



RESEARCH PAPER

No 35

**User guide to developing
an employer survey
on skill needs**



User guide to developing an employer survey on skill needs

A great deal of additional information on the European Union is available on the Internet.

It can be accessed through the Europa server (<http://europa.eu>).

Cataloguing data can be found at the end of this publication.

Luxembourg: Publications Office of the European Union, 2013

ISBN 978-92-896-1359-0

ISSN 1831-5860

doi: 10.2801/25552

© European Centre for the Development of Vocational Training (Cedefop), 2013
All rights reserved.

The **European Centre for the Development of Vocational Training** (Cedefop) is the European Union's reference centre for vocational education and training. We provide information on and analyses of vocational education and training systems, policies, research and practice. Cedefop was established in 1975 by Council Regulation (EEC) No 337/75.

Europe 123, 570 01 Thessaloniki (Pylea), GREECE
PO Box 22427, 551 02 Thessaloniki, GREECE
Tel. +30 2310490111, Fax +30 2310490020
E-mail: info@cedefop.europa.eu
www.cedefop.europa.eu

James J. Calleja, *Director*
Laurence Martin, *Chair of the Governing Board*

Foreword

The capacity to prevent skill gaps and shortages impacts on the competitiveness of the European economy, particularly when one considers the contribution of enterprises and their investments towards a more rapid economic recovery. Opportunities to cope with the present crisis will be greatly extended if efficient systems to identify and anticipate skill needs both at European level and in Member States will be available.

The EU initiative New skills for new jobs provides a context for activities which cover the analysis of skill demand, supply and skill mismatch. Associated is the idea for developing an employer survey on current and future skill needs in Europe. The resolution of the Council of the European Union on new skills for new jobs from 15 November 2007 (Council of the EU, 2007) emphasises new opportunities for citizens in Europe to improve their knowledge, skills and competences. There is a need better to align skills with the needs of both society and the economy, and to anticipate skill needs and identify skill gaps that may emerge in European labour markets.

Information concerning skills and their development is largely derived from household surveys covering areas such as labour market trends and data on trends in occupations and sectors from which conclusions on skill demand and supply are drawn. Detailed and regular analysis of cross-sectional data obtained from individuals over time indicates substantive structural changes but not the implications in terms of skill needs. For a clearer understanding of these complex issues one has to go beyond structural data to qualitative aspects of skill requirements and their development in the workplace.

The European Commission has entrusted Cedefop to develop a pan-European employer survey on skill needs. In close collaboration with experts, Cedefop has evaluated innovative approaches for the measurement of employer's skill needs. The principal aim of the pilot study was to test a task-based approach using questions on importance and related changes for both generic and specific skills in a range of sectors and occupations. In addition, information was collected on newly-emerging tasks and on drivers of change at the workplace, such as innovation and adaptation to environmental regulations. Finally, the study explored questions concerning the preparedness of the workforce to meet new requirements together with policies adopted by firms to address these areas of concern.

The present Cedefop publication is part of a toolkit, including the questionnaire and anonymised data set, for researchers to carry out related

employer surveys in other international or national contexts, and possibly in sectors and occupations not considered in the study. It illustrates an approach for identifying employer's perceived current and future skill needs tested in the pilot survey in nine Member States. The instrument is the result of development work carried out 2010 and 2011 and of pre-tests in 2011, and finally a large pilot survey. The publication offers conceptual and design considerations as well as specific recommendations for preparing a practical employer survey, including suggestions for sample size calculations, maximising response rates, and ensuring the validity of the instrument and subsequent inferences derived from it.

The development of an employer survey on skill needs (in Europe) makes a valuable contribution to linking the world of employers with that of education and training, and vocational education and training (VET) in particular. The cooperation of employers with those who shape VET policy, as well as the educators and trainers who develop and implement VET programmes, is needed more than ever. It is our hope that in providing these materials for further use we can contribute to this process.

Christian F. Lettmayr
Deputy Director

Acknowledgments

This publication is the result of a team effort and reflects the contributions of all involved in the project. Nick Sofroniou and Alena Zukersteinova from Cedefop supervised the study and edited and prepared the present report. Special thanks go to Pascaline Descy (Head of Area research and policy analysis) for her guidance and advice as well as to Cedefop experts Giovanni Russo and Marco Serafini for their invaluable ideas and support during the project. This publication also relies on the initial contributions and advice of Manfred Tessaring and Katja Nestler.

The publication is a result of the development work under a contract between Cedefop and TNS Infratest Sozialforschung (contract number 2009-FWC56/AO/RPA/AZU-TODUN/Pilot-European-Employer-Survey/015/09). The development work benefited considerably from contributions by Arnold Riedmann and Miriam Gensicke (TNS Infratest Sozialforschung, Munich), Bernd Dworschak and Alexander Karapidis (Fraunhofer Institute for Industrial Engineering, Stuttgart), Francis Green (LLAKES Centre, Institute of Education, London) and their expert advisory group.

Cedefop would like to acknowledge the Skillsnet members for productive debates, thoughtful inputs and useful ideas. Important suggestions were also made by Cedefop expert Pedro Moreno Da Fonseca who reviewed this publication, as well as Vladimir Kvetan, Konstantinos Pouliakas and Antonio Ranieri. Assistance from Roula Panagiotou is also much appreciated.

This paper is part of Cedefop's ongoing work on early identification of skill needs. In this context, Cedefop's activities of skill forecasting and exploring employer surveys are supported financially by the Progress programme, European Community programme for employment and social solidarity (2007-13), managed by the European Commission, Directorate General for Employment, Social Affairs and Inclusion.

Table of contents

Foreword.....	1
Acknowledgments	3
CHAPTER 1. Introduction	9
CHAPTER 2. The survey concept in brief – Key characteristics and fields of application	11
CHAPTER 3. Choice of the data collection methodology	14
CHAPTER 4. Sampling principles, survey universe and definition of respondents 16	
4.1. Sampling principles	16
4.1.1. Design of the sampling matrix	16
4.1.2. Setting of targets for the sampling matrix	18
4.1.3. Sampling sources	19
4.1.4. Structure of the net samples	20
4.2. Coverage of size-classes of organisations.....	21
4.3. Coverage of sectors	21
4.3.1. Which and how many sectors of activity are of interest for the survey	23
4.3.2. Sectors to be included – Criteria for their choice	24
4.4. Coverage of occupational groups	28
4.4.1. Pre-selection of occupational groups for the pilot survey.....	28
4.4.2. Why (pre-)selection is necessary	29
4.4.3. How many occupational groups can be covered	30
4.4.4. How to pre-select the occupational groups to be covered	30
4.5. Definition of the sampling unit and unit of enquiry.....	32
4.6. Definition and identification of the right respondent	33
CHAPTER 5. The questionnaire	35
5.1. The pilot master questionnaire and proposals for modifications.....	35
5.1.1. Contact phase (Q001 to Q005)	35
5.1.2. Screening phase (Q050 to Q099).....	36
5.1.3. Module 1: background information on the establishment, selection and characterisation of an occupational group	37
5.1.4. Module 2: generic task questions.....	39
5.1.5. Module 3: occupation-specific tasks and newly-emerging tasks	41
5.1.6. Module 4: drivers of changing the tasks: innovations in the establishment.....	42

5.2.	The mechanism for selecting an occupational group	43
5.2.1.	The selection process applied in the pilot survey	43
5.2.2.	Proposal for modification of the selection procedure	44
5.3.	Hints for the interviewer briefing	45
5.3.1.	Contact phase.....	46
5.3.2.	Q100c/Q110	47
5.3.3.	Q102: verification of the sector classification.....	47
5.3.4.	Q106b.....	49
5.3.5.	Q107_1a/b to Q107_3a/b.....	49
5.3.6.	Q110.....	50
5.3.7.	Q113.....	50
5.3.8.	Q114 to Q116	50
5.3.9.	Q200 et seq.	50
5.3.10.	Q200/Q201	52
5.3.11.	Q202/Q203	52
5.3.12.	Q200, Q202, Q204, ... Q234.....	52
5.3.13.	Q201, Q203, Q205, ... Q235.....	53
5.3.14.	Q240.....	53
5.3.15.	Q242.....	53
5.3.16.	Q300/Q301	53
5.3.17.	Q303.....	53
5.3.18.	Q304.....	54
5.3.19.	Q305.....	54
5.3.20.	Q306.....	54
5.3.21.	Don't know	54
5.3.22.	Don't know/no answer.....	54
CHAPTER 6.	Translation	55
6.1.	General translation guidelines	55
6.2.	Translation of official terminologies.....	57
6.3.	Specific challenges in translating the questionnaire.....	58
6.3.1.	Variable text elements.....	58
6.3.2.	Complex programming instructions	58
6.4.	Translations in case of repeating the survey.....	59
CHAPTER 7.	Programming and testing the questionnaires	60
7.1.	Specific challenges for programming.....	60
7.2.	Organisation of programming	61
7.2.1.	Specific hints on programming the CATI instrument.....	61
CHAPTER 8.	Fieldwork: timing, organisation and monitoring	63
8.1.	Forms of survey organisation	63
8.2.	Time planning.....	64
8.3.	Fieldwork monitoring	65

8.3.1. Weekly progress reports	65
8.3.2. Non-response information per cell of the sampling matrix	66
8.3.3. Final field reporting	67
8.4. Measures to improve response rates.....	68
8.5. Specific requirements for conducting the survey as a repeat survey...	70
CHAPTER 9. Data handling.....	71
9.1. Data checks	71
9.1.1. Central check of interim data sets	71
9.1.2. Checks of final data set.....	71
9.1.3. Data and consistency checks embedded in the questionnaire script.....	72
9.2. Coding of open-ended answers.....	72
9.2.1. Q304: newly-emerging tasks (not mentioned in the previous task questions).....	72
9.2.2. Q306_4: other measures applied (to address newly- emerging tasks for the occupational group).....	73
9.3. Data outputs.....	74
CHAPTER 10. The weighting	75
10.1. Organisation of the weighting process.....	75
10.2. Establishment-proportional weighting	75
10.2.1. Basic structure of the weighting process	76
10.2.2. Practical implementation of the weighting process	79
10.3. Employee-proportional weighting	79
10.3.1. Employee proportional weights as an additional option	79
10.3.2. Option 1: use of the numerical value on number of employees in the occupational group as factor.....	80
10.3.3. Option 2: recoding the numerical values of Q110 into size bands.....	80
10.4. Statistical information required for weighting	82
10.5. Design effects, effective sample size, weighting and standard errors .	82
CHAPTER 11. Possibilities for analysis of the data.....	85
11.1. Univariate descriptive analysis	85
11.2. Pairwise analysis.....	86
11.2.1. Generic tasks.....	87
11.2.2. Occupation-specific tasks	88
11.2.3. Newly-emerging tasks.....	88
11.3. Analyses with multiple variables	90

CHAPTER 12. Conclusions	93
12.1. General assessment of the survey concept.....	93
12.2. Key issues for sampling, programming and data collection	94
12.2.1. Sample size and coverage of the survey.....	95
12.2.2. Unit of enquiry.....	96
12.2.3. Data collection mode.....	97
12.3. Outlook: possibilities and restraints for any replications of the survey concept	97
List of abbreviations	101
References.....	102
Annex 1. Final draft questionnaire ‘pilot employer survey on skill needs in Europe’	103
Annex 2. Proposal for a motivation letter.....	138

List of tables

1. Levels of differentiation in the ISCO-08	13
2. Proposal for the sampling and weighting matrix.....	18
3. Aspects to be considered for application of size thresholds for the survey universe	22
4. Occupational groups addressed in the pilot survey.....	51
5. Time frame needed for survey preparation	64
6. Reporting template for weekly progress reports (Germany) Only the fields labelled 'actual' have to be completed; the targets are fixed, the % is calculated automatically	66
7. Template for recording final non-response reasons	67
8. Weighting matrix.....	79

CHAPTER 1.

Introduction

This manual is intended for institutions, researchers or fieldwork agencies that are planning to implement an employer survey on skill needs and can be inspired by the model developed by Cedefop. The aim of the manual is to provide scientists and survey practitioners with comprehensive information on the preparation and implementation of this challenging and innovative survey. In the implementation of the survey concept, different options and variations are possible, depending on the main aim of the survey, the available resources and the type of employers that are of interest. The manual discusses these options and their implications with the present survey instrument as a detailed example.

The development of the survey stretched over three project phases, starting in January 2010 and ending in November 2012. In the course of this development phase, the survey concept and draft questionnaire versions were repeatedly discussed with a broad array of skill needs experts, among them researchers, policy-makers and representatives of the employers. In addition, the survey was tested in practice in two stages: in summer 2011, a preliminary version of the concept was evaluated in five EU Member States (Germany, Ireland, Spain, Poland and Finland) by way of 125 standardised interviews and an additional 127 cognitive pre-test interviews. At the beginning of 2012, a further refined version of the concept and questionnaire were tested in a large-scale pilot survey with more than 8 500 interviews conducted in a selection of seven sector clusters in nine Member States (Czech Republic, Germany, Ireland, Spain, France, Italy, Hungary, Poland and Finland).

Apart from a presentation of the general survey concept and its aims, the information in the manual comprises concrete hints and instructions on all relevant technical and organisational aspects of the implementation of the survey, such as the sampling, programming, interviewer briefing, fieldwork surveillance and data processing.

The manual is largely structured in accordance with the timing of the different steps in the project cycle. It starts with the choice of the data collection methodology and some general considerations about the universe to be covered. Then, basic characteristics of the questionnaire are presented. Finally, the subsequent steps of fieldwork are detailed, starting with the translation of the master questionnaire and the programming of the script. A further chapter is dedicated to the organisation of fieldwork, including considerations about the time span required for implementation of a survey of this kind. The handbook then

offers recommendations regarding data handling and weighting of data. The latter is an indispensable step for this survey because of its disproportional sampling design. Finally, some possible analyses are delineated.

CHAPTER 2.

The survey concept in brief – Key characteristics and fields of application

The pilot employer survey on skills is a representative survey among employers. It aims at identifying skills, competences and qualifications needed at the workplace, from the perspective of employers. Its main focus is on working tasks performed at the workplace, their change in importance and the preparedness of the workforce to cope with tasks that are becoming more important. The survey concept aims at combining the collection of background information on possible drivers of change at the level of the whole organisation with the collection of information on qualitative changes regarding the tasks to be carried out by particular occupational groups within the organisation.

To this end, at the beginning of a survey interview session one particular occupational group is selected as reference group for all questions related to particular tasks. In the pilot survey two sets of tasks were differentiated. First, a set of generic or transversal tasks with some relevance for almost all types of workplaces is asked about. Though being identical for all occupational groups, these so-called generic-task questions have to be answered for one (previously selected) occupational group only, because in the application of the tasks there may be large differences between occupational groups. Second, for certain occupational groups a set of occupation-specific tasks that are normally characteristic for the respective group are asked about.

In sum, the master questionnaire tested in the pilot survey consists of the following four questionnaire parts or modules:

- Module 1: background information on the organisation; the selection of a particular occupational group for reference in most further questions;
- Module 2: questions on the importance and development of a series of generic tasks (more general tasks that are applicable to many types of jobs and workplaces) for the selected occupational group;
- Module 3: importance and development of a series of occupation-specific tasks (tasks that are characteristic for a particular type of job) and emergence of any new tasks, again for the selected occupational group only;

Module 4: background questions on major changes and innovations introduced in the organisation and their impact on the selected occupational group.

Generally, the survey could be carried out either with this combination of generic and occupation-specific task questions or it might be carried out with just one of these question blocs, either with the generic or with the occupation-specific question module. If replicating the survey with just one of these types of task-related questions, some modifications are, however, recommended: when concentrating on the set of generic questions only, it is worth considering broadening this set by taking on some aspects from the occupation-specific questions and trying to formulate them as generic questions. When asking occupation-specific questions only, in order to get a full picture, the sets of occupation-specific task questions need to be broadened by additionally taking up some crucial generic tasks and adapting them to the respective occupational groups. For any option including occupation-specific questions, it has to be noted that occupation-specific questions have so far been drafted for seven occupational groups only. For further occupational groups to be included in a replication of the survey, this type of question has to be newly developed. In addition, some items for the existing occupation-specific questions need to be reformulated because they turned out to be too complex and thus difficult to understand.

The selection of one particular group of employees as reference group for the majority of survey questions is an important and innovative feature of the survey. In the pilot survey, the selection of occupational groups was made based on the ISCO 3-digit level (ISCO-08 minor groups, see Table 1). An occupation is hereby defined in accordance with the official ILO definition as 'a set of jobs whose main tasks and duties are characterised by a high degree of similarity' (ILO, 2007, p. 1). A main reason for using ISCO in the survey is that as other skill needs forecast activities at European level also use ISCO the different forecast instruments can better complement one another.

The definition and selection of occupational groups can in principle also be made at a broader level such as the ISCO-08 major groups (ISCO-08 1-digit level) or submajor (ISCO-08 2-digit level) or in turn at a finer level such as the unit groups (ISCO-08 4-digit level). However, this would have major implications for the survey in terms of the expected outcomes, the sample size and inherent measurement errors and has not been tested empirically in the context of the pre-test or pilot conducted in this project. Details of how to carry out the survey on a different ISCO level are therefore not described in this manual.

Table 1. Levels of differentiation in the ISCO-08

Name of the ISCO differentiation	Level	Number of groups differentiated at that level
Major groups	1-digit	10
Submajor groups	2-digit	43
Minor groups	3-digit	130
Unit groups	4-digit	436

Source: Cedefop pilot survey 2012.

When carried out as a snapshot survey collecting cross-sectional data at one particular point in time, the employer survey on skill needs can still provide some interesting insights. For example, into skills and competences which in the future should be emphasised more in the vocational training of a particular occupational group or even in the general schooling system (if the issue concerns the working population in general).

The survey will gain a further explanatory power if it is repeated at regular intervals, for example every five years. This will enable researchers to detect any changes over time and thus be able to react quickly by designing adequate policy measures as response to the observed developments. Moreover, the design of the survey as a repeat survey will ease evaluation of the measures taken, for example of specific programmes directed at improving particular types of skills among pupils or employees. Another advantage is that only with a repeat survey will it be possible to infer clearly causation from the data. Based on cross-sectional data, even statistical analyses with multiple variables can only detect associations between certain phenomena, but cannot clearly allow one to infer that one phenomenon causes the other. With longitudinal data from a repeat survey, this type of insight becomes possible.

Such a repeat survey does not necessarily have to be designed as a panel survey, asking exactly the same establishments again in each further wave. Though such a panel survey allows for further analysis, it is a very costly concept, particularly if setting up such a panel survey in several countries simultaneously. As long as the sample definition and other important parameters are not changed between the waves, the repetition with different sets of establishments each time is a good enough option and it avoids problems due to attrition of the sample over time.

CHAPTER 3.

Choice of the data collection methodology

The employer survey on skills developed in this project, requires a computer-assisted data collection methodology such as CATI (computer-assisted telephone interviewing), CAPI (computer-assisted personal interviewing) or CAWI (computer-assisted web interviewing). For several reasons, it is not suitable for paper and pencil techniques, including the following:

- (a) identification of the selected group is not made in advance, but only during the interview. The different occupation-specific questionnaire modules (variants of questions Q300 and Q301 in Module 3 of the pilot master questionnaire) could therefore not easily be administered in the paper methodology since this would require the provision of all different variants of these questions to the respondent (or to the interviewer in case of paper-assisted personal interviews). They would then manually have to select the adequate set of questions, a procedure that would likely be very error-prone;
- (b) process of randomly selecting one occupational group also requires a computer-assisted data collection methodology. The selection mechanism is too complicated for an error-free application in a self-administered paper questionnaire;
- (c) once a particular occupational group is selected, a clear reference to this selected occupational group has to be ensured throughout all parts of the questionnaire that are directed at the selected group only. In the computer-assisted methodologies, this is made by fading in the name of the respective group selected earlier in the questionnaire. In a self-administered paper questionnaire, such a clear, repeated reference would not be possible and it would remain unclear whether the respondent really referred questions correctly throughout the interview.

For the pilot, the CATI methodology was preferred over other computer-assisted data collection methodologies such as CAPI and CAWI.

Carrying out the survey with face-to-face CAPIs is in general possible and may even lead to slightly higher data quality. But the CAPI methodology requires a considerably longer fieldwork period and implies much higher fieldwork costs. Similarly, face-to-face interviewing in organisations is not common in some countries. In these cases adequate face-to-face interviewers may therefore not be available as an option.

Online interviewing (CAWI) can currently not be recommended for this survey due to major sampling difficulties. The main problem in this context is that representative registers of e-mail addresses of enterprises or establishments are not currently available in most Member States. And even where such registers are available, the method is likely to face a very high and selective non-response if it is not accompanied by an interviewer-based method in the recruitment phase and for reminders. As long as the sampling and non-response problems of CAWIs among organisations (sampled from general address registers) are not solved, the CAWI alone is therefore not an option worth considering for the survey. Though it may be useful when incorporated as part of a hybrid approach involving, for example, initial telephone contact or mixed mode using CATI to fulfil the sample.

CHAPTER 4.

