economic Activities in the European Community, Rev. 1.1 (2002) (NACE Rev.1.1). <http://europa.eu.int/comm/eurostat/ramon/> .

The NACE classification has 848 partly overlapping items, covering five levels.³ The INDUSTRY chooser uses all NACE codes at the fourth level, except for manufacturing industry, where the third level is used only. This is done for two reasons: (1) NACE details the manufacturing industry to a much larger extent than other types of industry; (2) due to mergers and new technologies, these very detailed fourth level industries do not exist anymore as separate enterprises, making the third level a better approach of reality. The INDUSTRY chooser is a two-step chooser with 19

³ Level 1 : 17 divisions identified by one-letter alphanumerical codes (A to Q);
Level 2 : 32 classes identified by two- letter alphanumerical codes (AA to QA);
Level 3 : 62 groups identified by two-digit numerical codes;
Level 4 : 224 sub-groups identified by three-digit numerical codes;
Level 5 : 514 items identified by four-digit numerical codes.

first-step items and 208 second-step items. A few NACE codes were split into two items in the INDUSTRY chooser, because it was user-friendlier to place them.

In the *WageIndicator* dataset, the NACE codes are used to identify industries. The variable SECTO2_I has numerical, 4 digit codes, including two decimals, to identify the third/fourth level industries in NACE. The value labels are in English.

6. THE ISCO CLASSIFICATION AND THE OCCUPATION CHOOSER

The ISCO classification is the basis of the OCCUPATION chooser in the *WageIndicator* questionnaire. It is used for the question: 'What is your occupation'. We used the revised *International Standard Classification of Occupations* (ISCO88), which was approved by the 14th International Conference of Labour Statisticians 1987 of the International Labour Office (ILO), see <http://laborsta.ilo.org/>. The ISCO88(COM) version is adapted for use in European countries, see also at the site of the University of Warwick <http://www2.warwick.ac.uk/fac/soc/ier/research/isco88/>. The ISCO88(COM) is used by some national statistical agencies in Europe, whilst other statistical agencies have correspondence tables available. In international surveys, the classification is frequently used. The ISCO-codes can be converted to the International Socio-Economic Index of Occupational Status.

The ISCO88(COM) classification has 544 partly overlapping items, covering four levels of aggregation.⁴ The OCCUPATION chooser uses all ISCO88(COM) codes at the third and fourth level, but has added approximately 2000 more detailed occupations at the fifth level. This is done for several reasons: (1) ISCO88(COM) identifies service occupations, new ICT occupations and a few other occupational groups not to a very detailed extent; (2) distributions over ISCO occupational classes are very uneven, as national Labour Force Surveys show, urging to detail the large classes; (3) the division of labour within occupational groups across countries varies substantially, leading to detailed classification in some countries whereas this is not applicable in other countries. Finally, our four-year experience with the Dutch *WageIndicator* questionnaire shows that respondents feel a strong need to be much

⁴ Level 1 : 10 major groups identified by one-digit numerical codes (1 to 0);
Level 2 : 28 sub- major groups identified by two-digit numerical codes (10 to 93 + 1);
Level 3 : 116 minor groups identified by three-digit numerical codes (100 to 933);
Level 4 : 390 unit groups identified by four-digit numerical codes (1110 to 9330).

more precise about their occupation than about their industry. Therefore, the master list of all occupations has more than 2500 items, making the fifth level of aggregation a better approach of reality.

The OCCUPATION chooser is a three-step chooser with 21 first-step items and 113 second-step items, and the master list of 2500 occupations for the third step, sorted by alphabet. It is available in the following languages: DK, EN, ES, GE, FI, IT, NL, PO for all 390 unit groups at the 4th ISCO level of aggregation. At the 5th level of aggregation, however, many of the 2000 occupations are available in only one or a few languages. The main reason is that certain occupations were either not at all or not to a very detailed level found in all countries. In these cases, the items of the master list remain empty and the chooser will not display them. The number of occupations varies largely across countries. The Polish OCCUPATION chooser has some 700 occupations, the Belgian chooser has 980 occupations in French and Dutch, the German chooser has 650 occupations, the Finnish and the Spanish chooser have some 600 occupations, the Italian and the English chooser have some 1250 occupations, whereas the Danish chooser has approximately 700. These numbers can change in case occupations are added to the chooser.

