Martin Baethge / Frank Achtenhagen / Lena Arends / Edvin Babic / Volker Baethge-Kinsky / Susanne Weber

PISA-VET

A Feasibility-Study

Education Science

Franz Steiner Verlag

Martin Baethge / Frank Achtenhagen u. a.

PISA-VET

Martin Baethge / Frank Achtenhagen /
Lena Arends / Edwin Babic /
Volker Baethge-Kinsky / Susanne Weber

PISA-VET

A Feasibility-Study

Franz Steiner Verlag 2006

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN-10: 3-515-08968-3
ISBN-13: 978-3-515-08968-5

Druck: Printservice Decker & Bokor, München.
Printed in Germany

TABLE OF CONTENTS

PREFACE

The present book is based on a research project, which was assigned to the main authors and their institutes (Sociological Research Institute at the University of Goettingen [in charge] and the Institute of Business and Human Resource Education, University of Goettingen) by the German Federal Ministry of Economics and Labour (BMWA).

The concepts of this feasibility study were based on the idea that competence measurement in VET encompasses not only basic cognitive competencies but also aspects of competence development during VET as well as their usability in the labour market. The scope of aspects distinguishes a PISA-VET from PISA in compulsory schools. The task of the feasibility study was to examine whether such a broad approach is feasible or not.

The completion of the feasibility study within the short period of one year was only possible due to the support of an international research community with experts from 17 different countries in two international workshops (Annex A1, A2: participation lists). In this regard, the participation of an interministerial working group, with members of the Federal Ministry of Economics and Labour (BMWA), the Federal Ministry of Education and Research (BMBF), and the Federal Institute for Vocational Education and Training (BiBB), was very helpful. Moreover, the suggestions of our German colleagues, Steffen Kühnel (Institute for Sociological Research, University of Goettingen) and Jürgen Rost (IPN Leibniz Institute for Science Education, University of Kiel), essentially contributed to the outcomes of this feasibility study. We want to express our gratitude to all of them, including Walter Hoffman (head of division, BMWA) for his reliable and competent assistance.

Without the immense efforts of Erika Beller, Barbara Dehne and Ingelore Stahn (secretariat, Sociological Research Institute) as well as Christel Schikora and Heike Jachinke (secretariat, Institute of Business and Human Resource Education) the organisation of workshops and the formatting of diverse texts would not have been possible. We also express our gratitude to all of them.

After all, we hope that the feasibility study "PISA-VET" contributes towards fruitful discussions about the formation of vocational education and training and provides a basis for further developments, particularly with regard to the recent developments of a European Qualifications Framework (EQF).

On behalf of the authors

Martin Baethge
Frank Achtenhagen

1. INTRODUCTION: THE CONCEPT OF A FEASIBILITY STUDY FOR A PISA-VET

1.1 IMPACT, PERSPECTIVES, AND IMPLEMENTATION OF THE STUDY

The political and economic benefits of an international assessment of vocational education and training, implemented as a large-scale assessment, are obvious. Due to increasing internationalisation of economic-exchange relationships in goods and labour markets, as well as economic, political, and social standardisation in Europe under the condition of increasing knowledge intensity in working processes, educational systems have changed. In particular, vocational education systems have gained importance for providing competencies of occupational mobility and independent lifestyle, international competitiveness, and innovativeness. The European Commission has put forward the ambitious economic and social goal of becoming "the most competitive and dynamic knowledge-based economy in the world" (Europäische Sozialagenda, 2001). Improving participation in education and the labour market is therefore, playing a substantial role in reaching these goals. In addition to the development of excellent professional qualifications, VET needs to focus on those competencies which facilitate trans-national mobility and international cooperation.

The European Union has focused on the development of a European Qualifications Framework (EQF) and a European Credit transfer system for Vocational Education and Training (ECVET) to increase transparency in vocational education and training in Europe and facilitate comparability of vocational qualifications. The development of a set of common reference points located in an eight-level framework is one of the results of the Copenhagen Declaration of November 30, 2002 and the Maastricht Communiqué of December 14, 2004 (European Commission, 2005; Sellin 2005).

Regulations of comparability are necessary normative definitions to classify vocational tracks and -qualifications. The "actually" developed competencies in different VET programmes in Europe must be determined to fulfil their principal purpose of providing reliable and confident information for educational institutions as well as institutional and individual actors on the labour market. Regarding the variety of VET forms in Europe, this information will contribute to a mutual learning process. A comparative study of VET, modelled as a large-scale assessment along the lines of PISA (Programme for an International Student Assessment), will facilitate reaching this goal. The aim of a large-scale assessment is to assess the competencies, acquired by young people during VET in their native countries in an objective, reliable and valid way, and to link the results to the macro- and micro-structural factors. These factors, in turn, impact

the development, evaluation and utilisation of competencies at a later point in time in the labour market. Exploring the interdependencies between competencies acquired during vocational education and training and their impact on the labour market is a unique feature of VET in contrast to PISA. Moreover, a PISA-VET could contribute towards explaining the interdependencies between domain-specific vocational and general competencies with regard to individual life goals in different domains (e.g., work, family).