Sampling principles, survey universe and definition of respondents

To provide findings that are representative for the defined universe, the survey has to be conducted in a randomly selected set of organisations within the country or countries to be covered. In the definition of the universe, the sampling unit and the respondents for the survey, there is some leeway for a replication of the survey – it is not imperative to apply exactly the same sample design as for the pilot survey. The following paragraphs are based on the design applied to the pilot survey, but they also discuss some possible variations.

4.1. Sampling principles

For any replication of the survey, the application of a stratified random sample design is strongly recommended. In this sample design, the targeted universe is divided into a number of cells defined by key criteria such as the size-class, the sector of activity, the country or other territorial differentiations (sampling or stratification matrix). Within each cell, addresses have to be drawn at random from an address source that lists the organisations of the defined universe. Such a probability sample helps ensure the representativeness of the sample and the robustness of inferences drawn, unlike a quota sample.

4.1.1. Design of the sampling matrix

The stratification matrix is an important tool to ensure an adequate sampling procedure at all stages. If based on relevant correlated indicators, stratification in general leads to more precise estimators. If non-relevant criteria are chosen stratification is merely ineffective but will cause no damage. Moreover, in the special case of purposely disproportional designed samples the stratification table is not only a simple allocation formula, but also an essential tool for controlling and steering the total selection process. Hence it is recommended to use this matrix for drawing the gross sample from an address register, for the management of the sample during fieldwork and as framework for the later sampling ex-post: during the weighting. For all three steps, the matrix should be defined by exactly the same dimensions to ensure consistency of the sampling and weighting process.

Within each participating country, the sampling matrix to be used for the survey can best be defined by two dimensions: a differentiation by size-class (number of employees) and a differentiation by sector of activity. Though the sector differentiation is indispensable only if the sample is going to be disproportional with regard to sectors, it is recommended to use the sector as a stratification dimension also in samples that are meant to be proportional with regard to sectors of activity. The reason is that unit non-response rates can differ largely between different types of sectors so that a sector-proportionally drawn gross sample may well result in a net sample with rather disproportional sector structures, if the sector is not considered in the matrix used for steering the net sample.

The concrete shape of the sampling matrix should be adapted first to the net sample size and second to the definition of the universe in terms of sectors and size-classes to be covered. The total number of cells defined by the sampling matrix should not be too large, since the larger the number of cells, the more time-consuming and expensive fieldwork tends to become. However, the number of cells needs to be large enough to:

- (a) enable disproportional drawing of addresses for all dimensions for which disproportional drawing is considered important;
- (b) separately steer all key dimensions which can be supposed to be characterised by substantial differences in the degree of unit non-response.

As a rule of thumb, our recommendation is that in none of the cells should there be fewer than 30 interviews, though some exceptions to this rule might be necessary for the largest size-classes. A consideration of the minimum standard error required for an estimate may increase this quantity further as part of sample size calculations at the design stage.

As for the size-classes, a relatively rough categorisation into four or maximum five size-classes is sufficient for most purposes. The main aim of the stratification into size-classes is to avoid the large majority of interviews being conducted in organisations belonging to the smallest size-class(es) since the number of small organisations is in most sectors considerably higher than the number of middle-sized or large organisations. In terms of employment impact, the middle-sized and large organisations are more substantive than the small ones.

The four size-classes used in the pilot survey (5 to 9, 10 to 49, 50 to 249 and 250 or more employees) reflect standard size-class boundaries and are used in many European-wide employer surveys. However, a modification of the size bands is possible, but has then to be made uniformly for all participating countries. If envisaging an alternative division of size bands, it also has to be

assured beforehand that the statistical figures required for the weighting are available in the envisaged alternative breakdowns.

The pilot survey was conducted in a subset of sectors and not economy-wide. The sector dimensions in the model sampling matrix shown below (Table 2) reflect the seven sector clusters chosen for the pilot survey. If replicating the survey in the same (or an even finer) degree of ISCO differentiation as used for the pilot (the ISCO-08 3-digit level), a relatively fine differentiation by sector is indispensable since the display of occupational groups in the question block dedicated to the selection of an occupational group for the further interview (Q107 in the pilot master questionnaire) is determined by the sector of activity. For each sector cluster, a different set of occupational groups is shown in the interview. If the survey is to be replicated on a considerably less differentiated level (such as the ISCO major groups respectively ISCO-08 1-digit level), sectors can be summarised in broader categories because at this higher aggregated level often the same occupational groups are dominating ⁽¹⁾.

The 28-cell matrix used for the pilot can serve as orientation for the set-up of a sampling matrix for this type of survey. The 28 cells represent the maximum degree of differentiation recommended for a national net sample of n =1 000 interviews (per country).

Table 2. **Proposal for the sampling and weighting matrix**

NACE Rev. 2 sector (cluster)	5 to 9 (or 1 to 9) employees	10 to 49 employees	50 to 249 employees	250 or more employees
Sector cluster 1				
Sector cluster 2				
Sector cluster 3				
Sector cluster 4				
Sector cluster 5				
Sector cluster 6				
Sector cluster 7				

Source: Cedefop pilot survey 2012.

4.1.2. Setting of targets for the sampling matrix

For each of the cells in the sampling matrix, targets for the number of net interviews to be achieved need to be set. The targets should be set centrally to ensure a similar distribution of the interviews over the cells of the sampling matrix in each country. Targets need, however, to be confirmed by the institutes

⁽¹⁾ There are 10 major ISCO groups, two of which are specific to certain sectors of activity. In many sectors, the 'ranking' of the three to five most frequent occupational groups will be the same.

responsible for fieldwork to avoid the definition of unrealistically high targets for difficult cells. Within each cell of this sampling matrix, sampling then has to be made strictly at random.

For setting the targets in the cells, the following aspects should be considered:

- (a) the targets for the size-classes should not be set proportional to the real distribution of organisations. Instead, it is recommended to apply a mixture of an establishment- and an employee-proportional sample. Compared to an establishment-proportional perspective, this means an oversampling of larger units and a considerable under-sampling of the units in the smaller size-class(es). For the remaining intermediate size-classes, an almost even distribution of interviews is proposed;
- (b) the targets for the sector clusters can either be defined representatively (reflecting the real size proportions of the sectors to be covered) or some sectors can deliberately be oversampled at the expense of others. An oversampling of certain sectors may be appropriate, for example, to have a sufficient number of interviews available for separate analyses in the smaller sectors. Which of the two variants is to be preferred depends on the research interests. If results are mainly meant to be used at sector level, then oversampling of the smaller sectors is advisable.

The gross sample to be acquired before the launch of the survey should be adjusted to the set targets. It should be dimensioned rather generously to have some reserve addresses available in case general response rates, or the response rates for a particular type of organisations, turn out to be lower than expected.

4.1.3. Sampling sources

Sampling for the survey has to be made from publicly available address registers of organisations. The register to be used should be the most complete and updated register available in the respective country. The option of first choice is normally the national statistical business register. These registers generally exist in all Member States, but in some countries (such as Germany and Spain) they are not made accessible for survey purposes. Also, these registers are often collected at the level of enterprises/companies, only. Some statistical offices do additionally maintain registers at the level of establishments/local units, which is the level recommended for any replications of the survey (Section 4.5).

If an adequate register is not available from the national statistical office, or if the register maintained by the statistical office is not the best and most updated one available in the country, alternatives have to be identified. Potential

alternatives may be registers of federal labour agencies, social insurance corporations or other institutions. If these are not available either, the registers of national or international commercial address providers such as Dun & Bradstreet, Schober or Bill Moss may be the only remaining alternatives.

In countries where no suitable address register is available at the level of establishments/local units from any source, the best available register of companies/enterprises should be used instead and a screening procedure then applied to obtain a random selection of establishments out of the company-based register. A proposal for this screening procedure is contained in the pilot master questionnaires provided in Annex I (questions Q050 to Q099).

Before drawing the sample, each planned address source has to be checked for its completeness respective to its representative coverage of the target population. For these checks, the available statistics on the distribution of organisations should be used and the sources for the entries in the register verified. If the only available sampling source has clear structural weaknesses (systematic undercoverage of certain groups of addresses, such as public organisations in the sector 'human health and social work activities'), then these have to be adjusted by an oversampling of the underrepresented types of units in the gross sample. Care is needed with some commercial registers which may emphasise firms with larger quantities of temporary workers. If the number of addresses available in the foreseen address source is not sufficient for this aim, further addresses for the undercovered types of units need to be added from an additional address source.

4.1.4. Structure of the net samples

During fieldwork, the sample has to be steered according to the targets set for each cell of the sampling matrix. The agreed sample distribution needs to be fulfilled, though it is recommended to leave some leeway for minor deviations from the targets to ease fieldwork. Some leeway is particularly important for countries that have to apply the screening procedure, or to use address sources with outdated size information which often needs to be corrected during the interview. In case of deviations between the number of employees as indicated in the address source and in the questionnaire, the number of employees indicated by the respondent in the questionnaire should be the reference because this (and not the information from the address source) is the size of unit the respondent also has in mind when answering the remaining questions.

4.2. Coverage of size-classes of organisations

In the pilot employer survey on skill needs, only organisations with a minimum of five (dependent) employees qualified for participation in the survey. This size threshold is not imperative. It can be varied to a certain degree or even totally omitted, instead including establishments of all sizes. However, it always needs to be ensured that the same threshold is applied throughout the survey, in all countries (if it is a multinational survey) and in all sectors of activity (if different sample sources need to be used for the sampling in different sectors).

There are arguments in favour of, and against, applying a size threshold for the organisations to be sampled. For any replication of the survey concept, the people in charge of the survey conceptualisation can decide on the definition of the threshold. Considerations in Table 3 can serve as a guide for this decision.

The definition of the universe as units with five or more employees, as was done in the pilot survey, represents a kind of compromise reconciling the two positions to a certain degree.

Once any size thresholds are defined, a multistratified random selection procedure will ensure that all size-classes within the defined universe of the survey are covered to a sufficient degree (see Section 4.1).

4.3. Coverage of sectors

The survey can generally be carried out either economy-wide or in just a selection of sectors of activity, depending on the research interests and on the available budget which in turn determines the achievable net sample size. If the affordable net sample size is rather low, a concentration on specific sectors is recommended. It has the advantage of considerably enlarging the statistical power for any findings on those sectors that are included.

The economy consists of a multitude of different sectors of activity. The NACE Rev. 2 classification system differentiates between 21 sections (NACE 1-digit level), 88 divisions (NACE 2-digit level), 272 groups (NACE 3-digit level) and as many as 615 classes (NACE 4-digit level). In each sector, different occupational groups are prevalent. In the pilot employer survey on skills, a set of 12 NACE divisions were covered (NACE Rev. 2 sectors 28, 29, 30, 41, 42, 43, 46, 47, 62, 64, 84 and 86). These sectors were grouped into seven sector clusters (NACE Rev. 28/29/39, 41/42/43, 46/47, 62, 64, 84, 86). Within each of these clusters, the same set of occupational groups was selected.

Table 3. **Aspects to be considered for application of size thresholds for the survey universe**

Arguments for the definition of a size threshold	Arguments against a definition of a size threshold
<p>The smaller the sampled organisations, the more likely none of the pre-selected occupational groups is present; organisations where none of the pre-selected groups is present need to be filtered to END. This leads to enhanced screening costs.</p>	<p>Setting a size threshold considerably reduces the size of the available universe; this is particularly the case in sectors where the majority of organisations are (very) small. With a size threshold, the number of available addresses for a sector might be critically low in some cases.</p>
<p>In very small organisations, there is often just one employee within the selected occupational group; in such cases, it is difficult for the respondent to answer the task and skill questions for the respective job or position without having the individual job holder in mind. Strengths and weaknesses of individual employees may therefore strongly influence the results in small establishments and may lead to a certain distortion of the results.</p>	<p>Tasks and skill requirements may differ considerably between establishments of different sizes. For example, engineers in large firms or doctors in hospitals may have different tasks than their colleagues working in small workplaces (e.g. work in the larger units might generally be more specialised while in the small units, more all-round skills may be required).</p>
<p>The very small organisations (with less than 10 employees) are often only poorly covered by the available address registers and even where they are generally covered, the address quality is usually much poorer than for the larger organisations. This is mainly due to the considerably higher mortality and relocation rate among these smaller units. If address sources in the countries to be covered by the survey are known to have only poor coverage of the small units, it is recommended to exclude these units totally by setting a general size threshold. Otherwise, distortions of the results due to selectivity of the entries in the address sources may result.</p>	<p>People are generally schooled and trained for working in all kinds of organisations employing people of their occupational group. The size of the organisations is normally not relevant for the schooling and vocational education and training system. From this point of view, the exclusion of organisations from a certain size is not fully justifiable.</p>
<p>Similarly, the general availability and quality of statistics on the number of organisations by sector and size-class is in some countries problematic for the small-sized establishments. In particular, there are countries where the official statistics do not differentiate between organisations with 0 employees (self-employed persons) and establishments with one to four or one to nine (dependent) employees. This causes problems for accurate weighting of the data.</p>	<p>In some countries, particular sectors may be dominated by very small units. Also, some occupational groups may be concentrated very much on small organisations. If setting a threshold, an important share of the employment in the respective occupational group may thus not be covered.</p>

Source: Cedefop pilot survey 2012.

The sectors of activity are not the core category of the pilot survey – the survey is first and foremost interested in the analysis of the situation of different occupational groups, as vocational education and training (VET) systems are normally oriented at occupational groups rather than sectors of activity. There are occupational groups which exist only in one particular sector of the economy, but most occupational groups coexist in various sectors of the economy. Though the work tasks and the skills required to exert them may differ in accordance with the

sector where an employee is employed, the basic education and training for an occupational group (as defined at the ISCO-08 4-digit level) is in most cases generally the same, regardless of the sector. ICT technicians, for example, can be found in a broad variety of different sectors. Nevertheless, they usually have the same general educational background, regardless whether being employed in a firm of the production sector, in construction or in services.

In spite of the occupational group (and not the sector) being the core category of interest, for the definition of the sample universe the sector of activity is the relevant category because existing address registers of organisations usually do not provide any information on the occupational groups that their workforce encompasses. The sector of activity therefore has to serve as a kind of proxy for the identification of the occupational groups that are of interest.

In some systems, firms can have an additional NACE code (a secondary code), which is in many cases the relevant one. This occurs because firms may have a secondary activity that represents a minority share of its turnover and that the statistical authority may want to account for. In a relevant number of cases the firm might be stating its (real) secondary activity as the main one (which may be easy due to common accounting). The usual reason for this is that owners and partners may derive advantages (tax, contributions, subsidies) from classifying the firm in a different category. When this occurs, it is important that the firm is classified in the survey with the relevant NACE code, since otherwise some bias may be incurred (especially in the case of small countries). A tactful enquiry with national authorities or with Eurostat about this subject is advisable.

4.3.1. Which and how many sectors of activity are of interest for the survey

Before launching the survey, which sectors of activity and which occupational groups are of interest have to be clarified. The survey can be carried out economy-wide, but it can also be concentrated on just a few specific sectors – for example on sectors suffering from skill shortages or sectors undergoing major changes. Sectors considered as being of little interest with regard to skill needs forecasts can in turn be excluded.

The role of sector experts is absolutely fundamental to provide a sound rationale to make these decisions. One of the key roles of experts is in providing insight into differing strategic options across sector segments, thus providing critical insights on skills evolution. In addition, they may provide useful information concerning relevant value-chain linkages between sector clusters (frequently trade and production are relevantly linked with possibilities for vertical integration) identifying dynamics in occupations that could otherwise be excluded from analysis.

If the survey is to deliver information for the (re-)shaping of the VET system in general and/or on the general schooling system, an economy-wide survey is the option of choice. The drawbacks of full coverage of the whole economy are, however, a high complexity of the survey instrument and the need for a very large sample to have enough interviews available for the analysis of each of the occupational groups covered.

A concentration of the survey on particular parts of the economy which are of specific interest has the advantage that the complexity of the survey can be reduced since a lower number of occupational groups have to be considered. This is particularly relevant in the preparation of the selected groups (selection of the groups and examples, translation, programming, etc.) to be shown (Q107 in the pilot master questionnaire) and in the handling and analysis of the data. A limitation of the focus of the survey on few selected groups also offers the possibility to go for a finer differentiation of occupational groups (such as for the ISCO-08 4-digit level instead of the 3-digit level) because, with a given budget, the number of observations per sector and thus per occupational group can be much higher. This, in turn, improves the precision of the measurements because respondents have to consider a less heterogeneous group of employees when answering the questions. It also improves the usefulness of the data for the revision of VET curricula and other rather specific policy measures.

Therefore, it is recommended to check thoroughly before the launch of the survey which sectors of activity are of interest for later data users. Any sectors that can be excluded *a priori* from the sample will lead to a higher number of interviews available for the remaining sectors and will thus improve the empirical base for these. However, care has to be taken not to exclude any sector in which one of the occupational groups forming the focus of the survey has a large quantitative importance. Otherwise, there is the danger that, based on survey data, VET curricula are being adapted to the needs of the included sectors only, neglecting potentially different needs in other sectors excluded from the survey but which also employ people of that occupational group.

4.3.2. Sectors to be included – Criteria for their choice

Regarding the choice of the sectors to be included, several aspects should be considered.

4.3.2.1. Content-related criteria

The research interests are the most important criterion for the definition of the sectors to be included in a replication of the survey. The following aspects could guide the choice of sectors:

- (a) which sectors are currently undergoing major changes in terms of organisation, the introduction of new technologies and the emergence of new tasks. Information on this can be extracted, for example, from other surveys, from expert workshops or from other skill needs forecasting activities;
- (b) which sectors are considered as key sectors for the future and for strengthening the competitiveness of the European economy;
- (c) which sectors have recently faced major skill gaps or skill shortages. Information on this can be derived, for example, from surveys on (hard-to-fill) vacancies.

4.3.2.2. *Size of the sector*

The sectors to be included in the survey should each be sufficiently large. If the sectors are too small, there may not be enough interviews available in the end for any solid statistical analysis. How large the universe of a sector needs to be as a minimum depends on the following factors:

- (a) the number of addresses available for the sector in each country to be included in the survey (particularly if private address sources have to be used, the number of available addresses may be considerably lower than the number of units that exist in the sector according to the available statistics);
- (b) the response and cooperation rates to be expected for surveys of this kind (as a general rule, the gross sample should be 10 times as large as the minimum net sample size needed for the analysis within a cell);
- (c) the quality of the address register to be used (if the best available register is outdated, more addresses are needed to accomplish a certain net sample size for the cell than with an up-to-date register of high quality);
- (d) the number of occupational groups that are meant to be pre-selected as reference group(s) for the survey questions about tasks and skills. If five groups are to be pre-selected and one of them is chosen for the interview, a larger gross sample is needed for the sector than if just two or three groups are to be pre-selected. For example, if 100 net interviews are carried out in a cell and five occupational groups are pre-selected, on average $100/5=20$ interviews will result. This is already a critically low number of interviews for any analysis differentiating between countries and occupational groups at that level. Alternatively, if in turn just two groups are pre-selected, then on average $100/2=50$ interviews will result.

In a survey design with five pre-selected groups per sector of activity and the (rather optimistic) assumption of an overall response rate of 10%, at least 1 000 addresses are needed to achieve on average $N = 20$ interviews for every group.

In a formula, the required sector size can be expressed in the following way:

$$S = \frac{N * O}{R}$$

where

S = minimal sector size

N = targeted minimal number of interviews per cell/occupational group

O = number of occupational groups pre-selected per sector of activity

R = expected overall response rate

The above example with a response rate of 10%, an expected minimum of $n=20$ interviews per cell (an occupational group within a sector in a particular country) and a survey concept using five pre-selected groups can be expressed in this formula as:

$$S = \frac{20 * 5}{0.1} = 1\ 000$$

If 50 interviews are envisaged as minimum per cell, then already 2 500 addresses are needed for the sector.

The required sector size can be calculated with the use of this formula, whereby N , O and R can be adapted to the situation in the countries where the survey is going to be implemented. In the definition of N (the targeted minimal number of interviews per cell) it should be considered that despite the application of a random selection mechanism among the pre-selected occupational groups, the number of net interviews resulting for each of these groups will differ. This has two major reasons:

- (a) some occupational groups can be found in almost all organisations within a particular sector while others may be concentrated in just a few of them;
- (b) the results of any computerised random selection mechanism will show a certain variation as long as the number of interviews in the cell is low. The larger the number of interviews, the more even the distribution will become.

For these reasons, it is recommended to define N not too narrowly, if it is imperative that in each of the five groups pre-selected within a sector a defined number of interviews are meant to be obtained. The precise definition of N is up to the researchers – there is no generally acknowledged minimum number of interviews a cell should have for analysis. The 20 interviews in the example

above are certainly an absolute minimum and can provide only indicative results. A target of 50 interviews per occupational group and country would be a much better and more widely-acknowledged base. Even this is still a very low number of interviews if results are to be analysed at this level of breakdown by country, sector and occupational group. With $N = 50$ interviews, the results in the cell have a confidence interval of roughly $\pm 10\%$ at the 95% confidence level. If the survey indicates that for example 85% of the establishments of a particular cell report an increase in task x for their occupational group y , the true value lies somewhere between 75% and 95%. For binomially distributed observations with percentage values of around 50%, the confidence interval is even higher and will roughly be at $\pm 14\%$, with the true value then lying between 36% and 64%. To show confidence intervals of not more than $\pm 10\%$ for any percentage value, a sample size of around 100 cases is needed.