If a respondent cannot find its occupation, they are asked to tick nevertheless an item that comes closest or to tick 'other', and then use a text field to be more precise about their occupation. We expect to be able to trace 'new' or 'not-listed' occupations here. We aim to make a summary of these occupations after the first data-release and discuss the list with the national research partners.

The *WageIndicator* dataset has the codes for the master list of the OCCUPATION chooser. The value labels are in English, as far as the occupations could be translated. Otherwise they remain in their original language. The variable OCCUPAT_I is numerical. The occupation-groups at step 1 and 2 have a negative code and the occupations at step 3 have a 9-digit code, of which the first four digits are based on the 4-digit ISCO88(COM)code. If a respondent quits the chooser after step 2, and not completes step 3, the codes of the root are saved for the dataset.

7. THE ISO, UN AND WORLDBANK CLASSIFICATION AND THE COUNTRY CHOOSER

The UN classification is the basis of the COUNTRY chooser in the *WageIndicator* questionnaire. It is used for the questions: 'Please select the country where the foreign investment comes from' and 'In which country were you born and in which country were your mother and father born?'. The former is only asked when working for a foreign-owned company. The latter is only asked when ticked 'other country' in case the respondents self or their parents were not born in the country of survey.

The classification of countries is fairly standardised. The International Organisation for Standardisation (ISO, www.iso.org) sets the standards and assigns numerical and alphabetical codes to countries and territories. ISO assigns three codes to each country or territory: a two-letter code, a three-letter code, and a three-digit code.⁵ In June 2004, ISO-3166-1 included 240 countries, including not-inhabited territories. The Statistics Division of the United Nations (UN, www.un.org/) presents a list of the names of countries, using the 3-digit alphabetical ISO-codes. This list includes 232 countries. The World Bank Group (www.worldbank.org/) classifies economies according to gross national income (GNI) per capita. In July 2004, their list included 208 countries. For these countries, the World Bank has economic data available.

For the COUNTRY chooser, we didn't use the full ISO-classification, because it includes 10 countries that are of no use for our purpose, such as Antarctica or the Cocos Islands. We didn't use the World Bank-classification, because for our purposes too many countries are not included. We departed from the UN-classification. This classification includes two islands that are not present in the ISO list, quite likely because they are considered part of Great Britain (Isle of Man⁶ and the Channel Islands⁷). We included these two islands, because of their importance for the question of the foreign-owned companies. The UN-classification has two islands countries with very few inhabitants, notably Svalbard and Jan Mayen, and Tokelau. We excluded them from the COUNTRY chooser. Finally, neither ISO nor the UN

⁵ The numeric code identifies a physical territory, and the letters a country name. Thus when Germany reunified, it kept DE/DEU but changed from 280 to 276, while when Southern Rhodesia was renamed Zimbabwe it stayed 716 but changed from RH/RHO to ZW/ZWE.

⁶ Assigned code Z1.

⁷ Assigned code Z2.

includes Taiwan⁸ in their classifications. For the purpose of measuring country of birth, we decided to include Taiwan in our list, because of its size and the large numbers of immigrants from this island.

In total, the COUNTRY chooser consists of 236 countries, subdivided into 5 continents and 17 sub-continents. Per sub-continent, the countries are sorted alphabetically. The names of the countries are not always in line with the ISO or UN classification, because the chooser should facilitate web visitors to search easily for the name of the country. Therefore, instead of for example the name 'Netherlands Antilles' we rather use 'Antilles, Netherlands', because we assume that someone will search first on the A of Antilles and only if not found look for the N of Netherlands.

In the *WageIndicator* dataset, the countries have the three digit numerical ISO code. The value label includes the two-letter ISO code and the full country name, for example 'NA Namibia'.

8. THE NUTS CLASSIFICATION AND THE REGION CHOOSER

Three questions in *WageIndicator* questionnaire address regions, notably region of residence, region of workplace, and region of birth. To identify regions, a REGION chooser has been made, based on the NUTS regional classification from the RAMON server of Eurostat <http://europa.eu.int/comm/eurostat/ramon/>. The NUTS classification is a three-level hierarchical classification, identifying regions and sub-regions within the European countries. The question is 'In which region do you live?', followed by the REGION chooser. The following questions ask: 'Is your workplace also in this region?' and 'Are you born in the region where you currently live?'. If answered no, the REGION chooser pops up again.