4.3.2.3. *Possibilities and limitations of sector clustering*

The 12 NACE Rev. 2 divisions included in the pilot survey are grouped into seven sector clusters (NACE 28/29/30 Production, NACE 41/42/43 Construction, NACE 46/47 Trade, NACE 62 Computer programming and consulting, NACE 64 Financial services, NACE 84 Public administration and NACE 86 Health). The sampling, the steering of the net sample and the weighting were done based on these seven sector clusters, not on that of the 12 single NACE divisions which the clusters were composed of. Within each of the sector clusters, the same set of five pre-selected occupational groups is used as reference. The clustering thus alleviates the complexity of survey administration because only $7 \times 5 = 35$ occupational groups instead of $12 \times 5 = 60$ groups need to be selected and handled. The sampling and weighting matrices are also less complex, with $4 \times 7 = 28$ instead of $4 \times 12 = 48$ cells to be handled.

The clustering is a possibility for also reducing complexity in any further replication of the survey. However, clustering is not always an adequate means. Several aspects have to be considered:

- (a) only sectors (in this case: NACE divisions) in which the same occupational groups are prevalent should be pooled into a cluster. Which occupational groups are most prevalent within a NACE division can be counted on the basis of labour force survey (LFS) data showing the number of employees by occupational groups and sectors. If the occupational groups foreseen for a sector cluster do not fit for some of the sectors that the cluster is composed of, many interviews will have to be filtered to END because none of the pre-selected occupational groups exist;
- (b) should the tasks profiles of the occupational groups to be selected within a sector cluster largely differ between the single NACE divisions the cluster is

- meant to be composed of, clustering is not recommended. For an assessment of the degree to which the tasks differ, qualitative interviews with experts having a good insight into the respective sectors are recommended;
- (c) the unit non-response to be expected in the different sectors that a cluster is composed of should be at roughly the same level. If this is not the case, a separate steering of the samples by single sector is preferable to avoid the situation arising that in the net sample one of the sectors within the cluster dominates over the others in spite of not being larger in size ⁽²⁾.

The possibilities of clustering should be checked well in advance of fieldwork since the analysis of the LFS data for the identification of the relevant occupational groups and eventual expert interviews can be time-consuming. The clustering should be made using the NACE divisions. This level proved to be adequate for this aim. For a finer level of sector differentiation, the analysis of LFS data with regard to the prevalence of occupational groups would become problematic since the number of LFS interviews is likely to be too low for analyses at this very high level of disaggregation. A clustering on a higher level (summarising different NACE sections) is not an option since the 21 sectors are too different from one another to be clustered in a meaningful way for the purpose of this survey.

If only a selection of sectors is going to be included in the survey, interviews from other sectors will be filtered to END (in Q102) where the sector of activity is being verified. Similarly, interviews with establishments smaller than five employees will be filtered to END (in Q100c). These contacts would not count as full interviews.

4.4. Coverage of occupational groups

4.4.1. Pre-selection of occupational groups for the pilot survey

An important characteristic of the present approach is that many questions are not asked for the whole workforce but for a selected occupational group only. To

⁽²⁾ In the pilot survey in some countries problems of this kind emerged with the trade cluster composed of NACE 46 (Wholesale trade) and NACE 47 (Retail trade). Though in all countries the universe of establishments in NACE 47 is larger than in 46, the net sample of the trade cluster was in some countries dominated by units from NACE 46. The main reason for this was that in many shop chains there is a general policy not to take part in any survey. Since NACE 46 is much less dominated by large shop chains than NACE 47, this led to some distortions in the net sample of this sector cluster.

this end, in the pilot survey five occupational groups were pre-selected for each of the seven sector clusters in which the survey was conducted. The pre-selection was made based on information from the LFS. According to the LFS data, the five selected occupational groups were among the seven largest occupational groups within the respective sectors. Among these seven largest groups, a choice of five groups was made, in a mixture of quantitative criteria (number of employees in the group) and content-related criteria (which of the seven groups are most dynamic and/are considered as most interesting in terms of skill needs research).

4.4.2. Why (pre-)selection is necessary

As described above, the proposed survey concept is designed as an occupation-related approach. This means that the majority of the questionnaire (Module 2 on generic tasks as well as Module 3 on occupation-specific tasks) refers to a specific occupation within the interviewed local unit and not its total workforce. In this way, measurement errors are meant to be minimised because each respondent answers questions for a relatively homogenous group only.

Within large organisations, dozens of different occupational groups may coexist. The set of questions related to an occupational group takes about 10 to 15 minutes of interviewing time. It is obvious that this set of questions cannot be asked for each of the existing occupational groups, this would lead to an extremely long and repetitive interview. In the pilot, the questions were therefore asked about just one of the occupational groups present at the organisation. Repeating the entire set of questions for a second or third occupational group would lead to a too long and repetitious interview. There is also the possibility of confusion due to a switch between different reference groups (from questions about the whole establishment to questions about selected group 1 and then to questions about selected group 2 and finally back to the whole establishment) that may occur.

Since address registers listing the occupational groups present in an organisation do not exist, it is not possible to determine beforehand the occupational group about which questions are meant to be asked. This can only be done during the interview, after having clarified which of the occupational groups are present at the establishment. In the pilot questionnaire, this mapping was limited to five occupational groups. The limitation to five occupational groups is, however, not imperative. This number can be varied to a certain degree. Enhancing the number of groups leads to a lower number of observations for each occupational group and thus to a lower statistical power of the results. Reducing the number of occupational groups to be mapped and selected in turn leads to a higher number of net interviews for each of the selected groups, but

reduces the overall number of occupational groups that can be covered with the survey. A reduction also leads to more dropouts in the screening because establishments where none of the preselected groups is present should not be interviewed any further. The procedure adopted in the pilot survey in such cases is to ask these establishments for the largest occupational group that exists at the establishment and to continue the interview with reference to this group is not recommended for any replication of the survey. The occupational groups named in this open-ended question were very heterogeneous and many of them could not be clearly attributed to an ISCO group. For any quantitative analyses, these miscellaneous interviews were therefore of little value. In a replication of the survey, they would better not be conducted at all.

The heterogeneity of the answers to be expected on an open-ended question about occupational groups existing at the establishment and the difficulties to group these into the ISCO system is among the main reasons for preselecting a number of groups instead of asking the respondents to name the existing occupational groups in an open-ended question. Another aspect is that if asking the respondents to name a group in an open-ended question and then using this as reference group, answers would be concentrated around just a few occupational groups (those with the highest frequency), with only few interviews being conducted about the remaining groups.

4.4.3. How many occupational groups can be covered

The number of occupational groups to be covered in each sector cluster is largely determined by two factors: the net sample size envisaged per sector and the minimum number of interviews expected per occupational group. The latter is determined by the maximum level of standard errors and respective confidence intervals considered as tolerable.

4.4.4. How to pre-select the occupational groups to be covered

For the pre-selection of the occupational groups to be covered by the survey, both the research interests and practical criteria need to be considered.

4.4.4.1. Research interests

Depending on the aim of the study, some occupational groups may be of more interest for the survey than others. If the main aim of the survey is for example to give an input for a revision of VET curricula, then occupational groups for which VET is relevant should be the focus. Low-skilled occupational groups typically not requiring any VET, or groups made up of high-skilled academic workers, might be less relevant and could therefore be neglected in the pre-selection of the occupational groups to be targeted in the survey.

Should the main aim of the study be in turn to get more insight into the jobs that have been most dynamic in the recent past, then such particularly dynamic groups could be chosen with preference. Insights about the dynamism of a group can be gathered from different sources such as sector reports, expert interviews, information from other forecasting activities or the analysis of the quantitative development of an occupational group. At European level, the best source for this type of analysis is Eurostat's LFS which collects data about the number of employees by country, sector and occupational group. The large number of interviews collected annually for the LFS in each Member State guarantees that for even smaller sectors and occupations a reasonably high number of observations are usually available.

To analyse the development of sectors or occupational groups over time, LFS data from different years can be ordered and compared. In the comparison of LFS data from different years, two major changes in classification systems have, however, to be considered:

- (a) the switch from the NACE Rev.1.1 sector classification to the NACE Rev.2 system. In most countries, this switch was made in the LFS by the beginning of 2008, but in some countries (e.g. Germany) this only occurred a year later;
- (b) the switch from ISCO-88 to ISCO-08. This switch took place in all countries simultaneously at the beginning of 2011 so that all LFS data from 2011 onwards are coded by ISCO-08, while older LFS data are coded by ISCO-88.

Both revisions implied major changes for some of the categories. Comparisons of data from before and after these revisions are therefore problematic since it is not possible to clearly separate recodification effects from real quantitative developments. For data on occupational groups, comparisons between current data and data from before January 2011 should be treated cautiously, or not made at all. And for data on the sectors of activity, the data from 2009 onwards should not be directly compared with older data.

4.4.4.2. *Practical considerations*

As described above, it is recommended to concentrate the survey on a small set of pre-selected occupational groups and to terminate the interview, if none of these exists at the establishment. Therefore, from a practical point of view it is important to select groups from which it can be assumed that at least one of them is present in the majority of the establishments of the sector. This can best be ensured by including the quantitatively most important occupational group within

the sector as one of the pre-selected groups. In this way, it is likely that the large majority of interviews can be carried on.

For the choice of the remaining selected groups, the quantitative criterion should also be an issue of consideration. Interesting, but still quantitatively very rare occupational groups can be difficult to obtain in the survey, particularly if the few employees of that group are concentrated in just a few organisations within the sector. The inclusion of only such very small groups is likely to lead to many interviews being terminated prematurely because in many establishments none of the pre-selected groups exist. If choosing both small and very large groups, the majority of interviews will be conducted with these large groups – which are not necessarily particularly interesting groups.

4.5. Definition of the sampling unit and unit of enquiry

Surveys among organisations can generally be conducted at either the level of enterprises/companies or at the level of establishments/local units.

The terms establishment or local unit on the one hand, and enterprise or company, on the other hand are being used synonymously in this handbook.

For organisations consisting of just one production or service unit in the country (single-site organisations), the differentiation between establishment/local unit and enterprise/company is irrelevant. However, for all organisations that consist of more than one (legally dependent) unit in the country, the differentiation does matter: in an enterprise-based survey, the headquarters would be selected because this is the (only) address with which the enterprise is listed in the address register. The headquarters is then asked about the situation in the whole enterprise. In an establishment-based survey, in turn all units (the headquarters as well as all subsidiaries that are within the defined universe) have an equal chance to be selected. In the interview, the selected units are then asked about the situation in this local unit only (which may be a subsidiary or the headquarters). The sampling on the establishment level also implies that more than one unit of the same enterprise may be selected and interviewed.

For any replication of the employer survey of skill needs, it is recommended to once again use the establishment/local unit rather than the enterprise/company as both the sampling unit and the unit of enquiry. The main reason for this is that in case of multisite organisations, people responsible for personnel at the local unit are considered as much better informed about the

concrete work practices going on at the workplaces than the people at the central human resources department at the (possibly remote) national headquarters. Particularly where the tasks to be performed by employees of a specific group may vary widely between the different local units or between the local units and the headquarters, it can be assumed that the central human resources department is not the ideal respondent for the survey.

There had been considerable debate in the preparation phase of the survey on whether or not the targeted persons at the local units would have enough knowledge to answer the questions adequately. This issue was therefore closely observed in the pre-test. The pre-test did not provide any obvious indication that the establishment level would not be the right one, or that the enterprise level might in turn be the better unit for this type of survey.

From a practical point of view, using the establishment level provides some more difficulties than using the enterprise or company as the sampling unit. The choice of the establishment level implies the selection of an address source that systematically lists all establishments/local units of a multisite enterprise to ensure that each unit has an equal chance of being selected. Such registers exist only in about half of the current EU Member States. In countries where no suitable address register of establishments/local units exists, the best available register of companies/enterprises needs to be used instead and a specifically-developed screening procedure has to be applied to get a random choice of establishments/local units out of this company/enterprise-based register. This screening procedure is already contained in the attached master questionnaire (Q050 to Q099). It was used in the pilot for the interviews in the Czech Republic and Hungary and it has already been used in a very similar form in previous EU-wide surveys such as Eurofound's European company survey (ECS) 2009 or EU-OSHA's Esener 2009.

4.6. Definition and identification of the right respondent

The questionnaire is meant to be answered by the person with the best overview of the tasks to be performed by the workforce in the establishment/local unit. In larger establishments this will normally be the local human resources manager. In smaller establishments which do not have any human resources department or a human resources department with only administrative tasks (such as payroll administration), the managing director or branch manager will normally be the most appropriate respondent. Managers of specific departments or lines within the establishment are not adequate respondents for the entire questionnaire –

the interview needs to be started with somebody having an overview across all occupational groups working within the establishment.

All questions are designed so that they can normally be answered by the described target person. This holds even for those questions that are only meant to be answered for a specific occupational group (particularly Modules 2 and 3). In some cases the chosen respondent might nevertheless feel unable to have enough insight into the tasks and skills of the selected occupational group and might therefore want to refer the interview to a colleague, for example the line manager/department manager responsible for the chosen occupational group in day-to-day work. Generally, such switches can be allowed after Q109 but they should be applied only exceptionally. They should in any case be made only after the target person has been presented with the type of questions asked in Modules 2 and 3. If an interviewee is sure that they are unable to answer this type of question correctly for the selected occupational group, then a switch to somebody else in the establishment might be appropriate. When switching the telephone to another respondent, it is in any case crucial to ensure that the last part of the questionnaire (Module 4) is again answered by somebody with sufficient overview of the whole company to answer correctly these questions about changes and drivers of changes.

In cases where respondents at the local unit level are not allowed to participate in an interview but refer to the headquarters instead, the interview may be passed on to the central human resources at the headquarters. However, this should be granted only exceptionally because the headquarters will often not be able to refer answers directly to the situation in the previously contacted establishment. This is essential if the interview is being switched from the level of the chosen local unit to the headquarters – the interview must always refer to the initially contacted establishment.

CHAPTER 5.

The questionnaire

The questionnaire used in the pilot survey and printed in Annex 1 to this handbook should be considered as a suggestion and toolkit. Variations of this master questionnaire are possible when replicating the survey. If major changes are made to the questionnaire, then additional pre-testing before the launch of the full survey is strongly recommended. The pre-testing could concentrate on those issues which were modified and not tested in the pilot survey (or in any other previous survey).

5.1. The pilot master questionnaire and proposals for modifications

The master questionnaire shown in Annex 1 is the result of a long process of discussions and subsequent revisions. It has in this form been tested on a large scale in the pilot survey conducted at the beginning of 2012 in nine countries, with a total net sample of 8 523 interviews. In the course of the pilot fieldwork and in the posterior analysis of the data, some ideas for variations or improvements of the questionnaire became obvious. These are presented and discussed in this chapter.

5.1.1. Contact phase (Q001 to Q005)

Start of interview

The hints to the programmer have to be adapted to the survey design, in particular the codes of the countries and sectors of activity to be included need to be in line with the definition of the universe for the replication of the survey.

Q001

The hints to be used by interviewers in the introduction of the survey to potential respondents need to be adapted. In the replication of the survey, the aims and focus of the project, the client institution and other aspects might be different from the pilot.

5.1.2. Screening phase (Q050 to Q099)

The screening questions only need to be applied in countries for which no establishment-level register is available. The aim of the screening questions is to enable a random selection of establishments/local units in countries where no register of establishments/local units exists and where the sampling, therefore, has to be made on basis of an enterprise register. Additionally, the screening questions provide information necessary for an appropriate weighting of the data from the screening countries: Q051a and Q051c ask about the number of local units an enterprise has within the defined universe. This information allows for an additional weighting factor to be introduced in the screening countries. This factor enables a correction of the lower selection probabilities of the local units of multisite enterprises, compared to a sampling based on a register listing all local units (see Chapter 10 for more details on the weighting).

The screening part was considered quite complicated by the national institutes that had to apply it in the pilot (Czech Republic and Hungary). Any simplifications of the procedure are, however, likely to lead to a loss of representativeness regarding the selected local units and are, therefore, not recommended.

A drawback of the screening procedure in its present form is the selection of just one local unit per enterprise. Particularly if the survey is to be replicated in sectors of activity which are dominated by few large multisite enterprises with many subsidiaries, a modification of the screening procedure along the following lines might be worth considering. In the pilot, the selection of just one unit led to only very few interviews within the banking and insurance sector (NACE 62) in the two screening countries because in this sector there are only a few bank and insurance chains with dozens of subsidiaries spread over the country. The number of interviews could theoretically be improved by asking the headquarters for the contact details of more than just one of the local units which might lead to some more interviews. On the other hand, many headquarters might be unwilling to allow for more than one of their subsidiaries to be interviewed. If in a replication of the survey interviews are meant to be concentrated on sectors where few large chains dominate the universe, it might be worth testing whether or not the selection of more than one unit is a practicable option, for example, during pre-testing.

5.1.3. Module 1: background information on the establishment, selection and characterisation of an occupational group

Q100c

The number of employees in the establishment is an important piece of information, among others for the steering of the sample which should be done according to the size information from the questionnaire and not the size information from the address sources. In spite of the clear reference of this question to the local establishment, there are hints from the pilot results that the question was sometimes answered for the entire enterprise with all its local units. Therefore, an even clearer hint on the local unit should be inserted. This can be made by programming and reading out an additional clarifying text in all cases where $Q100=2$, where the selected unit is one of a number of establishments belonging to a larger enterprise or organisation.

Q102

This question aims at verifying the sector of activity. For sector clusters (composed of several single sectors), it is recommended to summarise the names of the different sectors the cluster is composed of for this question instead of asking about all the sectors separately. The most important aspect of this question is whether the attribution to the sector cluster is correct since the occupational groups to be shown in the survey depend on the sector cluster, not the single sectors a cluster is composed of.

Q104

This question on the market for the goods and services of the establishment was designed as a single-punch question in the pilot, with the focus on the main market. If all relevant markets are of interest (and not just the main market), then the question may be designed as a multipunch question (taking out the reference to the 'main' market in the formulation of the question).

Q109

This question requesting the largest occupational group was asked in the pilot if none of the pre-selected groups mentioned in Q107 existed at the establishment. The group named in this open-ended question was then taken as reference group for the further questions. For any replication of the survey, it is recommended to terminate the interview if none of the pre-selected groups exist. Following this recommendation, Q109 can be deleted.

Q111/Q112

Naming the rough percentage of the workforce that is under 30 years or 50 years or older provided some difficulties particularly for respondents in smaller establishments. They sometimes had problems in expressing the answer in terms of percentages and could usually better provide figures on the absolute number of persons in these age groups. Modifying the question by asking for the number of persons (instead of their percentage share) would help respondents' answers from small establishments. However, respondents from larger establishments can often better express the answer in terms of percentage categories so that this change should be implemented only if the majority of interviews is to be carried out in small establishments.

Q113

Current hard-to-fill vacancies as asked about in this question were reported by only a small share of establishments in the pilot survey. Moreover, the existence of vacancies seemed to be very much influenced by the economic crises as countries hit particularly severely by the crisis had only very few hard-to-fill vacancies. A broadening of the reference period by for example asking about hard-to-fill vacancies in the past two years instead of 'currently' would help to reduce the impact of economic cycle effects in this question. The main aim of the vacancy question in this survey is not a precise mapping of skill needs at the moment of the interviews but the analysis of possible correlations between vacancies, change and innovation in the establishment and changes in the task portfolio. Against this background, adding one or two (open-ended or closed) questions about the type of skills that was lacking in applicants for the vacant positions might be worthwhile, to get some more qualitative insight into the reasons for the vacancies.

Q114-Q116

This set of questions aiming at identifying the level of formal education required from applicants of the selected occupational group caused problems in the national adaptations of the questionnaire and partly also during fieldwork. When replicating the survey, this set of questions should be rethought.

Q116

In the current form, particularly Q116 caused frustration in some countries. This is reflected for example by an average rate of 3% for 'don't know/no answer' responses (unweighted figures). In Poland, as much as 10% of the respondents answered the question in this way (unweighted). These are clear signs for a need to reformulate this question.

Where there is questionnaire space available, for example in a more focused occupation or sector-specific study an additional question on career management might be considered, given the growth of mentoring activities, work-based learning and career development.

5.1.4. Module 2: generic task questions

The module with the generic task questions as tested in the pilot consists of three types of questions:

- (a) Q200, Q202, Q204, ...Q240 ask for the importance of a set of 17 generic tasks, that is tasks applicable to a broad set of different jobs;
- (b) Q201, Q203, Q205, ...Q241 ask for the change in importance of these 17 generic tasks;
- (c) Q242_1 to Q242_21 ask for each of the generic tasks reported to be on the increase whether the employees of the selected group are well prepared for these tasks (that are increasing in importance).

All three types of questions appeared to work well in the pilot. Feedback from the fieldwork partners did not reveal any general problems of understanding for any of these questions. The rates of 'don't know/no answer' were on average also low for these questions, in particular for the set of importance questions. The rate of 'don't know/no answer' for the change in importance and the preparedness questions was slightly higher, but in none of the questions did it surpass an average of 3.5% (unweighted). Also, each of the generic tasks mapped in the pilot turned out to have at least some relevance for a clear majority of respondents. Though a few of the generic tasks did not apply for a sizeable minority of up to about a third of the respondents (26% for manual dexterity, 28% for making speeches and presentations, 34% for reducing the use of raw materials; unweighted figures), the generic task dimensions mapped in the pilot master questionnaire can be considered as relevant enough for being kept in the questionnaire.

This indicates little need for revisions in this module as regards the selection and wording of the generic tasks. An issue of general consideration for any replication of the survey as a cross-national survey is however the use of the importance scale: judging the data based on a general knowledge of skills issues and the vocational training situation in the countries raised some doubts regarding the full international comparability of the questions using an importance scale: 'How important is task x (very important, fairly important/not important/not applicable). Is the importance of this task staying about the same, increasing or decreasing?' The pilot data showed some unexpected country results on this set

of questions which might derive from culturally different interpretations of what is considered important.

Using a less subjective measurement scale might reduce potential country biases. In the preparatory phase, the use of a frequency scale (how often is task xy performed by an occupational group in an average week/month?) instead of the importance scale had been discussed but was finally rejected. The reason for this rejection was that the frequency does not necessarily coincide with the relevance or importance of a task for a specific job. For example, a doctor may do dozens of blood tests per week, but maybe just one operation. Drawing conclusions from the frequency on the importance of a task and thus the respective skills required for it on part of the employee might be misleading as this example shows – the operation skills will not be considerably less important for this doctor than his skills in terms of blood tests.

The pilot results do however reconfirm that for comparability between countries the importance scale seems to be problematic and that a frequency or competence-level scale might be a better choice, at least for some tasks. The number of times something is used is a precise, countable category hardly influenced by any cultural response biases. An importance scale leaves more leeway for interpretation and cultural bias – what in country A might be considered as important might be rated as only fairly important in country B, just because of cultural habits in answering surveys (such as the avoidance of going for extreme poles of a scale such as ‘very’ or ‘not at all’ in some countries) or because of nuanced differences in the translation or translatability of the importance scale into some languages. At the very least some items with frequency scales could be used as anchors for the remaining items that involve an importance scale as could the addition of anchoring vignettes.

If deciding to replicate the survey with generic tasks only (without the occupation-specific questions), there is some room for a few additional generic task dimensions. Since this is a very repetitive part of the survey that is not exactly a pleasure to answer for most respondents, it is strongly recommended to limit the number of additional generic tasks to ask about to a maximum of up to five further tasks.

In principle, any further type of tasks can be added to the generic tasks as long as the task is easily understandable for respondents of different types of workplaces and can be supposed to be relevant for a large part of the targeted universe. It is possible to add new aspects of work that are meant to be of particular relevance for the future, for example similar to the environmental tasks that had been tested in the pilot. For any new tasks to be added, some prior pre-testing is recommended, for example in form of a few cognitive interviews. Care should be taken to minimise the overlap of any new tasks with some of the

existing task dimensions. Only dimensions really adding new, relevant information should be added.

5.1.5. Module 3: occupation-specific tasks and newly-emerging tasks

5.1.5.1. Module 3a: questions on occupation-specific tasks and newly-emerging tasks

The module with the occupation-specific task questions exists and has been tested for seven occupational groups only. For any replication of the survey intending to use the occupation-specific task modules, it will therefore be necessary to draft different versions of this module for all occupational groups for which this module has not already been developed and tested in the pilot survey project.

The occupation-specific questions developed for the pilot were closely oriented to ISCO-08. For each occupational group, between four and eight occupation-specific task dimensions were elaborated and asked about in the survey. Some of these questions turned out to be problematic. Some of the task dimensions were so general as to cause frustration with respondents (such as the task ‘constructing, maintaining and repairing buildings and other structures by using traditional or modern building techniques’ asked for the ‘building frame and related trade workers’). In other cases, the ISCO level at which the occupational groups were differentiated caused problems since some of the tasks differ considerably within the ISCO 3-digit groups: the tasks ‘providing care and support to women and newborns following childbirth’ or ‘assessing progress during pregnancy and childbirth’ asked for ‘nursing and midwifery associate professionals’ is for example relevant for a part of this group only, namely for the midwifery associate professionals which at the ISCO 4-digit level form an ISCO-08 unit group of their own.

For any replication and extension of the occupation-specific questions, it is recommended to use again the ISCO-08 job descriptions as an initial orientation, but to take the time and discuss these intensely with (international) experts familiar with the respective occupational groups. In these expert rounds, the portfolio of occupation-specific tasks to be asked for the group should be discussed. As a result of these discussions, the ISCO job descriptions may be further refined. In the development of questions for the occupation-specific part, care should be taken to keep obvious overlaps with the generic task questions to a minimum.

5.1.5.2. *Module 3b: newly-emerging tasks*

The questions on newly-emerging tasks proved to be rewarding in the pilot survey. It is recommended to maintain this part without major changes.

The newly-emerging task questions bring in the element of the future developments which are a core issue for an employer survey on skill needs. Though answers will vary considerably in terms of depth and quality, the open answers to this question will reveal developments relevant for the near future. The open answers can also be used to inform the set of task questions for a next wave of a survey.

5.1.6. Module 4: drivers of changing the tasks: innovations in the establishment

The questions of this last module serve for an analysis of the impact of different types of changes or innovations on the shape of tasks. Questions Q400 to Q403 proved to work well, though there were some complaints reported in the pilot about the (high) degree of abstractness in the formulation of some items. Since the questions have been thoroughly tested in the pilot and in other surveys, we recommend keeping this set of questions as it is. In the case of a sector-specific survey some benefit may arise from providing examples of what constitutes an innovation in each sector (cluster).

Questions Q404 and Q405 asking whether the selected group is the one most affected by changes within the organisation can be deleted when replicating the survey. This information mainly had the function of an indicator of the relevance of the survey.

It is possible, instead, to introduce some further questions referring to the whole organisation here. If a government is for example interested in some local phenomenon or policy, for example how some policy tool is associated with skill demand, an appropriate question could be added.

In general terms, any issues of particular national interest can in principle be added to the questionnaire, be it in Module 4 or in other parts. In the case of adding specific national parts, it will however be important that these amendments do not have repercussions on the parts designed for cross-national comparison. From this perspective, the best position to place national-specific elements is at the end in Module 4.

Any institution intending to replicate the survey should consider that task-based research is a quite new field of research and not already a fixed, well established science. There is currently a lot of ongoing research in this field, for example at the German Federal Institute for Vocational Education and Training (BiBB), from which new insights can be expected, for example for the interpretation of the task-related questions. Particularly if replicating the survey

only in a couple of years, a first step should therefore be to get an update on this kind of research. This might provide some new ideas for possible variations or refinements of the survey concept.

5.2. The mechanism for selecting an occupational group

5.2.1. The selection process applied in the pilot survey

In the pilot survey, for each of the seven relevant sectors of activity, three occupational groups are shown initially (Q107_1a/b, Q107_2a/b, Q107_3a/b). If only one of these groups exists at the establishment, this group has to be selected and referred to in various questions throughout the interview (as [group_select]). If more than one of the three groups exist, one of the two or three existing ones has to be selected in a random selection process. If none of these three occupational groups exist, the fourth pre-selected group (Q107_4a/b) has to be shown. Finally, if this one does not exist either, the fifth group is shown (Q107a/b). If none of these five groups exist, the respondent is asked to name the largest occupational group within the establishment (Q109) in an open-ended question. This group is then selected as the reference group for large parts of the interview. In a total, there are thus 35 different pre-formulated text elements that need to be programmed for the variable text element [group_select], plus the open naming of an occupational group in Q109 (applying if none of the pre-selected groups Q107_1a/b to Q107_5a/b exist).

Once the occupational group has been selected in Q107, this information is stored by the computer in the form of the variable text element [group_select]. This text element is repeated in several of the questions referring to this particular occupational group only. In this way, it is ensured that the answers to these questions really are referring only to employees of this specific group, and not to the whole workforce or to the personal situation of the respondent. To which occupational group reference was made in an interview is determined by the NACE code and the answers to questions Q107_1a/b to Q107_5a/b respectively Q109 in questionnaire Module 1.

The following questions of the pilot master questionnaire refer to the selected occupational group only:

Module 1: Q110 to Q116

Module 2: All questions in that module (i.e. Q200 to Q242)

Module 3: All questions in that module (i.e. Q300, Q301, Q301b, Q303 to Q308)

Module 4: Q403

The remaining questions that are not listed here refer to the establishment/local unit in total.

In the pilot master questionnaire, the reference group ([group_select]) is not repeated in each single question that refers to this group only so as not to make the questionnaire too long or monotonous. In those questions where the [group_select] is not directly named, it needs to be clear from the context whether the selected group or the whole workforce is meant. In some languages, this may require slight adaptations of the concerned national question versions. One solution for these languages could be to show the variable text element [group_select] in each of the questions, though this will make the interview more repetitive and a bit longer.

5.2.2. Proposal for modification of the selection procedure

Based on the experiences from the pilot, it is recommended that future implementations modify the selection procedure used in the pilot in two ways:

- (a) in the pilot, occupational groups 1, 2 and 3 had been preferred over groups 4 and 5 to get a substantial number of interviews for the first three groups. This worked in principle, but it turned out that the number of interviews for groups 4 and 5 was mostly too low to enable any sound statistical analysis. These interviews were therefore of limited value. In a replication of the survey, this can be avoided by applying the full random mechanism to all pre-selected groups instead of applying the two-stage procedure of the pilot. The number of pre-selected groups shown in the survey should not be too large in order to have enough interviews for each single group in the end. We recommend selecting three to five groups per sector, with fewer groups increasing the precision of estimates for the groups actually selected;
- (b) if none of the five pre-selected groups were present in an establishment, the pilot interviews were conducted with the group indicated by the respondent as the largest one in the establishment (Q109). Since the groups named in these cases were very heterogeneous, they could generally not be analysed with any quantitative methods. It is recommended not to offer this option in a replication of the survey, but to terminate the interview if none of the pre-selected groups are present at the establishment, recording a count of the number of such terminated interviews.

With these modifications, the questionnaire part where the occupational group is selected would be as follows:

Q107		Does exist	Does not exist	Group remains unclear	Don't know	No answer
		1	2	7	8	9
_1a	[group 1*]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_1b	Show only if Q107_1a = 7 [group 1 example]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_2a	[group 2*]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_2b	Show only if Q107_2a = 7 [group 2 example]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_3a	[group 3*]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_3b	Show only if Q107_3a = 7 [group 3 example]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_4a	[group 4*]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_4b	Show only if Q107_4a = 7 [group 4 example]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_5a	[group 5*]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_5b	Show only if Q107_5a = 7 [group 5 example]	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*** Programmer:**
please note that the names of the occupational groups to be inserted in Q107_1a/b to Q107_5a/b vary, depending on the NACE sector of activity of the interview. Insert the corresponding occupational group from column B in the Excel sheet 'ISCO groups', in accordance with the NACE sector the establishment belongs to according to the information from the address source (column A of the Excel sheet).
The variable text element [group_select] which is to be inserted in various later questions is defined in question Q107. The definition is as follows:

- if only one of the items Q107_1a/b, Q107_2a/b, Q107_3a/b, Q107_4a/b or Q107_5a/b is ticked with 'does exist' (= code 1), then this is the occupational group to be selected for [group_select];
- if more than one of the items Q107_1a/b to Q107_5a/b are ticked with 'does exist' (= code 1), then the occupational group to be used for [group_select] has to be randomly selected among these up to five groups that exist in the establishment. Please implement an adequate random selection mechanism;
- if none of the items Q107_1a/b to Q107_5 a/b are ticked with 'does exist' (Code 1), the interview is filtered to END and is terminated at this stage.

Once the occupational group has been defined in this way within each interview, this information has to be stored and will be used in several further questions as text for replacing the variable text element [group_select].

5.3. Hints for the interviewer briefing

Due to its complexity and innovative elements, a thorough interviewer briefing is a prerequisite for the successful implementation of this survey. The briefing should be done by the project managers in charge of the project at the national survey institutes in charge of fieldwork, together with the CATI studio supervisors that will be in charge of monitoring the interviews. In the briefing, interviewers should be given ample opportunities for testing the questionnaire in different

variations (with addresses from different sectors of activity and thus different occupational groups).

In the interviewer briefing, the general information given about the survey aims, the sampling unit, the target persons and the selection of occupational groups have to be explained to the interviewers. Particular emphasis needs to be placed on the mapping of occupational groups, the process applied or the selection of a particular group, and on the reference of many questions to this previously selected occupational group only. A correct reference is crucial for the quality of the data.

Apart from these general issues, the following hints on specific questions should be part of the interviewer briefing:

5.3.1. Contact phase

For the contact phase, the following arguments should be provided to interviewers (see also the hints in the introductory part of the interview):

- (a) name the client institution, its legal status and its aims (the survey is conducted on behalf of *insert institution name* is *insert key details*. The web portal is *insert Web page address*).
- (b) the motivation letter is meant to help interviewers to convert initial refusals (especially soft refusals) into interviews. The letter should be offered to all respondents who do not spontaneously agree to be surveyed and in all cases where the contact person first has to consult superiors or colleagues about permission to participate. Experience has shown that an official recommendation letter can help a lot in converting soft refusals at this stage into interviews, particularly if the motivation letter is from an acknowledged official institution. This letter might also include endorsement by a relevant sector organisation that also sends newsletters to their members and advertise important events/initiatives on their websites.

What is the benefit of participating in the survey?

- (a) The survey aims to identify future skill needs of employers in Europe at an early stage. By bringing in their own views and experiences in the survey, participating establishments can exert influence on the education and training schemes. This will help to improve future skills of the labour force;
- (b) well-trained job applicants having the competences and qualifications required at the workplaces are an important competitive advantage;
- (c) if feasible: participating establishments will receive an exclusive short report of the main results for free download before the results will be published officially. To this end, respondents should send an e-mail request to *insert e-mail address of research organisation*.

5.3.2. Q100c/Q110

In the case of multisite enterprises, it is important to get the number of employees working in the contacted establishment/local unit, not the number of employees working in the whole enterprise (with all its subsidiaries).

The reference to employees on the payroll is meant to exclude external contract staff such as freelancers, temporary agency workers, etc., who are working at the establishment, but do not receive their salaries directly from the firm or organisation the establishment is part of. Whether an employee receives the salary directly from the local establishment or whether it is being remitted from the headquarters of the firm is not relevant here. Thus, employees who work in a subsidiary, but receive their money directly from the headquarters of the firm or organisation are nevertheless to be counted as employees on the payroll of the local establishment.

5.3.3. Q102: verification of the sector classification

In this question, the information on the sector of activity that comes from the address source is being verified. The correct attribution of the sector of activity is of particular importance for this survey because for each relevant sector, another set of occupational groups is shown in the interview (Q107) and because the survey covers only a selection of all existing sectors. Interviews with establishments from sectors that are not among the sectors targeted for the pilot would be useless. Therefore, establishments with a wrong sector codification in the address source are filtered out in Q102. If you experience that the respondents are unsure of whether or not an establishment with a particular activity belongs to the targeted sector group, the following list can help to come to the correct decision (this is the list of sectors that were included in the pilot; this list has to be adapted to the universe defined for the replication of the survey):

- (a) NACE 28, 29, 30: manufacture of machinery, equipment, motor vehicles or other transport equipment:
- manufacture of general-purpose machinery;
 - manufacture of office machinery and equipment (except computers);
 - manufacture of hand tools;
 - manufacture of agricultural and forestry machinery;
 - manufacture of machinery for food, beverage or textile production;
 - manufacture of motor vehicles, trailers and semi-trailers;
 - manufacture of parts and accessories for motor vehicles;
 - building of ships and boats;
 - manufacture of railway locomotives and rolling stock;
 - manufacture of air and spacecraft machinery;
 - manufacture of military vehicles;

- manufacture of bicycles, motor cycles or other transport equipment;
- (b) NACE 41, 42, 43: construction of buildings, civil engineering or other specialised construction activities:
 - development of building projects;
 - construction of residential and non-residential buildings;
 - constructions of roads, railways, pipelines or irrigation systems;
 - construction of water projects, power plants, etc.;
 - demolition of buildings;
 - earth moving and preparation of building sites;
 - drilling;
 - installing electrical installations, heating, plumbing or other construction installation;
 - finishing of buildings, for example plastering, floor and wall covering, cleaning of newly constructed buildings;
 - roofing;
- (c) NACE 46 and 47: wholesale or retail trade:
 - wholesale agents for all types of all types of goods;
 - retail sale of all types of goods (except motor vehicles) in stores, stalls, vial internet, etc.;
- (d) NACE 62: computer programming, consultancy and related activities:
 - development of systems software, databases, webpages;
 - planning and designing of computer systems;
 - installation of personal computers or of software;
- (e) NACE 64: financial service activities:
 - monetary intermediation by banks, saving banks and credit unions;
 - activities of holding companies;
 - management of investment and other funds;
 - credit granting;
- (f) NACE 84: public administration and defence; compulsory social security:
 - executive and legislative administration at the local, regional or central level;
 - administration of waste collection and disposal operations;
 - foreign affairs;
 - defence activities (military);
 - justice and judicial activities;
 - public order and safety activities, for example police, fire fighting and fire prevention;
 - compulsory social security activities, for example administration of unemployment insurance, retirement pensions, etc.;
- (g) NACE 86: human health and social work activities:

- hospital activities;
- general medical consultation and practice;
- special medical practice;
- dentists.

5.3.4. Q106b

The examples named for other types of training might not be familiar to all respondents. Please explain if necessary:

- (a) on-the-job training: on-the-job training is characterised by planned periods of training, instruction or practical experience in the workplace using the normal tools of work, either at the immediate place of work or in the work situation;
- (b) secondments: a planned temporary transfer to another enterprise/organisation with the primary intention of developing skills of the workers involved;
- (c) learning circles: learning circles are groups of persons employed who come together on a regular basis with the primary aim of learning more about the requirements of the work organisation, work procedures and workplaces;
- (d) quality circles: quality circles are working groups having the objective of solving production and workplace based problems through discussion. The primary aim of the persons attending should be learning.

5.3.5. Q107_1a/b to Q107_3a/b

The main aim of this question is to identify a relevant occupational group as reference group for large parts of the further interview. Since many further questions will refer to the occupational group selected at that stage, it is very important that in this question respondents get a clear idea of the groups the survey is about.

In a first step, in Q107_1a, 2a, 3a, 4a and 5a general description of the relevant groups is given. If from this description a respondent gets a sufficiently clear idea about the employees being referred to, then the respondent's answer should be entered (group does exist or does not exist). If the group remains unclear – for example, if respondents have spontaneous questions about whom of their employees is part of that group and whom not – then code 7 'group remains unclear' should be entered in Q107_1a, _2a, _3a, _4a and/or _5a. Then, some more concrete examples for the occupational group are shown in the following subquestions Q107_1b, _2b, _3b, _4b and 5b. Having now been shown the examples, the respondent's answer can now be entered. Only if the respondent still does not have a clear idea about an occupational group, code 7 'group remains unclear' shall be entered. This group will then not be considered in the selection of an occupational group for the remainder of the interview.

Based on the answers in Q107, the computer system selects one particular occupational group as reference group for the further interview. If none of the groups mentioned in Q107 exist in the establishment, the interview is terminated immediately after Q107.

Interviewers should familiarise themselves with the occupational groups and the examples given in Q107. In the pilot, the following groups and examples were used (the group with an x in the column 'group 1' is the group for which the occupation-specific questions Q300/Q301 were additionally asked).

5.3.6. Q110

All questions from Q110 to Q308 refer to the selected occupational group only. In Q110 to Q112, first some general characteristics about the employees of this specific occupational group working within the establishment/local unit are asked about. These questions just concern this particular selected group, not the whole workforce.

5.3.7. Q113

With vacancies proving hard to fill we mean vacancies that cannot be filled since a couple of weeks or months because of lack of applicants having the required skills. The time from when on a vacancy is defined as a 'hard to fill vacancy' has deliberately not been defined precisely.

5.3.8. Q114 to Q116

These questions are asked of all establishments, not only to those currently having any vacancies that are hard to fill. It is important for the analysis of the survey to know which requirements, in terms of formal education, the establishment normally has for applicants of that occupational group – even if the last recruitment in this group was already some time ago.

5.3.9. Q200 et seq.

The task questions in Modules 2 and 3 are to be referred to employees of the selected group only. It might occur that the working tasks of the employees from the same occupational group vary to some extent within the establishment. In this case, respondents are meant to give their 'best estimate of their importance on average' (see introductory text to Module 2). If a task is for example very important for two out of five employees of an occupational group, but not particularly important for the other three employees of that group, the correct answer would be 'fairly important' as indication for the 'average' importance of the task for that group.