Each national version of the *WageIndicator* questionnaire has its own country-specific REGION chooser. Therefore, the items in the REGION chooser are only available in the language(s) of the country. Countries with two languages, such as Belgium, have language specific versions of the country-specific REGION chooser. In order to trace boundary work and boundary residence, the names of the surrounding countries have been added to the REGION chooser. These country names are in the national language(s) of the country.

⁸ Assigned code Z3.

In the *WageIndicator* dataset, the regions have a numerical code, composed of the three digit numerical ISO code, a follow-up number and the numerical part of the NUTS codes. The first numbers indicate the country, using the ISO code, for the United Kingdom for example 826. The following numbers indicate the regions, using a follow-up code. The last numbers indicate the sub-regions, using the numerical part of the NUTS code. For example, the UK sub-region Brighton and Hove with the NUTS code UKJ21 has the *WageIndicator* code 8261021. The value label includes the two-letter ISO code and the NUTS region code, in this case 'UKJ21 Brighton and Hove'.

In many continental European countries, respondents may have their regions of workplace and residence in bordering countries. To tackle this phenomenon, the REGION chooser does include a list of neighbouring countries for each country of survey. The neighbouring countries are coded with the three digit numerical ISO code for the country of survey, followed by the number 99 for neighbouring country, and then followed by the three digit numerical ISO code for the neighbouring country. For example, Ireland is a neighbouring country of the United Kingdom and it is therefore coded 82699372. The value label includes the two letter ISO code for the country of survey, Z indicating a neighbouring country, the two letter ISO code for the neighbouring country and the full country name, in this case 'UKZIE Ireland'.

APPENDIX PRINCIPAL CHARACTERISTICS OF THE SIX LEVELS IN THE ISCED EDUCATIONAL CLASSIFICATION

LEVEL 1 BASIC EDUCATION (PRIMARY SCHOOLS)

This level covers primary schools.

LEVEL 2 LOWER SECONDARY OR SECOND STAGE OF BASIC EDUCATION

The contents of education at this stage are typically designed to complete the provision of basic education which began at ISCED level 1. In many, if not most countries, the educational aim is to lay the foundation for lifelong learning and human development on which countries may expand, systematically, further educational opportunities. The programmes at this level are usually on a more subject-oriented pattern using more specialized teachers and more often several teachers conducting classes in their field of specialization. The full implementation of basic skills occurs at this level. The end of this level often coincides with the end of compulsory education where it exists.

LEVEL 3 (UPPER) SECONDARY EDUCATION

This level of education typically begins at the end of full-time compulsory education for those countries that have a system of compulsory education. More specialization may be observed at this level than at ISCED level 2 and often teachers need to be more qualified or specialized than for ISCED level 2. The entrance age to this level is typically 15 or 16 years.

The educational programmes included at this level typically require the completion of some 9 years of full-time education (since the beginning of level 1) for admission or a combination of education and vocational or technical experience and with as minimum entrance requirements the completion of level 2 or demonstrable ability to handle programmes at this level.

LEVEL 4 POST-SECONDARY NON-TERTIARY EDUCATION

ISCED 4 captures programmes that straddle the boundary between upper-secondary and post-secondary education from an international point of view, even though they might clearly be considered as upper-secondary or post-secondary programmes in a national context.

ISCED 4 programmes can, considering their content, not be regarded as tertiary programmes. They are often not significantly more advanced than programmes at ISCED 3 but they serve to broaden the knowledge of participants who have already completed a programme at level 3.

Typical examples are programmes designed to prepare students for studies at level 5 who, although having completed ISCED level 3, did not follow a curriculum which would allow entry to level 5, i.e. pre-degree foundation courses or short vocational programmes. Second cycle programmes can be included as well.

LEVEL 5 FIRST STAGE OF TERTIARY EDUCATION (NOT LEADING DIRECTLY TO AN ADVANCED RESEARCH QUALIFICATION)

This level consists of tertiary programmes having an educational content more advanced than those offered at levels 3 and 4. Entry to these programmes normally requires the successful completion of ISCED level 3A or 3B or a similar qualification at ISCED level 4A.

All degrees and qualifications are cross-classified by type of programmes, position in national degree or qualification structures (see below) and cumulative duration at tertiary.

LEVEL 6 SECOND STAGE OF TERTIARY EDUCATION (LEADING TO AN ADVANCED RESEARCH QUALIFICATION)

This level is reserved for tertiary programmes which lead to the award of an advanced research qualification. The programmes are therefore devoted to advanced study and original research and are not based on course-work only.