Table 4. Occupational groups addressed in the pilot survey

Code	(Occupational) groups 1 to 5	Examples for groups 1 to 5	Group 1
291	Assemblers	Assemblers of mechanical machinery, electrical or electronic equipment	X
292	Machinery mechanics and repairers	Mechanics and repairers of motor vehicles or industrial machinery	
293	Engineering professionals (excluding electrotechnology)	Industrial or production engineers, civil engineers, environmental or mechanic engineers	
294	Physical and engineering science technicians	Mechanical or electrical or civil engineering technicians	
295	Sheet and structural metal workers, moulders and welders, and related workers	Welders, flamecutters, structural-metal preparers and erectors	
411	Building frame and related trade workers	Bricklayers, stonemasons, concrete placers or carpenters	X
412	Physical and engineering science technicians	Mechanical or electrical or civil engineering technicians	
413	Building finishers and related trades workers	Roofers, floor layers, glaziers, plumbers or air condition mechanics	
414	Engineering professionals (excluding electrotechnology)	Industrial or production engineers, civil engineers, environmental or mechanic engineers	
415	Mining and construction labourers	Civil engineering or building construction or mining labourers	
471	Sales, marketing and public relations professionals	Advertising, marketing or technical sales professionals	X
472	Shop salespersons	Shop keepers, shop supervisors or shop sales assistants	
473	Material-recording and transport clerks	Stock, production or transport clerks	
474	Cashiers and ticket clerks	Cashiers and ticket clerks	
475	Health professionals	Pharmacists or optometrists	
621	Software and applications developers and analysts	Systems analysts, software, web or multimedia developers, applications programmers	X
622	Information and communications technology operations and user support technicians	ICT operations and user support technicians, web or computer network and systems technicians	
623	Database and network professionals	Database designers, or administrators, systems administrators or computer network professionals	
624	Client information workers	Enquiry clerks, receptionists or contact centre information clerks	
625	Business services and administration managers	Human resource managers, finance managers or business services and administration managers	
641	Finance professionals	Accountants, investment advisers or financial analysts	X
642	Business services and administration managers	Human resources managers, finance managers or business services and administration managers	
643	Financial and mathematical associate professionals	Credit and loans officers, finance dealers and brokers or statistical associate professionals	

Code	(Occupational) groups 1 to 5	Examples for groups 1 to 5	Group 1
644	Tellers, money collectors and related clerks	Bank tellers, bookmakers or money lenders	
645	Numerical clerks	Accounting, bookkeeping, statistical, finance or insurance clerks	
841	General office clerks	General office clerks	X
842	Protective services workers	Fire fighters, police officers, prison or security guards	
843	Regulatory government associate professionals	Police or customs inspectors, government tax or social benefits officials	
844	Administrative and executive secretaries	Office supervisors or legal or administrative secretaries	
845	Domestic, hotel and office cleaners and helpers	Domestic, hotel and office cleaners and helpers	
861	Nursing and midwifery associate professionals	Nurses, midwives	X
862	Personal care workers in health services	Health care assistants, home-based personal care workers or other personal care workers	
863	Medical doctors	Generalist or specialist medical doctors	
864	Medical and pharmaceutical technicians	Medical laboratory technicians, medical and dental prosthetic technicians or pharmaceutical technicians and assistants	
865	Domestic, hotel and office cleaners and helpers	Domestic, hotel and office cleaners and helpers	

5.3.10. Q200/Q201

Here, respondents should be informed that we are not talking about reading abilities in general, but about employee's ability to read and comprehend instructions, guidelines, manuals or reports. Therefore the full text should be read out.

5.3.11. Q202/Q203

Respondents should be informed that we are not concerned about writing abilities in general, but about employee's ability to write instructions, guidelines, manuals or reports. Therefore, full text should be read out.

5.3.12. Q200, Q202, Q204, ...Q234

These questions on the importance of a task always refer to the selected specific occupational group only, not to the whole workforce (or to the person being interviewed). What is meant here is whether the respective task is important in the job of these employees, i.e. whether it is an important part of their work in this establishment.

5.3.13. Q201, Q203, Q205, ...Q235

These questions on changes in importance also always refer to the selected specific occupational group only, not to the entire workforce or the interviewee. What is meant here is whether the importance of the task for the job of these employees is changing. If respondents ask to be given a time horizon as anchor point for these questions, interviewers should instruct them that the questions are about changes that have taken place in the recent past and can still be observed at the moment of the interview.

5.3.14. Q240

Explanations in brackets should always be read out to respondents. Without these explanations, the scale is insufficiently clear because, at least in the English version, the differentiation between 'straightforward' and 'moderate' and between 'complex' and 'advanced' is not self-evident.

5.3.15. Q242

Some respondents might ask what is meant by 'well prepared'. Here, respondents should just refer to the current state of preparedness of their employees, no matter whether they are well prepared (or not sufficiently well prepared) by the education system, by training provided by the employer or from any other 'source' of training or knowledge.

5.3.16. Q300/Q301

In the pilot survey, this set of questions was asked in part of the interviews only – namely in those where the selection mechanism for the occupational group in Q107_1a/b to Q107_5a/b led to the selection of 'group 1' for interview. Questions Q300/Q301 were different for each sector cluster.

In a replication of the survey, these occupation-specific task questions will probably exist either for all occupational groups covered in the survey or for none of them. This part of the questionnaire and the interviewer hints will therefore have to be adapted.

5.3.17. Q303

This question is in general asked to all respondents. But respondents who had to answer questions Q300/Q301 before get a different entry text to the question than those coming directly from Module 2.

In a future implementation of the survey, these occupation-specific task questions will probably exist either for all occupational groups covered in the survey or for none of them. This section of the questionnaire and the hints for interviewers will need to be adapted.

5.3.18. Q304

This question about newly-emerging tasks for the selected occupational group is crucial for the success of the survey which ultimately aims at identifying new demands in terms of competences and qualifications employees should have. Please make sure to obtain and enter clear and concise answers here.

If a respondent names only one newly-emerging task, this task is to be entered as open answer in Q304_1 and in Q304_2 then 'no further task' should be ticked. If a respondent has two relevant tasks to name, then the first task is to be entered as open answer in Q304_1 and the second task as open answer in Q304_2. In Q304_3, then 'no further task' would have to be ticked and so on.

5.3.19. Q305

The tasks named in the previous question have to be shown.

5.3.20. Q306

This question refers to all newly-emerging tasks named in Q304/305.

5.3.21. Don't know

One should try to minimise 'don't know' answers by prompting interviewees to make their best estimate, if they are not totally sure about their answer. As long as they are reflecting the, albeit vague, impression of the respondent, best estimates are still preferable to an item non-response.

5.3.22. Don't know/no answer

The differentiation between 'don't know' and 'no answer' is important for this survey which still has a pilot character. If somebody would be willing to give an answer, but does not feel well enough informed to do so, 'don't know' is the right category to be coded. If the respondent does not want to answer a particular question, for example considered as too sensitive, the code 'no answer' should be offered.

'Don't know' or 'no answer' should only be recorded as alternative responses; they should not be included in the response scale.

CHAPTER 6.

Translation

6.1. General translation guidelines

If the survey is going to be replicated as a cross-national survey, the elaboration of adequate translations is indispensable. For the general translation procedure, different options are possible. Some international survey researchers rely on additional back-translations as a means to ensure full comparability of the national versions. Others prefer procedures in which two people translate the questionnaire separately and then compare and discuss the two variants, possibly with the participation of some kind of arbitrator.

Which of the two methods is to be preferred for this survey depends on the personal preference of the responsible researchers and on the profile of the people available for the elaboration and check of the various national versions. Both procedures are shortly described below:

(a) variant A: translation and back-translation:

- (i) in a first step, a professional translator who is a native speaker of the target language translates the questionnaire into the respective language. Ideally, this version is additionally proofread by another translator before being delivered;
- (ii) the delivered national language version is then checked by an independent person who is native speaker of the target language and ideally at the same time an expert in the subject matter (here: skill needs research and/or survey research in general). This type of check can, for example, be done by a network of international experts maintained by the client institution commissioning the survey or by employees of the institutes in charge of the national fieldwork. The checks should result in questionnaire versions containing comments and concrete proposals for revisions;
- (iii) the resulting annotated questionnaires should then centrally be checked by the coordination unit. The main aim of this check is to ensure the consistency of the proposed national revisions with the original master questionnaire. Revisions that are fully in line with the master can be accepted while others have to be discussed and subsequently modified;
- (iv) the revised national-language versions can then be passed on to another translation agency for back-translation. The back-translation

- should be done by a translator who is not familiar with the English master version;
- (v) the back-translated versions are then compared to the original English master version. This comparison should best be done centrally by the coordination team. All persons in charge of these checks need to be well familiar with the English language. Any differences with repercussions on the content need to be annotated and discussed with native speakers, for example with those that had been in charge of checking the first version of the translation (step ii). If differences between the two versions are due to mistakes in the national version, the national version has to be revised accordingly;
- (b) variant B: elaboration of two independent translations and comparison of the translations:
- (i) the first step consists of the simultaneous translation of the English master questionnaire by two translators working independently from each other;
 - (ii) these two translated versions are then compared to each other. This comparison can either be made by the two original translators who then discuss any differences with each other. Or they can be made by a third person, an independent arbitrator. The arbitrator then annotates any differences in substance between the two translations and decides question by question which of the translations is the more accurate and fluid one. If none of them fully meets the expectations, the arbitrator may also implement an own version of the translation. In case of doubts about the best way of translating an issue, the arbitrator can discuss the two versions with the original translator. With the elaboration of the merged version by the arbitrator, the translation process can be terminated;
 - (iii) if time and financial means allow for a further step, a second independent arbitrator can be contracted to check the decisions of the first arbitrator and discuss any issues of dissent with the first arbitrator.

Both translation procedures have their advantages and disadvantages. An advantage of the back-translation is that it enables the project coordinators to have the chance for a final check of all changes introduced in the checking process without necessarily being familiar with the respective national language. This possibility is lacking in the second option – there, the coordinators have to fully rely on the work of the native speakers in charge of the checks. The chance of a central final check could be particularly helpful in cases where the optimisation of the national language versions might have led to some

discrepancies with the original master version, thus potentially endangering the full comparability of all national versions. It may also usefully feed into further revision of the master version to remove ambiguities that occur when it is translated into some languages.

Variant B, in turn, tends to be less costly and more time efficient, at least as long as it is terminated with step (ii). If a further arbitrator is contracted, this advantage gets lost, but the quality of the translations is further improved. In general, for variant B it is important to avail for each of the national versions of a person who is native speaker of the target language and has some expertise in the subject matter of skills research and/or survey research more in general.

6.2. Translation of official terminologies

In two parts of the master questionnaire, international codification systems are being used:

- (a) in Q051b/Q102, the NACE Rev. 2 classification of sectors of activity is used;
- (b) in Q107, the ISCO-08 codification of occupational groups is used.

The translation of these terms, in particular of the occupational groups according to the ISCO system, is crucial for achieving full comparability of the data. If there are differences in the denomination of the occupational groups to which reference is made during the interview, this may result in a comparison of apples with oranges and may thus render the survey worthless in terms of international comparability.

Therefore, the official terminologies introduced in these two sections of the questionnaire should not be translated in the same way as the rest of the questionnaire because for both classification systems, acknowledged official translations into all EU languages do already exist. These official national language versions should be used. They are accessible via internet under the following web links:

- (a) for the ISCO-08: <http://eur-lex.europa.eu/JOHtml.do?uri=OJ%3AL%3A2009%3A292%3ASOM%3AEN%3AHTML> (European Commission, 2009);
- (b) for the NACE Rev.2 classification (up to the level of NACE Rev. 2 divisions): <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32010R0715:EN:NOT> (European Commission, 2010).

6.3. Specific challenges in translating the questionnaire

The questionnaire for the skill needs survey contains a few specific difficulties for the translations which are described in the following paragraphs.

6.3.1. Variable text elements

The occupational group selected in Q107 ([group_select]) is repeatedly shown in subsequent questions to make clear to the respondent that questions are still referring to this particular group. The respective questions are formulated in such a way that the insertion of the name of the occupational group works well in the English language and in most other languages. The pilot has, however, shown that for some languages the grammatical form of the terms for the occupational groups might need to be adapted to result in a grammatically correct interrogative sentence. This needs to be checked for each language and for each single question where these variable text elements are to be placed. If adaptations should prove to be necessary, the grammatically correct versions of the terms will have to be provided for each of the occupational groups and for all questions where the variable text [group_select] is to be inserted.

6.3.2. Complex programming instructions

Some parts of the questionnaire contain quite complex hints for the programmers. This is particularly the case in the following parts:

- (a) screening part Q050 to Q099 to be used in some countries;
- (b) the selection mechanism for the occupational group in Q107 to Q109;
- (c) instructions for handling the variable text elements for the occupational groups.

Experiences from the translations of the pre-test versions have shown that professional translators may have difficulties with this type of subject-specific terminology leading to mistakes in the translations. Translation mistakes in these parts may in turn easily lead to major mistakes in the data. It may therefore be safer to leave these complex programming instructions in the English language, particularly since most programmers are familiar with written English, as it is the most widely used language for programming. If using a centralised programming for all country versions, the translation of the programming instructions is not necessary anyway.

The translation of some crucial terms that are repeatedly used in the questionnaire is not easy. While the term 'task' still has direct equivalents in most languages, the term 'skill' for example is in many languages somewhat arbitrary

for the translation. For any important terms that might cause difficulties for the translation, the set-up of a translation glossary will be helpful.

6.4. Translations in case of repeating the survey

For the observations of trend developments over time, a replication of the survey in certain time intervals would be very useful. In any replication using the existing pilot master questionnaire or substantial parts of it, the handling of the existing translations is an important success factor. For all questions taken from the pilot master questionnaire, the existing translations should be respected and used as starting point. Though the translations elaborated for the pilot survey were checked in several stages and by different persons, they should be checked once again for any replication of the survey.

To this end, translators should receive the existing translations together with all new survey parts to be translated from the English language. The translators should then check the existing translations for correctness. Questions that are wrongly translated in the opinion of the translator in charge of this check should be annotated and proposals for improved formulations should be made. An independent second person who is native speaker of the target language and expert in skill research and/or in survey research in general should then compare the proposals for reformulation to the original translation and should then take a decision on the best version. The previous translation should be replaced only if it was clearly faulty or misleading and not for mere issues of language style. The alternative proposals and the reasons for the decision should be documented so that the decision process can easily be reconstructed by any further persons.

All national translations elaborated for the pilot survey are stored at Cedefop in MS-Word format and can be made available for any replication of the survey. In addition to the English master version, the following language versions exist: Spanish, Czech, German, French, Italian, Hungarian, Polish and Finnish.

CHAPTER 7.

Programming and testing the questionnaires

7.1. Specific challenges for programming

The questionnaire elaborated for the pilot employer survey on skills is a complex instrument. It can nevertheless be programmed on most of the common CATI (computer-assisted telephone interviewing) software platforms, for example in Bellview, Quancept or NIPO Odin in the case of implementing a CATI survey. A prerequisite for the software to be used is that it offers the possibility to fade in variable text elements in the course of the survey which are defined by the entries to questions the respondent answered earlier in the survey (question bloc Q107).

If the survey is to be replicated with occupation-specific questions for all or a broad number of sectors, software which supports the embedding of some kind of data bank is strongly recommended to be able to administer the multitude of variants (different occupational groups) in a safe and efficient way.

In general terms, the most difficult parts of the scripting for this survey are:

- (a) the check of the correct sector attribution;
- (b) the mapping of the occupational groups in Q107, including the presentation of examples in dependence of the respondents answer;
- (c) the random selection of one occupational group ([group_select]) among all of the pre-selected groups that exist in the establishment;
- (d) the explicit reference to this particular group in a number of subsequent questions by fading in the name of the group;
- (e) the administration of the question bloc with the occupation-specific questions (if applicable).

These parts of the questionnaire need specific attention not only in the elaboration of the CATI questionnaire script, but also in its testing. It has to be ensured that the following points are met:

- (a) for each interview the sector attribution according to the information in the address source(s) is shown so that it can be checked with the respondent whether or not the sector attribution is correct;
- (b) the correct occupational groups and the examples for the groups are shown for each interview. The occupational groups to be shown differ in accordance with the sector an address is attributed to;

- (c) for each interview, one occupational group is selected as reference group, either in one of the questions of question bloc Q107. This information has to be stored so that it can be reused in later parts of the questionnaire;
- (d) the random selection procedure works properly, i.e. that each of the pre-selected occupational groups that actually exists at the establishment has the same chance of being selected for the interview (wherever more than one of the groups exists);
- (e) the term for the selected occupational group is shown in all questions where it is foreseen to show this element. In this context, it also has to be checked whether the element is grammatically correct in all languages and in all relevant questions;
- (f) the terms for the occupational groups fit into the grammatical structure of the sentence in which they are embedded.

7.2. Organisation of programming

Due to the high complexity of the survey, for multinational studies with a larger number of participating countries it is recommended to programme centrally and test a master script that can be shared by all countries. This improves efficiency and is likely to improve the script quality by avoiding individual national programming mistakes. This recommendation pre-supposes that all participating countries use the same CATI platform.

If the script is not centrally programmed, it is important to provide a central data map, the use of which should be made obligatory for each national institution in charge of fieldwork. In this data map, the names and positions of all variables should be defined. Such a central data map helps to harmonise the national data sets and facilitates the interim and final data checks.

7.2.1. Specific hints on programming the CATI instrument

After the pilot survey, suggestions were made from part of some fieldwork partners to show the selected occupational group not only in those questions where the variable text element [group_select] is integrated into the question, but to display it on the screen in all those questions where reference is to be made to the chosen specific occupational group to remind interviewers about whom we are talking (such as ‘this question is to be answered for [group_select] only’).

For the questionnaire parts related to the sample management system (interview entry), national adaptations to the respective local CATI system should be allowed, within certain limits. For example, local institutes need to be free to include their national texts about data protection and the handling of the data in

the entry phase. The entry text itself and the texts for the identification of the proper respondents should, however, not be changed, to avoid any national influences in the recruitment phase.

For the programming of nation-specific text parts in the introduction to the interview, the central data map should foresee some columns for free use so that all countries have enough space to store their local sample management information in the central data file.

CHAPTER 8.

Fieldwork: timing, organisation and monitoring

8.1. Forms of survey organisation

In principle, the survey can be coordinated and carried out either by the national statistical offices or by any survey research institution with reliable and experienced cooperation partners in all countries to be covered. Preconditions on the part of the institute(s) responsible for the replication of the survey are the availability of a suitable software platform for the programming of the questionnaires and experiences in the programming and administration of complex survey instruments, a computerised sample management system and a sufficiently large staff of interviewers with experience in interviews among organisations.

If the survey is going to be carried out as a multicountry study, a central coordination unit should be set up to ensure that the survey is being implemented according to the same parameters in all countries. The role of the coordination unit should at least comprise the following tasks:

- (a) coordination of the time planning for fieldwork preparations and fieldwork itself in all countries;
- (b) finalisation of the master questionnaire (if any changes to the pilot master questionnaire are envisaged) and the motivation letter;
- (c) coordination of the translation process and in particular of the translation checks to guarantee that the national versions are fully compatible with one another;
- (d) coordination of the sampling, in particular the set-up of the sampling frames (including a check of the adequateness of the address registers, definition of the categories of the sampling matrix, setting the net sample targets for all cells);
- (e) coordination of the programming in the participating countries (finalisation of a master script with all hints to the programmers, clarification of difficulties in the programming, provision of a data map to be used in all countries) or the setting-up of a central questionnaire script and elaboration of the national versions;
- (f) centralised check of interim data sets for correct filtering, correct data structure, etc.;

- (g) central surveillance of fieldwork progress and collection of information on the fieldwork (reporting and clarification of difficulties, reporting of unit non-response, etc.);
- (h) provision of central guidelines for interviewer briefing.

The national fieldwork partners need to be obliged to follow strictly the rules set by the coordinating agency. The coordinating unit must clearly show on which aspects of fieldwork the national institutes have some leeway and where not.

Due to the complexity and innovativeness of the survey, the coordination unit should have ample prior experience in the administration and coordination of employer surveys in a variety of settings and with non-standard survey designs. It is unlikely that the survey will be replicated in exactly the same way as the pilot. Therefore, in the preparation phase, various decisions on details of the survey design will have to be made and compromises will have to be taken. Not all the possible variations and adaptations of this survey concept could be dealt with in this pilot.

8.2. Time planning

Due to its rather complex design, the survey requires a relatively long preparation phase. Once the survey concept is defined and eventual modifications of the model questionnaire are made, roughly the following time spans should be made available as minimum for the respective working steps:

Table 5. Time frame needed for survey preparation

Working step	Time frame
Sampling preparations (collection of information on the universe, selection of an address source, ordering of addresses, drawing of the gross sample, etc.)	Three to four weeks
Translation process (including checks of national versions)	Six to eight weeks
Programming and testing of the questionnaire script	Two to three weeks
Total	11 to 15 weeks

Source: Cedefop pilot survey 2012.

The subsequent fieldwork period for the study should be dimensioned generously. A sufficiently long fieldwork period is a prerequisite for the achievement of high response rates. In surveys among human resources managers or other leading management representatives, much of the interviewing cannot be done spontaneously, but requires prior appointments. Frequently, these appointments have to be postponed due to other tasks or events on part of the respondents. There should be enough time to allow for such

unforeseen postponements. Also, some cells of the sampling matrix may be more difficult to fill than others. This, too, should be considered in the time planning.

If the survey is going to be carried out with around 1 000 interviews per country, a fieldwork period of five to eight weeks should be foreseen as a minimum. With larger sample sizes, the fieldwork period should accordingly be longer. For a sample size of 2 000 interviews, for example, about eight to 12 weeks should be scheduled as minimum. However, these are just approximate values for orientation. The fieldwork period actually required depends on a number of further factors such as the interviewing capacities of the fieldwork partners (number of CATI stations, number of available interviewers with experiences in interviewing companies or establishments) and the size of the available gross sample (if the available gross sample is small in relation to the expected net sample, all efforts have to be made to get as many interviews out of the gross sample as possible). Should the survey be concentrated on the larger countries and on a broad array of sectors of activity, the fieldwork period will tend to be shorter than if small countries are included and/or the survey is concentrated on very few sectors of activity only.

Before the launch of the project, a detailed time plan should be drafted and agreed with all participating local institutes.

8.3. Fieldwork monitoring

8.3.1. Weekly progress reports

During the fieldwork phase, an overview over the number of interviews reached in each cell of the sampling matrix should be collected by the coordinating agency on a weekly basis. For this purpose, a reporting template has to be developed. In this sheet, the targets for each cell should be shown and fixed. Figures on the number of interviews per cell actually reached have then to be completed by the local cooperation partners (see Table 7 as an example, taken from the pilot survey).

Table 6. **Reporting template for weekly progress reports (Germany)**
Only the fields labelled 'actual' have to be completed; the targets are fixed, the percentage is calculated automatically

NACE sector		5 to 9 employees	10 to 49 employees	50 to 249 employees	250 or more employees	Total sector
Sector cluster NACE 28, 29, 30	target	40	45	50	25	160
	actual	5	5	5	5	20
	in %	13%	11%	10%	20%	13%
Sector cluster NACE 41, 42, 43	target	40	50	42	8	140
	actual	7	7	7	7	28
	in %	18%	14%	17%	88%	20%
Sector cluster NACE 46, 47	target	35	45	45	15	140
	actual	8	8	8	8	32
	in %	23%	18%	18%	53%	23%
NACE 62	target	45	50	40	5	140
	actual	4	4	4	4	16
	in %	9%	8%	10%	80%	11%
NACE 64	target	35	40	45	20	140
	actual	3	7	4	2	16
	in %	9%	18%	9%	10%	11%
NACE 84	target	30	40	40	30	140
	actual	3	10	4	7	24
	in %	10%	25%	10%	23%	17%
NACE 86	target	35	40	35	30	140
	actual	4	6	11	3	24
	in %	11%	15%	31%	10%	17%
All sectors	target	260	310	297	133	1 000
	actual	34	47	43	36	160
	in %	13%	15%	14%	27%	16%

Source: Cedefop pilot survey 2012.

8.3.2. Non-response information per cell of the sampling matrix

Every two to three weeks, and after finalisation of fieldwork, additional information on the progress per cell of the sampling matrix should ideally be collected:

- number of interviews reached in the cell;
- number of addresses still available for the cell (in the drawn gross sample);
- number of definite quality neutral non-responses for the cell (i.e. line dead, wrong telephone number, line always busy, etc.);
- number of addresses out of scope (wrong sector, < 5 employees, private household);
- number of definite refusals for the cell (refusals for any reason);
- incomplete interviews;
- number of appointments (ideally differentiated between weak and fixed appointments).

This information facilitates the central control over the fieldwork period and the anticipation of any fieldwork difficulties.

8.3.3. Final field reporting

The coordination team will write an integrated technical report on the fieldwork for the survey. To this end, the local fieldwork agencies need to deliver the following input:

- (a) information on the start and end dates of fieldwork;
- (b) a documentation of all difficulties occurring in the preparation, fieldwork or data-coding phase;
- (c) the number of interviewers working on the project;
- (d) proposals for an improvement of the survey (if there are any);
- (e) an overview of the item non-responses for the survey.

Table 7. **Template for recording final non-response reasons**

Gross sample I: total number of addresses bought for the survey	
• Addresses not used/not needed	0
▼	
Gross sample II: total number of addresses used/'touched' for the survey	
• Establishment/local unit does not exist	0
• Line dead; fax/modem, wrong telephone number	0
• Telephone not answered after a minimum of 10 futile contacts	0
• Line always busy or answering machine	0
• Private households (out of scope)	0
• Wrong sector according to questionnaire	0
• Wrong size (less than five employees according to questionnaire)	0
• Quota for the cell already full	0
▼	
Gross sample III: total number of addresses with contact	
• Refusal to participate in the interview	0
• No interview possible within fieldwork period	0
• Interview incomplete	0
• Other non-responses	0
▼	
Completed interviews	0
Net interviews as % of gross sample II	0%
Net interviews as % of gross sample III	0%

Source: Cedefop pilot survey 2012.

8.4. Measures to improve response rates

Achieving reasonably high cooperation rates is important for the success of the survey. If the survey is going to be carried out as a survey with voluntary participation, response rates will inevitably be considerably lower than for a survey with obligatory participation. Especially in surveys among organisations, response rates have been dropping consistently in the last couple of years.

Against this background, all reasonable measures should be applied to maximise response rates (Groves, 1989). In the following, some possible measures will be discussed:

- (a) as pointed out in Section 8.2, a sufficiently long fieldwork period should be foreseen to allow for a high flexibility in terms of the scheduling and postponement of individual interview appointments;
- (b) interviewing should be started with a small number of addresses per cell only. Only if it is clear that this number will definitely not be sufficient to reach the envisaged target, additional addresses should be drawn for the cell. This kind of staggered sample release can either be administered centrally or it can be implemented at the local level;
- (c) before an address is classified as a definitive non-response, at least 10 contact efforts should be undertaken at different days of the week and at different times. The contact efforts should stretch over a couple of weeks before classifying an address as final non-response;
- (d) a thorough interviewer refusal training before the launch of the survey will help to improve cooperation rates. In the entry part of the questionnaire and in the hints for the interviewer briefing, a number of possible arguments are listed that can help to increase cooperativeness;
- (e) a motivation letter should be drafted. It should include information on the client institution of the survey, on the survey aims and on the most appropriate respondent within the organisation. Also, data protection measures applied for the survey data should be mentioned. It is recommended to send this letter out as a fax or e-mail on the request of potential respondents or if interviewers consider the letter to be helpful for convincing people to participate. Sending the letter in advance to all establishments within the gross sample is generally also possible (provided that the address source contains the full postal addresses of the units). But this is a costly measure and its added value is questionable since the recipient of the letter and the (first) contact of the telephone call might often not be the same. The motivation letter should ideally be (co-)signed by representatives of the client institution and by representatives of relevant

- sector associations. Annex 2 to this handbook shows the motivation letter used for the pilot employer survey as an example;
- (f) in addition to the motivation letter, it is advisable to provide information on the survey on the webpage of the client institution, preferably in each language of the study. This enables respondents to get some more insight into the aims of the survey and into the client institution and the type of research it is dedicated to. The information also serves as a verification of the seriousness of the survey. Often, respondents have the possibility to directly open the indicated webpage while still talking with the interviewer;
 - (g) a short and interesting survey instrument is an important prerequisite to obtaining high response rates and only a low share of prematurely terminated, incomplete interviews. In surveys among organisations, respondents usually ask for an indication of the time needed for answering the survey. The respondent should not be misled about the assumed duration. For an interview taking 15 minutes, cooperation is already much more readily granted than for a 20- or 25-minute interview. The pilot interviews took slightly less than 20 minutes on average. The successful shortening of the questionnaire after the pre-test was reflected in the subjective impression of respondents and interviewers: in contrast to the pre-test (which had an average duration of 22 minutes), the interview duration and a monotony of the interviews were not an issue of major complaints in the pilot;
 - (h) being able to offer an additional online version of the questionnaire might convince one or the other additional person to participate in the survey. The advantage of an online version is that, in this way, the questions directed to the specific occupational group can be easily passed on to the person who is most familiar with the work of this particular group. There is, however, a danger that most of the respondents asking for the online version might not complete the survey because of its complexity, the repetitiveness of several question blocs and the overall duration of the questionnaire. Similarly, respondents with limited interest in the topic might terminate the survey without completing it, which in turn could lead to a greater non-response bias;
 - (i) another possibility way to improve response rates is to offer respondents a summary of results, preferably even before the publication of the results to the public. This may have an additional motivating effect.

8.5. Specific requirements for conducting the survey as a repeat survey

Designing the survey as a potential repeat survey, however, requires particular care in the set-up of the survey instrument since trends can be observed only when for each repetition the same basic survey design and questionnaire are used. Therefore, preparations for the first wave need to be done with particular care and any major changes to be introduced to the existing pilot survey concept should be pre-tested (including changes to the order of items). In addition, some further precautions have to be taken for the design of a repeat survey:

- (a) the sample universe needs to be defined carefully since it should be the same for all waves of the survey. This includes the definition of size thresholds, of the sectors to be included as well as of the occupational groups to be covered;
- (b) the targeted sample should be representative for the sections of the economy it covers and large enough to enable statistically sound inferences;
- (c) it needs to be ensured that the procedures used for the first wave are carefully documented to expedite replication of the survey with exactly the same design and according to the same principles;
- (d) the translations elaborated for the first wave need to be stored and used again for any future survey cycles (except for translations that turned out to be faulty).

CHAPTER 9.

Data handling

9.1. Data checks

9.1.1. Central check of interim data sets

If a decentralised programming arrangement is chosen, the main responsibility for an error-free questionnaire instrument is with each local fieldwork institute. In view of the complexity of the survey instrument, it is strongly recommended to add a centralised review of the programmed questionnaires. This can be done in form of a check of the first real interview data: after the first few dozen of interviews, an interim data set should be requested from all national fieldwork institutes. This interim data set can then be centrally checked by the coordination team, with a programmed check syntax confirming the following aspects:

- (a) is all information stored at the right place, in the columns foreseen in the central data map;
- (b) is an answer code entered for each of the questions;
- (c) is the answer code contained among the codes foreseen according to the data map and the master questionnaire;
- (d) are the answers to the open numerical questions (Q100c and Q110 in the pilot master questionnaire) within the allowed ranges;
- (e) for all single-punch questions (questions for which only one answer item is to be ticked): are there any interviews where more than one answer is ticked;
- (f) does the mechanism selecting the occupational group choose one of the existing occupational groups;
- (g) does the selection mechanism provide a balanced selection of occupational groups in all those interviews where more than one of the groups is present in the establishment;
- (h) are all filters correctly applied, is the sequence of all questions correct.

If the data set is not error-free, the local institute has to take steps to correct the questionnaire script accordingly. To ensure that the revised script works correctly, a further data set with the next couple of interviews should be sent to the coordinators for another data check.

9.1.2. Checks of final data set

The check routines done for the interim data checks should be repeated for the final national data sets.

9.1.3. Data and consistency checks embedded in the questionnaire script

Computer-assisted data collection methodologies have the big advantage that it is possible to programme checking routines into the questionnaire script. This is particularly useful for the check of numerical data which is logically interconnected.

In the pilot master questionnaire, numerical information is only collected about two issues: the total number of employees working in the establishment (Q100c) and the number of employees of the selected occupational group working in the establishment. Here, a crosscheck needs to be made during the interview, that the number of employees in the occupational group (Q110) is not higher than the total number of employees in the establishment (Q100c). This check is already been part of the pilot questionnaire script and is described in the programming instructions of the master questionnaire (Annex 1).

9.2. Coding of open-ended answers

The pilot survey contained four questions with open-ended items that needed to be coded:

- (a) Q109: largest occupational group in the establishment; to be named only if no pre-selected group is present;
- (b) Q304: newly-emerging task(s);
- (c) Q306_4: application of other measures to address newly-emerging tasks;
- (d) Q405: occupational group undergoing the greatest changes (if not the selected group).

Two of these open-ended questions (Q109 and Q405) are recommended for deletion in any replication of the survey. As for the two remaining open-ended answers, it is recommended to use the same principles for the post-coding as applied for the pilot survey. In the pilot survey, all verbatims with the answers to the open questions were first translated into the English language by a professional translation office. Then, the verbatims were centrally checked and coded. It is suggested to follow this procedure and to code the open questions according to the following principles.

9.2.1. Q304: newly-emerging tasks (not mentioned in the previous task questions)

In the coding, the newly-emerging tasks (Q304) should be attributed to the generic tasks of the questionnaire. Some of the newly-emerging tasks named in this question can probably be attributed to more than one generic task. The coding should therefore allow for up to three generic tasks being attributed to one

newly-emerging task. For newly-emerging tasks that do not fit with any of the generic tasks from the questionnaire, new codes should be created, possibly including a miscellaneous occupation-specific category for counting purposes.

9.2.2. Q306_4: other measures applied (to address newly-emerging tasks for the occupational group)

This open question asked for 'other measures applied' (to address newly-emerging tasks for the occupational group). The 'other' in this case refers to the types of measures presented in the preceding closed questions Q306_1, Q306_2 and Q306_3. During preparation of the coding for the pilot survey it turned out that in many cases, the measure(s) named under 'other measures applied' did not present any new measures, but just specifications of the measures from the closed questions Q306_1 to Q306_3. It was therefore decided to create a specific code for these cases, namely code 3 'open answers could be attributed to questions Q306_1, Q306_2 or Q306_3'. In the (few) cases where any of Q306_1 to Q306_3 should have been ticked with 'yes' according to the open-ended answer given in Q306_4, the answer to the respective closed question ('no', 'don't know' or 'no answer') was recoded into 'yes'. These were for example cases where the open-ended text in Q306_4 mentioned some type of training courses, but where Q306_1 'training of available staff' had not been ticked with 'yes'. It is recommended that one apply this recoding step in the same way for any replication of the survey.

Overall, the following codes are foreseen for this question:

- 2 no (no other measure)
- 6 other measure not understandable
- 7 no other measure specified
- 8 don't know (whether any other measure is applied)
- 9 no answer (on the question whether any other measure is applied)
- 11 training (general, specific content, etc.)
- 12 internal reorganisation
- 13 recruitment
- 14 apprenticeships, internships, trainees
- 15 external support
- 16 outsourcing/subcontracting/temporary work/freelancers
- 17 information
- 111 other

9.3. Data outputs

The results of the survey should be processed in form of a micro data set in the SPSS SAV and ASCII formats, preferably as comma-delimited values.

Apart from the micro data sets, it is recommended to deliver a set of standardised tables tabulating each single question, differentiated by the following dimensions:

- (a) country;
- (b) sector of activity;
- (c) size-class of the establishment and number of employees in the selected occupational group;
- (d) selected occupational group.

CHAPTER 10.

The weighting

10.1. Organisation of the weighting process

The weighting of the survey should be done centrally by the coordination team to ensure full consistency of the weighting across all countries. The local agencies should be made responsible for providing the coordinators with the requested statistical figures (distribution of establishments/local unit over the various cells of the sampling matrix). If this information does not exist in exactly the requested form, the best available alternative figures need to be supplied to elaborate estimates on this distribution. If for example, the statistical information is not available for the unit of establishments/local units, then information on the distribution of companies/enterprises needs to be collected instead, as an approximate orientation concerning the size of the universe.

10.2. Establishment-proportional weighting

As described in detail in Section 4.1, it is recommended to use a disproportional sample design for any replication of this survey. The sample should be disproportional with regard to the size of the establishments, with an oversampling of the large units. If the survey is to be conducted in various countries, it is also recommended to choose a design that is not fully proportional with regard to the size of the economy of the countries to be included. Otherwise, a large share of the interviews would have to be conducted in the large countries while for the smallest of the countries the number of interviews would not be sufficient for any solid statistical analysis. Disproportionality with regard to the sector is, however, not a must. It depends on the research aims whether the sectors should better be sampled proportional to their real quantitative importance or whether a further disproportional should be introduced here.

Within this disproportional sample design, weighting of the data is necessary because of two reasons: first, to allow for any representative statements, the disproportional structure of the samples has to be corrected for analysis. Second, experience shows that the willingness to participate in a survey depends, among other things, on the size and sector of the establishment. This pattern can induce disproportional non-responses, which are to be corrected in the weighting procedure.

The establishment proportional weighting factor adapts the net sample so that the distribution of establishments over the cells of the stratification matrix used for sampling is reproduced. Analyses based on establishment proportional weighted data allow one to making statement of the following type: 'X % of the establishments within the defined universe regularly check the skill needs of their employees'. Since most establishments are small and medium-sized units, analyses of this type are strongly influenced by the situation in these smaller units.

10.2.1. Basic structure of the weighting process

Up to three weighting steps are necessary to establish proportional weighting.

The first step applies only to countries using the screener procedure and only to multisite companies. Here an initial selection probability weighting factor has to be introduced. This weighting factor will correct for the unequal selection probabilities of establishments from multisite organisations. The exact value of this weighting factor will be determined on basis of the answers to the screening questionnaire: there, the number of establishments the organisation has within the defined universe in the corresponding country will be asked about. This figure will be the base for the calculation of the selection probability weighting factor. The factor is equal to the total number of establishments the originally selected multisite company has within the limits of the universe defined for the replication of the employer skill needs survey. To give outliers not too big an influence (such as a bank or shop chain might have hundreds of local units spread over the country) and to keep weighting factors reasonably homogeneous, we propose limiting this factor to the maximum value 5, generating trimmed weights. This means that if the interviewed unit belongs to a company with more than five subsidiaries (with five or more employees) the maximum factor 5 shall be applied, regardless of how many local units the organisation has. However, the exact definition of this threshold can be varied. For all single-site companies in screener countries and for all establishments/local units in the countries without a screener procedure the selection probability weighting factor will have the value 1.

The second step adjusts the disproportional structure of the net sample in each country to a strictly proportional structure. This factor can be called a post-stratification weight. This weight actually combines the correction for two different effects: first, the deliberately disproportional sample structure, as described in Section 4.1, and second it corrects any remaining disproportionate unit non-response which could not be controlled during fieldwork. For practical reasons these two corrections should be done in one single working step. That is, by correcting the actual distribution of the net sample (after application of the

selection probability weighting factor) so that it corresponds to the structure of the universe. A theoretically possible split into two separate working steps would mathematically lead to exactly the same factors, if the weighting is confined to the structure as defined by the weighting matrix. Post-stratification weighting will basically use the structure of the sampling matrix (size-class by broad sector of activity) and can include additional adjustments (such as by region or by a finer breakdown of sectors of activity) where reference data are available and where there are enough interviews for all cells of this finer breakdown. Usually post-stratification weights are adjusted so that the weighted number of cases exactly corresponds to the unweighted number of interviews in each country. Another possibility is to calculate the weights as extrapolations to the universe. The respectively weighted extrapolated number of cases then corresponds to the number of establishments existing in reality in the country.

In a third step the different size of the national universes will be considered. These differences are normally only partially reflected by the different target sample sizes. Usually, in surveys of this kind staggered sample sizes are used, for example $n = 1\,000$ interviews for small, $n = 2\,000$ for medium-sized and $n = 3\,000$ interviews for large countries. The application of the population weights allows one to aggregate the data at European level and to make direct (weighted) analyses related to Europe as a whole or to any groupings of countries.

To make analysis more comfortable, these factors can be integrated into one final weighting factor, the establishment weight (WF_{est}), which can be used for all analyses, be it at national level or European-wide. For the calculation of the final weighting factor the following algorithm can be applied:

$$WF = WF1 * WF2 * WF3 \quad (1)$$

where

- WF = final weighting factor (establishment-proportional)
- $WF1$ = selection probability weight, i.e. in the case of multisite companies in screener countries the number of eligible establishments of the originally selected multisite company (limited to a maximum of 5), in all other cases $WF1 = 1$.
- $WF2$ = post-stratification weight, i.e. the correction factor for adapting the actual structure of the net sample to the target structure of establishments
- $WF3$ = population weights or international weights, considering the different sizes of the national universes of establishments

The post-stratifications weights ($WF2$) will be calculated as follows:

$$WF2_{ijY} = (N_{ijY}/n_{ijY}) * (n_{totY})/(N_{totY}) \quad (2)$$

where

$WF2_{ijY}$	=	post-stratification weighting factor for cell ij of the sampling matrix in country Y
N_{ijY}	=	total number of establishments in cell ij of the sampling matrix in country Y
n_{ijY}	=	total number of completed interviews in cell ij of the sampling matrix in country Y
n_{totY}	=	total unweighted net sample size in country Y
N_{totY}	=	total number of establishments in country Y

Population weights or international weights ($WF3$) are constant factors for each country. They can either be calculated as projections to the universe (formula 3a) or be standardised to the total number of net interviews in all countries (formula 3b). In the first case the weighted number of interviews will show the total number of establishments (or employees) in the universe (nationally and European-wide). In the second case the weighted number of cases will add up to the total net sample size if one looks at the entire European sample, while weighted national figures will be different from the national net sample sizes. We recommend using the first option (formula 3a) for calculating the establishment-proportional weights. The two alternative ways of calculating the international weights will only affect the total number of weighted interviews, but the decision for one of them does not have any effect on the obtained distributions. As far as the distributions are concerned, both alternatives (3a) and (3b) can be used and will produce exactly the same results. Population weights or international weights ($WF3$) will be calculated as follows:

$$WF3Y = (N_{totY}/n_{totY}) \quad (3a)$$

or

$$WF3Y = (N_{totY}/n_{totY} * n_{toteurope}/N_{toteurope}) \quad (3b)$$

where

n_{totY}	=	total unweighted net sample size in country Y
N_{totY}	=	total number of establishments (or respectively, employees) in country Y
$n_{toteurope}$	=	total unweighted net sample size in Europe (i.e. in all countries involved)
$N_{toteurope}$	=	total number of establishments (or respectively, employees) in Europe

10.2.2. Practical implementation of the weighting process

For the weighting, the matrix used for the sampling will be used. For the pilot, this matrix had 28 cells that were defined by the seven sector clusters and the four size-classes covered with the pilot. For any replication of the survey, the sampling and weighting matrix will need to be adapted to the relevant universe.

Table 8. **Weighting matrix**

NACE Rev. 2 sector (cluster)	5 to 9 (or 1 to 9) employees	10 to 49 employees	50 to 249 employees	250 or more employees
Sector cluster 1				
Sector cluster 2				
Sector cluster 3				
Sector cluster 4				
Sector cluster 5				
Sector cluster 6				
Sector cluster 7				

Source: Cedefop pilot survey 2012.

10.3. Employee-proportional weighting

10.3.1. Employee proportional weights as an additional option

A further issue to be considered is the calculation of an additional employee proportional weight. An employee proportional weight referring to all employees within the establishment does not make sense for this survey since most of the questions refer to a particular group of employees only and not to the whole workforce. The selection mechanism for the occupational group applied in Q107 of the pilot master questionnaire selects one of the existing groups at random and not according to any quantitative criteria. An establishment with 1 000 employees might have just one employee of the selected occupational group while another establishment with five employees in total might have five employees of that particular group. Calculating an employee proportional weight based on the overall number of employees in the establishment would therefore be misleading.

There is, however, the possibility to calculate an employee proportional weight that takes into account the size of the selected occupational group working in the establishment. For the calculation of such a weight, the questionnaire of the pilot employer survey on skill needs contains a question asking for the number of employees of the selected occupational group within the

establishment in numerical terms (Q110). There are two options on how to use this information for the weighting.

10.3.2. Option 1: use of the numerical value on number of employees in the occupational group as factor

In this case, the total number of employees within the occupational group would have to be summed, using the establishment proportional weight. Then, the individual answers to Q110 would be divided by this overall value.

An example

The pilot data set contains $n = 369$ interviews where the reference group were the assemblers. The number of employees in this group ranges from one employee (in 15 establishments) to 800 employees (in two establishments) in the unweighted data set. In total, the employees in this group sum up to 14 115 employees, again unweighted.

Using the extrapolation weight, it is 1 328 establishments with just one employee in the occupational group and 23 establishments with 800 employees. In total, with the weighting the employees in this group sum up to 325 624 employees.

In this example, each interview from an establishment with just one assembler would count with 1 and each interview with 800 assemblers with 800, so these interviews would get 800 times the weight as the ones with just one assembler. Though the situation in these establishments with many assemblers is in fact much more important for the labour market of the assemblers than the situation in the establishments with just few assemblers, the differences are sufficiently large that it causes some uneasiness for implementing this type of weighting. This is especially the case because respondents in large establishments already had to construct an average when answering the questions on tasks – possibly the situation of perhaps 200 of these 800 establishments is quite different from this average situation. Further, outliers with unusual answer behaviour could get a very large influence over the overall results for an occupational group, particularly if the number of interviews available per group is only small.

10.3.3. Option 2: recoding the numerical values of Q110 into size bands

Another option for an employee proportional weighting related to the occupational groups would be the grouping of the number of employees named in Q110 into size bands. These size bands could then be taken for a weighting of the answers. To this end, the data set already contains a variable Q110_grp with the following values:

- code 1 = 1 to 4 employees within the selected occupational group;
- code 2 = 5 to 9 employees within the selected occupational group;
- code 3 = 10 to 49 employees within the selected occupational group;
- code 4 = 50 to 249 employees within the selected occupational group;
- code 5 = 250 or more employees within the selected occupational group.

The code for this variable could then be taken as a weighting factor. In this way, the data from an establishment with just one, two, three or four employees in the selected group would be counted with the factor 1, data from establishments with 250 or more employees would receive the factor 5. Thus, the problem of outliers influencing the results too much would be avoided while the situation in the establishments with many employees of a specific occupational group would still influence the overall results more than the situation in establishments with just very few employees within a group. The major drawback of option 2 is that the weights (1 to 5 in the proposal above) are quite arbitrary – instead of the values 1 to 5, one could also set other values or define the size bands differently. Therefore, a kind of compromise between the two options would be to implement option 1, but cut outliers rigorously at a quite low threshold to give trimmed weights, for example at 20 employees. Another possible compromise would be to take the calculated average value for each of the size bands shown above as factor so that the weight of the outliers would to a certain degree also be reduced.

Any employee-proportional weighting related to the occupational groups, using options 1 and 2 and the indicated possible compromises, will remain problematic if the number of interviews available for each occupational group is not larger than in the pilot survey.

The use of the information from the questionnaire is the only possibility to calculate an employee proportional weight referring to the selected occupational group when there is no official statistical information available for a weighting by the number of employees within a particular occupational group. The LFS provides information on the various occupational groups that employees belong to, and an indication on the size of the workplace/establishment. However, it does not provide information on the number of employees of the same occupational group working in the establishment. Neither is this information available for any other European-wide data source. One possible way to cope with the lack of statistical data on the distribution of occupational groups would be to modify the questionnaire and ask for the number of employees in each of the pre-selected groups, not only the one finally chosen as reference group for the further interview.

10.4. Statistical information required for weighting

For the establishment-proportional weighting described above, reliable statistical data about the distribution of establishments (or companies) in a country are required. This information needs to be collected in the same breakdown by size-classes and sectors of activity that has been used for the sampling.

The most reliable source for statistical information about the distribution of establishments or enterprises is usually the respective national statistical office. In some countries the statistical offices, however, only collect statistical information on the number of enterprises, not the number of establishments/local units. In these countries, other sources such as the national labour agency, social insurance providers or the statistics of commercial address providers need to be investigated. If these sources do not provide the required type of statistics either, estimates about the distribution of organisations will have to be made.

To our current knowledge, in almost half of the EU Member States, reliable statistical figures on the distribution of establishments/local units are not available from any of these sources. The Eurostat structural business statistics data provide counts of establishments in sectors by NUTS2 region, but unfortunately they are not broken down by establishment size which limits their use in the present context.

Even in the countries where statistical information about the distribution of establishments/local units and/or enterprises/companies is generally available, this information is not necessarily complete. For example, the databases of national statistical offices sometimes do not contain reliable information on the distribution of organisations in the public administration.

For the weighting of the data from a cross-national employer skills survey, it is important to apply the same weighting principles to all countries included in the survey. This implies, for example, collecting data with exactly the same size bands and with the NACE sector classification used in the sampling matrix. Where these data are missing, the best estimates need to be elaborated based on existing data that can give an indication on the distribution of establishments.

10.5. Design effects, effective sample size, weighting and standard errors

The sampling design described above implies unequal inclusion probabilities at two stages. The first stage is the stratification of the sample resulting in a disproportional structure of the net sample. The second stage applies to countries using the screener procedure only: as described above the establishments from multisite companies are characterised by a lower selection

probability than single-site companies. To adjust for these unequal selection probabilities design weights are used, which are part of the final weighting factors.

Although the weighting corrects biases that would be a result of the unequal selection probabilities, it has to be considered ‘that for weighted estimates the reduction in bias may also bring increases in the variances due to weighting’ (Kish, 1990, p. 129). The corresponding increase in variance can be measured by the design effect which is ‘the ratio of the sampling variance (squared standard error) of a particular sample estimate using a specified (non-simple, random) sample design to the sampling variance for the same estimate based on a simple random sample with the same number of cases’ (Frankel, 2010, p. 97). Thus, it measures ‘the inflation of variance of an appropriate estimator for a population parameter θ under a complex sample design compared to the variance of an appropriate estimator for the same parameter under simple random sampling with replacement’ (Ganninger, 2011, p. 966). The design effect due to unequal selection probabilities ($deff_p$) can be estimated according to the following formula (Ganninger, 2011, p. 966):

$$deff_p = n \cdot \frac{\sum_{i=1}^n w_i^2}{\left(\sum_{i=1}^n w_i\right)^2} \quad (4)$$

where:

n = unweighted sample size

w_i = design weight of element i .

The variance inflation implies that the standard errors of an estimator are too small and thus confidence intervals are too narrow, if computed on the basis of the actual sample size. This can be overcome by using the effective sample size (n_{eff}) to compute the standard errors. It ‘is defined as the actual sample size (number of cases) divided by the design effect’ (Frankel, 2010, p. 97), that means:

$$n_{eff} = \frac{n}{deff_p} = \frac{\left(\sum_{i=1}^n w_i\right)^2}{\sum_{i=1}^n w_i^2} \quad (5)$$

For example, assume the unweighted sample size n comprises 1 000 cases. If one finds that 60% of the establishments in the sample report an increase in task x for their occupational group y , the 95% confidence interval is $\pm 3\%$, so the true value can be thought of as lying somewhere between 57% and 63%. But when according to some unequal inclusion probabilities one faces a design effect of, for example, 1.25, the effective sample size n_{eff} is only 800. That means that the 95% confidence interval will be broader, having now a value of $\pm 3.4\%$, so we can infer that the true value lies somewhere between 56.6% and 63.4%.

CHAPTER 11.

Possibilities for analysis of the data

The employer survey on skill needs in Europe depicts current and future skill needs of establishments in countries, sectors and within different occupations. It provides information on the following issues relevant to a range of policy priority areas:

- (a) importance of generic tasks from the point of view of employers;
- (b) relevance of VET;
- (c) ageing and its impact on skills as perceived by employers (where feasible);
- (d) changing skills in the context of innovation and environmental awareness/regulations (looking at qualifications required, change in importance and preparedness for tasks that are increasing in importance);
- (e) changing skills in relation to training and identification of skill needs by employers;
- (f) the changing and emerging skill needs in sectors of particular policy interest;
- (g) the changing and emerging skill needs of small and medium-sized enterprises (SMEs);
- (h) current and prospective skill shortages (hard-to-fill vacancies, preparedness questions, etc.).

The data set can be analysed in terms of univariate, pairwise and multivariable analyses. To make sound inferences concerning estimated quantities it is important that standard errors or confidence intervals also be provided. These measures of precision enable one to place a given finding in the context of the uncertainty arising from its corresponding sampling and measurement error, to separate the 'signal' from the background 'noise'. Example analyses with indicative findings that may serve as suggestions for future research are provided in the companion summary report on the pilot employer survey on skill needs (Cedefop, forthcoming).

11.1. Univariate descriptive analysis

Univariate analysis is the simplest form of analysing the employer survey data set. The analysis is carried out with the description of single variables and it attributes to reveal first hints and findings. Further, this analysis can be the basis for further pairwise and multivariable analyses.

To obtain a first overview of the results of the survey, tables for all variables should be provided. The major objective at this stage is to reveal first findings on the basis of frequency distribution concerning developments in tasks and key drivers of change. Measures of central tendency (mean, mode, median) and/or range and standard deviation might be used for analysis depending on the context. Moreover, this stage might serve as a basis for further selection of variable sets for pairwise cross-tabulation and analysis.

11.2. Pairwise analysis

Pairwise analysis involves the analysis of two variables for the purpose of determining the empirical relationship between them. The major differentiating point between univariate and pairwise analysis, in addition to looking at more than one variable, is that the purpose of a pairwise analysis goes beyond being simply descriptive. These analyses can be helpful in testing simple hypotheses of association. In addition, it may prove useful to use third-level split variables where there is reason to believe that two-way associations vary systematically between groups, stratified two-way tables. Such multiway tabulations are a useful preliminary to statistical models with multiple variables, as well as being a useful device for summarising interactions between sets of variables.

These pairwise analyses might focus on some key outcome indicators such as levels and changes in the importance of skills, and the extent of preparedness. To capture change, one possibility is to derive a variable for the number or proportion of skills that are increasingly important. To capture preparedness, a possibility is to derive a variable for the proportion of changing skills for which employers are prepared.

For the start of the two-way analysis of the survey cross tabulations might be produced and analysed for a selected number of variables (based on hypotheses see below) by:

- (a) country;
- (b) establishment size;
- (c) sector;
- (d) occupational group.

These cross tabulations for different variables by countries, establishment sizes, sectors and occupational groups might include, for example:

- (a) Q114 and Q115, education level grouped into low, medium and high;
- (b) older versus younger workforce;
- (c) innovative versus non-innovative establishments;
- (d) Q105 reviewing skill and training needs;

- (e) Q113 hard-to fill vacancies for this occupational group;
- (f) Q104 geographical market;
- (g) Q106 participation in paid continuing vocational training.

On the basis of the cross tabulations for these variables researchers might look for differences between countries, establishment sizes, sectors and occupational groups to draw conclusions on the impact of different VET systems, on differences concerning human resources management activities of big establishments and SMEs as well as on VET activities in different sectors and for different occupational groups. This includes the comparisons of those establishments with VET activities with other non-VET training activities and the notion of training versus non-training establishments.

Additionally, the data can be analysed along three pillars:

- (a) generic tasks;
- (b) occupation-specific tasks;
- (c) newly-emerging tasks.

11.2.1. Generic tasks

Apart from general frequencies and the cross tabulations on this topic (Q200-Q242) mentioned above, the following list show examples of some worthwhile research questions:

- (a) are there differences for the outcome variables between innovative and non-innovative establishments (Q400)? It might be expected that innovative establishments show more activities in VET and more dynamics in the development of tasks;
- (b) are there any differences concerning skill needs reviewing versus non-reviewing establishments and generic tasks especially in terms of preparedness? It might be expected that establishments reviewing skill needs show more activities in VET and for this reason are better prepared for tasks gaining in importance;
- (c) are there any differences related to environmentally aware versus non-aware establishments and generic tasks? Do these establishments show other sets of tasks compared to the others? This can also be interlinked with the variable on potential impact of the adaptation of practices, products or services due to environmental awareness;
- (d) do training and non-training establishments show differences in generic tasks and skills especially in terms of preparedness? Here hints for VET activities can be expected in different countries, establishment sizes, sectors, occupational groups;

- (e) do establishments with different age structures in particular occupational groups show varying sets of tasks? Do tasks depend in any way on establishment size?

11.2.2. Occupation-specific tasks

Apart from general frequencies and the cross tabulations on this topic (Q300/Q301) mentioned above, the following research questions can be focused on:

- (a) do training and non-training establishments show differences in occupation-specific tasks and skills especially in terms of preparedness? Here hints for VET activities can be expected in different countries, establishment sizes, sectors, occupational groups, etc.;
- (b) are there differences concerning innovative versus non-innovative establishments for occupation-specific tasks, especially in terms of preparedness? It might be assumed that innovative establishments show more activities in VET and more dynamics in the development of tasks;
- (c) are differences of skill needs present for reviewing versus non-reviewing establishments and occupation-specific tasks, especially in terms of preparedness? It can be assumed that establishments reviewing skill needs show more activities in VET;
- (d) does one observe any differences concerning environmentally aware versus non-aware establishments and occupation-specific tasks? Do these establishments show other sets of tasks than the others;
- (e) do establishments with vacancies show other sets of tasks than those with no vacancies;
- (f) is it the case that establishments with very different age structures in particular occupational groups report other sets of tasks?

11.2.3. Newly-emerging tasks

Apart from general frequencies and the cross tabulations on this topic (Q303-Q308) mentioned above, the following research questions can be explored:

- (a) do training and non-training establishments show differences in newly-emerging tasks, especially in terms of preparedness? This would underline the need for VET. Which ways of addressing these tasks do they choose? Do they have problems in finding offers and providers? If there are problems, this would be a hint for VET providers to adapt their offers. Do they report problems in recruitment? ... in different countries, establishment sizes, sectors, occupational groups ...;
- (b) are there differences concerning innovative versus non-innovative establishments and newly-emerging tasks especially in terms of

- preparedness? It might be assumed that innovative establishments show more activities in VET and more dynamics in the development of tasks;
- (c) do innovative establishments show other kinds of newly-emerging tasks than non-innovative establishments? If so, this would be an evidence for a direct link between innovation and the need of VET and would offer new information for many stakeholders;
 - (d) are there any differences concerning skill needs reviewing versus non-reviewing establishments and newly-emerging tasks especially in terms of preparedness? It can be assumed that establishments reviewing skill needs show more activities in VET;
 - (e) do we see find any differences concerning environmentally aware versus non-aware establishments and newly-emerging tasks? Do these establishments show other sets of tasks than the others;
 - (f) do establishments with vacancies show other sets of tasks or fewer newly-emerging tasks than those with no vacancies;
 - (g) are establishments with very different age structures in particular occupational groups reporting other sets of tasks?

The qualitative data on newly-emerging tasks require a different data analysis strategy than the quantitative data. In particular for the kinds of newly-emerging tasks (Q304) methods of qualitative and quantitative content analysis are needed and should be used. In this context, central procedures of qualitative content analysis, inductive development of categories and deductive application of categories should be tested. Starting with an inductive development of categories, research might look for possible clusters or structural similarities of tasks. If inductive development of categories turns out to be without success, categories can be applied in a deductive way, for example the categorical system of domains of generic tasks (cognitive, interaction/social, physical, learning, green, self-direction) or the 18 generic tasks themselves to derive trends or directions of future developments within occupational groups. Software packages such as TextPack or TextQuest (mainly quantitative) or Atlas.ti or MaxQDA (mainly qualitative) may prove useful in this context. The summative content analysis involves counting and comparisons (keywords/content), and an interpretation of the underlying context.

Detailed qualitative analyses can be produced for all sectors and occupations included on the basis of the open-ended items concerning newly-emerging tasks.

However, it is of crucial importance to interpret the data on generic, occupation-specific and newly-emerging tasks not in isolation but jointly with reference to the corresponding occupational group. By means of a detailed

analysis of sectors, general developments in task profiles consisting of generic, occupation-specific and in particular newly-emerging tasks can be identified. The latter, especially, will have the potential to suggest future trends in the development of tasks and skills.

In cases where the survey is conducted in two or more countries, differences between countries in the relationships between variables could also be of interest. For example, it might be found that the skill needs for environmental activities are being better anticipated, or better prepared for, in some countries than in others. Such a hypothetical finding would, for example, have policy implications for those countries where greater proportions of establishments are not well prepared.

11.3. **Analyses with multiple variables**

Where significant and predictable associations are found between two variables in pairwise analyses one cannot immediately infer that these are causal relationships. It is quite possible that, additional variables may be affecting both of the variables in a pairwise analysis, or that there is a process of reciprocal causation, in both directions.

A useful way to begin to address the issue of causation in a systematic way is to propose and specify a model of behaviour, stating which drivers are expected to affect the outcome variables. For example, for changing skill needs the underlying model might be that technology, including how innovative the company is, and organisational features are among the factors driving the changing use of skills. Given this theory, one would then assume a particular specification about how the causal variables are affecting the outcomes. Usually one assumes a linear component for the model, unless there is a good reason to expect a non-linear relationship, simply because many relationships can in any case be approximated by linear terms (possibly supplemented with polynomial terms), particularly within the wider family of generalised linear models. With this assumption, and the standard statistical assumptions about the distribution of unobserved and unknown factors, one arrives at statistical analyses based on multiple regressions. Explanatory variables of interest for such analyses could include size, sector, country, occupational group, changes in the company, training/non-training establishment, and so on. Dummy coding of categorical explanatory variables enables them to be entered as discrete factors rather than continuous covariates in the model.

The response variable in this instance will not always be a continuous outcome variable. Indeed, the outcome will sometimes be a categorical variable

(skill increased/skill not increased) or a count variable (such as number of skills increased). In these cases, statistical theory provides a number of conventional estimation methods. Where variables are categorical, with a relatively small number of scale points (such as with individual tasks), it will be appropriate to use ordinal logistic or ordinal probit methods. With just two outcomes, logistic or probit methods are appropriate. Poisson distribution based count models might also be appropriate in the case of count variables, where the count is not too large. Several commercially-available statistical packages render the estimation of such models comparatively straightforward, avoiding the need for researchers to programme estimation algorithms. However, the use of off-the-shelf packages has to be combined with an awareness of the assumptions underlying such methods.

Statistical models with multiple explanatory variables allow one to control for the confounding effects of other variables, and so attempt to isolate the effect in which one is interested. For example, if one is interested in the impact of innovation on skill needs, it is important to control the effects of company size or sector. Typically, the raw association estimates between innovation and skill needs will differ from the partial estimates from the regression analysis, which controls for other variables.

To take another example, if one is studying the differences in skill needs between countries, one might wonder whether the observed differences are due to differences in the industrial structures of the two countries, or whether there is something about the cultures of the two countries that leads managers to do things differently even in the same industries. In that case, one could fit statistical models incorporating both a set of industry dummy variables and a set of country dummy variables alongside other control variables. The focus then would be on the estimated coefficients of the country dummy variables, both with and without controlling for industry differences. Even here, however, care would have to be taken in estimating the standard errors of the country dummies, to take into account possible correlations within each country. More parsimonious models can be obtained using multilevel models where random coefficients take into account between country differences and enable the fitting of additional country-level variables. Random coefficient models also allow the estimated effects of an explanatory variable such as innovation to vary across countries, rather than assuming a constant effect.

It is important to bear in mind the limitations of what can be tested in a cross-section analysis. Although one can control for observable variables, it is also likely that there are other confounding factors that are not observed since it has not been possible to acquire data about them through the survey. This limitation is partly owing to the practical limitations of survey time, but it also arises from

the data collection process. A particular problem is that the same person, the interviewee, is informing both about the independent variables and about the outcomes. If the informant is misinformed in any way, or is affected by social esteem bias, the error could affect both outcome and some independent variables, thereby causing spurious correlations. This problem of common observer bias is just one example of the biases that can arise from unobserved fixed effects. Unobserved fixed effects are a widespread limitation on the extent to which causal inferences can be made from cross-section models. Typically, longitudinal data are needed to be reasonably sure of eliminating possible biases in the estimates. In the cross-section analyses that will be used here, one can aspire to reduce biases by including as many control variables (which are not themselves outcomes) as are available and which appear to have a relationship with the outcomes. However, one still has to remain aware of the limitations over asserting that causality has been proved. Typically, it is possible only to show that estimates are consistent with the proposed model and subject matter knowledge is important in the building and interpretation of statistical models.

CHAPTER 12.

Conclusions

The following final discussion and conclusion focuses on technical and practical aspects of the implementation of an employer survey on skill needs. Issues related to content such as the type of information that can be derived from the survey, the usefulness of the results for the policy process and for the modification of VET systems are outlined in the companion report *Piloting a European employer survey on skill needs: illustrative findings* (Cedefop, forthcoming).

12.1. General assessment of the survey concept

Practical experiences with the survey concept piloted in nine European countries in 2012 were generally positive. In spite of its relatively high complexity, the survey generally worked well in practice. The process of selecting a particular occupational group worked, respondents were largely able to answer the questions asked in the survey and to relate their answers to the selected occupational group as intended. Item non-response was thus low, the targeted number of interviews could be achieved and feedback from respondents and interviewers on the survey was mostly positive. As far as can be judged, based on the first set of analyses done with the survey data, answers appear generally in line with expectations. The survey provided well differentiated results on the importance of the investigated tasks and skills for the different occupational groups as well as on recent changes in importance. Also quite promising were the open-ended answers on newly-emerging tasks. Further, the survey provides additional material for further analyses not yet undertaken, for example secondary analyses of the impact of different drivers of change at the workplace level on the development of tasks and skills, using statistical methods that enable one to fit multiple variables.

Some difficulties occurred with the occupation-specific task questions: dropout rates for certain occupations were relatively high because some items were relevant only to a part of the occupational group, but not to all employees of that group (such as the task 'assessing progress during pregnancy and childbirth' is highly relevant for midwives but not for nurses which are both summarised in the ISCO 3-digit level group 322 nursing and midwifery associate professionals). Also, some of the items addressed aspects of the work of the occupational group

that were so basic and general that they sounded rather odd to some respondents. This is because in order to stick to officially-acknowledged job profiles, the compilation of the specific task dimensions was very closely oriented at the ISCO job descriptions provided by ILO and these had to be further summarised due to restrictions regarding the interview length. In this regard, for future implementations of the survey one should consider giving up the aim to cover fully all relevant tasks, as described in the ISCO job descriptions, but rather to concentrate on selected tasks or aspects that seem of particular importance with regard to a potential change in skill demands. To arrive at more meaningful tasks list for the occupational groups, one could utilise expert workshops or focus groups (social partners, occupational associations, training providers) to select the most relevant tasks (prioritised) and reduce the multidimensionality of the ISCO tasks lists. Another possibility would be to focus this part of the survey on different occupations in different years (IT occupations, health occupations, green occupations, etc.) and for that reason to switch to the 4-digit level in this part of the survey, and to adjust or completely rearrange the tasks lists.

Another, more radical option, would be to discard totally the occupation-specific task domains for a future survey. This will however be at the expense of the depth of the data to be derived from the survey and their usefulness for input into the curricula of VET for specific occupational groups. If deleting the occupation-specific tasks, it is worth considering a slightly broader set of generic skills by including some more transversal differentiation concerning technical tasks.

12.2. Key issues for sampling, programming and data collection

The questionnaire concept applied in the pilot survey proved to be very complex. The complexity results from the chosen path of relating the majority of survey questions to a particular occupational group within an establishment, rather than to the whole workforce of an establishment. This way has been chosen to avoid collecting information that would be so general that it would not produce any meaningful insights into skill needs. In view of the large number of existing occupational groups, this is however a path that inevitably leads to a large complexity of the survey concept, even if the survey is restricted just to a number of quantitatively important occupational groups (as in the pilot). Conducting the survey in its present form in all sectors of activity would be technically feasible, but would imply a very high rate of complexity due to the large number of occupational groups that would then have to be taken into account in the

formulation, translation and programming of occupation-specific questions and in the handling of the data. Even if one were to omit the occupation-specific questions, the survey concept remains quite complex when data collection is to be done at the ISCO 3-digit level. Although omitting the occupation-specific questions would save a large amount of preparatory work related to the formulation of questions, the translation process, the programming and the data processing, the selection processes to be done before the interview (pre-selected occupational groups) and within the groups would still be rather complex.

The degree of complexity of the survey can be effectively reduced by two measures. First by a concentration of the survey on a relatively small set of sectors, or respectively occupational groups, either with or without any occupation-specific questions. Or, second, by defining the reference groups at another, more aggregated level such as the ISCO 1-digit level of differentiation. In this case, meaningful occupation-specific questions will be difficult to draft so that it is probably better to limit the survey to a mapping of generic tasks if taking this option. Both measures would evidently have repercussions on the data to be obtained from the survey.

If the same or an even finer level of differentiation than in the pilot should finally be chosen for a later replication of the survey, a centralisation of the survey programming and testing would be desirable when conducting the survey in a larger number of countries. For the pilot, it was not done because the survey was originally meant to be conducted in just five countries and only later expanded to further countries. With centralised programming and data checking, a somewhat higher degree of survey complexity is possible than with a decentralised organisation of these working steps.

12.2.1. Sample size and coverage of the survey

The number of observations that were obtained in the present survey, with its 1 000 cases per country and its limitation to a set of sectors, can be considered as an absolute minimum if the survey is meant to produce results on a national level and for different occupational groups. In the analysis of the survey, cells with less than 20 interviews were not taken into account because of concerns about the generality of these data. This implied that for Ireland, where the sample size was only 500 interviews, finally the observations on a couple of occupational groups could not be taken into account. Similarly, many observations for occupational groups 4 and 5 could generally not be used for a country level analysis because of the applied selection procedure that takes these groups into consideration only if groups 1 to 3 do not exist. If a larger number of sectors of activity should be covered by a later survey, this means that the total number of interviews to be realised per country would have to be improved considerably.

The pilot survey concentrated on a couple of sectors of activity. These were selected to represent very different sectors, among them dynamic and rather stable ones, private sectors as well as largely public sectors and large sectors (in terms of the number of establishments and the share of employment) as well as smaller ones. The survey basically worked in all these sectors. However, it has also become obvious that particularly for the smaller sectors and for the larger size-classes, there are clear limitations as regards the number (and distribution) of achievable interviews. If the pilot survey was to be replicated on a larger scale, with more interviews and covering more sectors of activity, these limitations would have to be taken into account. They result in basically two difficulties:

- (a) in countries with a very limited universe, the number of observations to be obtained within the present survey design might in the end be very small, in some cases so small as not to allow for any statistically assured conclusions on a section of this State's economy. The findings could still be used for an assessment of the tasks and the skill needs within an occupational group and a sector European-wide, but (in case of the small countries) not necessarily for a comparison between each single country. A census approach contacting all establishments, rather than a probability sample, makes sense in this case, though care must be taken not to increase non-response in future surveys due to survey fatigue;
- (b) the universe of very small establishments (five to nine employees) is considerably larger for most sectors than that of the larger establishments. Modifying the sampling matrix, to include even smaller units (those with one to four employees) to cope with the limitations of the sample size, would not necessarily solve the problems related to a limitation of the universe because: (a) the survey has shown that many address registers have considerable weaknesses regarding the degree to which addresses for these small units are updated; and (b) statistical modelling with multiple variables has suggested that the size of the unit has quite an impact on the reported degree of preparedness (though only little influence on the importance and the dynamics of importance).

12.2.2. Unit of enquiry

As discussed, the survey can generally be carried out at both the company and the establishment level. It is however recommended to use the establishment level because knowledge of the concrete tasks and the degree of preparedness of employees to perform these tasks is likely to be available at this level in a more precise and detailed way than from a possibly remote headquarters. Another aspect in favour of the establishment level is that the number of establishments is considerably higher than the number of companies so that, at

the company level, the discussed limitations of the universe become even more severe. Some flexibility with regard to the unit of enquiry could however be allowed for those cases where there is nobody at the establishment level who could answer the questions because all human resources decisions are taken at a central level.

12.2.3. Data collection mode

In terms of data collection modes, it has become clear that the tested survey design requires a computerised administration of the questionnaires. Therefore, any variant using paper questionnaires is not an option for a replication of this survey. A realisation of the survey by CAPI (face-to-face) or CAWI (online) are in turn generally feasible alternatives to the telephone mode, though the former would imply considerably greater costs and the latter problems with sampling and with increased non-response. The method of choice for any replication of the survey would therefore be CATI. Though hybrid approaches using both initial telephone interview contact and CAWI may be worthy of further exploration.

12.3. Outlook: possibilities and restraints for any replications of the survey concept

In the development phase, several crucial decisions had been taken concerning the survey concept to be designed. In the pilot survey, the developed survey design was then extensively tested in nine countries and in a broad variety of sectors of activity. If the concept is to be varied in any key aspects, a further intense pre-testing is highly recommended. This holds particularly in cases where any of the following areas are to be modified:

- (a) the survey was tested with a data collection at the ISCO 3-digit level. For any replication of the survey on a higher aggregated level, it should be thoroughly tested whether the denomination of occupational groups is sufficiently clear to respondents. So that they have a clear understanding about the employees they should relate their answers to; and that this understanding is widely shared by respondents from different types of establishments (different sizes, different sectors, etc.) and across countries. That is, the survey really does deliver reliable and fully comparable results over sectors, sizes and countries;
- (b) a replication of the survey on the even finer 4-digit level differentiation would in turn require less extensive testing because it can be supposed that problems with the understanding of the reference group will not become virulent there. Further, some of the ambiguities in the reference as

- discovered in the pilot will not be applicable there since the 4-digit level uses very precise job titles with which most of the targeted respondents are probably familiar. As far as a replication at the 4-digit level is concerned, the most critical point will be the statistical reliability of the data collected. To reach an acceptably low level of standard errors, a replication at the 4-digit level will either have to use very large national sample sizes (in the magnitude of around 5 000 to 10 000 interviews per country) or it will have to be limited to a small set of sectors of activity and/or occupational groups;
- (c) in the pilot survey, respondents were asked to answer questions on one particular occupational group only. If in a replication of the survey it is intended to ask questions for more than one occupational group, this should be previously tested. Asking the questions for two or more occupational groups might overstretch the willingness and/or capacity of respondents to answer the questions in a precise way and might thus lead to a considerable loss of data quality. Ideally the duration and survey load should be retained, with a target time of around 20 minutes, to minimise non-completion rates;
 - (d) as discussed, the analysis of the pilot data brought to light some unexpected country differences with regard to the answers on the task/skill questions using an importance scale. Before undertaking a replication of the survey in a cross-country context, it should therefore be further examined how far these unexpected differences may have been caused by culturally different interpretations of the importance scale. For a replication of the survey at national level only, this is likely to be a less problematic issue. For example, the large-scale O*net survey in the US, which uses a similar concept (and served as a model for some aspects of Cedefop's pilot employer survey on skills), also uses importance scales, without the problems that can arise by a necessity for comparability of the data. An alternative to the importance scale worth testing is the use of a frequency scale (the frequency with which a task is done respectively a skill applied). The use of the frequency measure had already been considered in the preparation phase of the pilot survey because of its higher degree of objectivity as compared to the importance scale which always implies some kind of subjective assessment. A task exerted only from time to time in a job is however not necessarily a task of minor importance for the job – though both dimensions will often coincide, there is no direct link between importance and frequency. The inclusion of frequency-based anchoring items and anchoring vignettes are worth exploring, if an importance dimension is retained;
 - (e) should a decision be made to limit the survey to particular occupational groups, a limitation that is inevitable if replicating the survey on the 3-digit level, care will have to be taken in the pre-selection of the occupational

groups to be taken into account. As described, in the pilot this selection had been mainly based on quantitative criteria (all pre-selected occupational groups were among the seven most frequent ones in the respective sector). In this case, the quantitative importance of occupational groups within a country or a sector can be derived from an analysis of labour force survey (LFS) data. However, the quantitatively most important groups are not necessarily the ones that are most interesting in terms of the prognosis of future skill demands. If one chooses to select newly-emerging, but quantitatively still-small occupational groups instead, there is in turn the danger that these groups will be encountered only in a small share of the relevant establishments. The survey concept as developed for the pilot is able to detect and map such still-small groups, where the employees of these groups are concentrated on very few establishments (of the selected sectors), the screening efforts are high and many interviews will need to be terminated immediately after the screening should none of the pre-selected (small) groups be present at a given establishment;

- (f) in addition to a pre-selection of occupational groups to be explored in the survey, there is a further selection process to be applied. Once the existence of the pre-selected groups is identified in an interview, a selection will again have to take place among these groups, in all cases where several of the pre-selected groups are present at the establishment. For this purpose a random selection among the pre-selected groups was carried out in the pilot. Other options that were discussed in the preparation phase were a selection by quantitative criteria (the group with the largest number of employees within the establishment) or a selection of the group that is most interesting in terms of task and skills developments from the point of view of the interviewed employers. The latter two options have the advantage of making the interview more interesting and rewarding for respondents since these are then not obliged to answer questions for a particular occupational group which is maybe of minor importance within the establishment. The big drawback of this option is, however, that the resulting data would in no way be representative. If selecting the largest group, data would concentrate on maybe just one or two groups within a sector, at the expense of the other groups (some of the pre-selected groups might hardly be selected at all when using this option for the selection process). Simply selecting the most interesting or dynamic group from the point of view of the employer can lead to survey results that would tend to overemphasise change and might thus mislead political actions derived based on the survey outcomes.

Regardless of the survey concept that might finally be chosen for a possible replication of the survey, it is highly recommended to plan the survey with a broad time horizon in mind. Though the survey concept proved to be promising as regards the results that can be deduced based on an ad-hoc survey, this type of survey will only show its full potential when repeated every couple of years. Such repetitions will enable the observation of changes over time and allow the survey's full potential for forecasting task and skill developments to be subject to a rigorous practical proof: if a larger part of the developments forecast in previous waves show up as strong trends in reality, this will confirm the usefulness of the survey concept and the ability of employers to predict important future skill developments.

List of abbreviations

CAPI	computer-assisted personal interviewing
CATI	computer-assisted telephone interviewing
CAWI	computer-assisted web interviewing
ISCO	international standard classification of occupations
ISCO-08	international standard classification of occupations, approved in 2008
ISCO-88	international standard classification of occupations, approved in 1988
LFS	labour force survey
NACE	European industrial activity classification
NACE Rev.2	new version of the European industrial activity classification
SME	small and medium-sized enterprise
VET	vocational education and training

References

[all links accessed on 28.6.2013]

- Cedefop (forthcoming). *Piloting a European employer survey on skill needs: illustrative findings*. Luxembourg: Publications Office.
- Council of the European Union (2007). Council resolution of 15 November 2007 on the new skills for new jobs. *Official Journal of the European Union*, C 290, pp. 1-3. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/c_290/c_29020071204en00010003.pdf.
- European Commission (2009). Commission recommendation of 29 October 2009 on the use of the international standard classification of occupations (ISCO-08). *Official Journal of the European Union*, L 292, 10.11.2009, pp. 31-47. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:292:0031:0047:EN:PDF>.
- European Commission (2010). Commission Regulation (EC) No 715/2010 of 10 August 2010 amending Council Regulation (EC) No 2223/96 as regards adaptations following the revision of the statistical classification of economic activities NACE Revision 2 and the statistical classification of products by activity (CPA) in national accounts. *Official Journal of the European Union*, L 210, 11.8.2010, pp. 1-21. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32010R0715:EN:NOT>.
- Frankel, M. (2010). Sampling theory. In: Marsden, P.V.; Wright, J.D. (eds). *Handbook of survey research*. Bingley: Emerald, pp. 83-137.
- Ganninger, M. (2011). Designing minimum-cost multistage sample designs. *Joint statistical meeting, Miami Beach, Florida, 30 July to 4 August 2011*. Alexandria: ASA American Statistical Association. http://www.amstat.org/sections/srms/proceedings/y2011/Files/300742_65566.pdf.
- Groves, R.M. (2005). *Survey errors and survey costs*. New York: Wiley. <http://dx.doi.org/10.1002/0471725277>.
- ILO (2007). *Resolution concerning updating the international standard classification of occupations*. Geneva: International Labour Organization. <http://www.ilo.org/public/english/bureau/stat/isco/docs/resol08.pdf>
- Kish, L. (1990). *Weighting: why, when, and how?* Ann Arbor: University of Michigan. http://www.amstat.org/sections/srms/proceedings/papers/1990_018.pdf.

Annex 1.

Final draft questionnaire 'pilot employer survey on skill needs in Europe'

The answer options 'don't know' and 'no answer' and other answer options below the line are not to be read out.

Instructions to the interviewers are highlighted in grey.

Instructions to the programmers are highlighted in orange.

Not all questions have to be answered by each respondent. Filters are set out after the answer categories ('go to'; exit filters) and additionally before the questions (entry filters; see instructions to the programmer). If there is no filter the question which immediately follows is to be asked.

The titles of the modules are not to be read out in the interview. They only serve for indicating the structure of the questionnaire.

Annex 2.

Proposal for a motivation letter

Subject: employer survey on skill needs in Europe

Dear Sir/Madam,

An interviewer of the research company [national TNS company] has recently contacted your establishment to request your participation in a telephone survey that aims at identifying changing tasks at the workplace and skill needs as perceived by employers in Europe. It is being carried out on behalf of and together with Cedefop, the European Centre for the Development of Vocational Training, an agency of the European Union. Cedefop has the strategic objective to strengthen European cooperation in vocational training, and supports the European Commission, Member States and social partners in designing and implementing policies for an attractive and relevant vocational education and training.

What issue is at stake?

To be competitive, the availability of adequately qualified and skilled employees is a main issue for any firm or organisation. To this end, it is essential to have information about relevant developments of skill requirements at an early stage. Therefore the survey asks for increasingly important working tasks and drivers of change. Your answers will help in meeting labour market and employers' requirements more adequately in the future.

Whom do we want to talk to in the interview?

The interviewer would like to talk to the person who has the best overview of working tasks and skill requirements of the different occupational groups working in the establishment. This might be the human resources manager or the managing director respectively or the branch manager of your establishment.

Data privacy protection and anonymity

Participation in the survey is voluntary. However, to ensure that results are representative, it is important to ensure the highest possible participation from the randomly chosen establishments. Therefore, we would be extremely grateful for your valuable input. [National TNS company] and Cedefop will safeguard a strict

anonymity of your answers towards everyone outside these two institutions so that any connection to your establishment or workplace is excluded.

What is your benefit?

We will send you an exclusive short report of the main results before they will be published which will allow you to benchmark your company against others. In case you are interested, please send us an e-mail to Skill.Needs@tns-infratest.com.

Participating in the survey also promises longer-term benefits as the findings will feed into education and training schemes to improve future skills of the labour force.

If you have further questions please contact [name and telephone number of project manager of field institute]. To find out more about Cedefop and its work please visit our website (www.cedefop.europa.eu).

Thank you in advance for your cooperation.

Yours faithfully,



CEDEFOP

European Centre for the Development
of Vocational Training

User guide to developing an employer survey on skill needs

Luxembourg:
Publications Office of the European Union

2013 – VI, 139 p. – 21 x 29.7 cm

ISBN 978-92-896-1359-0
ISSN 1831-5860
doi: 10.2801/25552
Cat. No: TI-BC-13-006-EN-N

Free download at:
http://www.cedefop.europa.eu/EN/Files/5535_en.pdf

Free of charge – 5535 EN –

User guide to developing an employer survey on skill needs

TI-BC-13-006-EN-N

This Cedefop publication is part of a toolkit, including questionnaire and data set, for researchers to carry out employer surveys in international or national contexts. It illustrates an approach tested in 2012 for identifying employer's skill needs, in nine Member States. The guide offers recommendations for a survey covering a range of occupations, sample size calculations, maximising response rates, and ensuring the validity of the instrument as well as subsequent inferences.

The European employer survey on skill needs makes an innovative contribution to linking the world of employers with that of education and training. In providing these materials for further use Cedefop aims to contribute to this process beyond the realm of the present study. The expression of skill needs by employers and cooperation with those who shape VET policy, as well as educators and trainers who develop and implement VET programmes, will play a valuable role in the future.

Europe 123, 570 01 Thessaloniki (Pylea), GREECE
Postal address: PO Box 22427, 551 02 Thessaloniki, GREECE
Tel. +30 2310490111, Fax +30 2310490020
E-mail: info@cedefop.europa.eu

visit our portal www.cedefop.europa.eu

5535 EN
http://www.cedefop.europa.eu/EN/Files/5535_en.pdf