Following PISA, we will adopt the abbreviation "PISA-VET" (Programme for an International Student Assessment for Vocational Education and Training). A feasibility study requires clarification of:

- The methodical conditions and problems which are to be solved in an international, comparative, large-scale assessment;
- the opportunities and problems of an international comparison of vocational education and training systems and -institutions;
- the sample definition of the research design (longitudinal- or cross-sectional);
- the relevant competence dimensions, in particular the relationship between subject-specific and generic skills[1];
- the relevant individual background variables (socio-economic and biographic factors) and;
- the organisational and financial problems of implementation.

There was agreement between the contractor and the authors that the feasibility study does not include presenting instruments for empirical research since the complex and time-consuming development of valid and reliable assessment tools is impossible within the available time of one year. After all, the task of highlighting possibilities for the development of measurement tools was fulfilled out in the feasibility study.

Implementation of the feasibility study

The focus of the feasibility study was an extensive review of existing national and international research literature on competencies and the corresponding measurement tools in VET, international comparative research in VET, and methodological aspects of design for international large-scale assessments. Compared to PISA, a PISA-VET survey proves more complicated regarding at least two aspects: First, vocational education and training must take into account vocational subject-specific competencies in addition to competencies of general education, and second, vocational education and training systems are so diverse (participants ages at the beginning of a programme, overall duration) that a comparison is very complex. Taking into consideration the characteristics of VET, the following aspects are discussed in the feasibility study:

With regard to the measurement of competencies:

(a) Discussion of the term competence and concepts of measurement, taking into consideration cultural contexts and scientific disciplines, and identifi-

1 A detailed discussion of different competence dimensions will be provided in *Chapter 2.*

cation of core elements of different concepts of competence and their operationalisation;
(b) proposal for a working definition of competence and perspectives for operationalisation in the context of an international comparison of VET;
(c) discussion of the relevant occupational fields and competence domains in the context of an international comparison of VET; and
(d) analysis of adequate concepts for measuring the relevant competence domains.

Regarding the micro- and macro-structural aspects and institutional factors influencing the development of competencies:

(a) Comparability of VET contents and levels of qualification (e.g., ISCED classification);
(b) discussion of institutional factors the influencing the quality in VET;
(c) determination of socio-economic and biographical factors for data collection of individual educational conditions; and
(d) agreement on the most relevant occupational fields for sample construction taking into consideration internationally comparable classifications of occupations.

The preparation, implementation, and analysis of two international workshops was the second focus of the feasibility study. The workshops, titled "PISA-VET", were conducted in Goettingen in November 2004 and April 2005 with active participation. Altogether, 22 experts from 13 different countries participated (Australia, United States, and 12 European countries), including two representatives from European research institutions for VET (CEDEFOP, European Training Foundation, Turin)[2].

In the workshop the problems and solutions were highlighted in a completely different way than would have been possible solely on the basis of a study of the existing literature. Moreover, a research-network was developed providing a basis for implementation as well as for opening access to resources in different countries at a later point in time. The participation of experts from the Federal Ministry of Economics and Labour (BMWA), the Federal Ministry of Education and Research (BMWF), and the Federal Institute for Vocational Education and Training (BIBB) provided them an opportunity to be involved in the discussion. [3]

The results of the literature review and the workshops are summarised in this report: Firstly, the state of the art is discussed critically, and secondly, perspec-

2 See attached list of participants (*Appendix A1, A2*)

3 In addition to consulting international expertise, scientific know-how at the national and international level was evaluated—Martin Baethge and Frank Achtenhagen presented the PISA-VET concept at national („Sektion Berufs- und Wirtschaftspädagogik der Deutschen Gesellschaft für Erziehungswissenschaften in Siegen", October 4, 2004), and international conferences (European Association for Research on Learning and instruction, -Learning and Professional Development- in Regensburg, October 11, 2004). In additions a workshop with scientists from the Federal Institute for Vocational Education and Training (BIBB) was very helpful.

tives for a PISA-VET are developed; the second chapter is focused on the discussion of different concepts of competence and measurement tools; the third chapter discusses the relevant micro- and macro-structural influencing factors of vocational and educational training; the fourth chapter outlines problems of sample construction and international comparative large-scale assessments. Finally, the results of the three chapters are summarised in the fifth chapter and compressed into two alternative models of research design. The report concludes with an outlook for further steps in the future.

1.2 THE CONCEPT OF VET IN THE FEASIBILITY STUDY

An international comparison of vocational education and training with the focus on competence measurement must be based on a common understanding of the goals for VET. This common understanding cannot be implicitly postulated, but must be mutually developed from a scientific research and policy point of view. On the one hand goals for VET can be based on a relatively narrow approach, which is focused on the required workplace skills. On the other hand they can be based or on a broader approach, which incorporates, in addition to subject-related competencies, those skills individuals need to participate effectively as members of a flexible, adaptable, and competitive workforce and in lifelong learning.

In accordance with ongoing scientific discussion in Germany, we have agreed upon three central goals which educational systems must address at the system level. They function as reference points for the definition of competencies, which must be developed in processes of vocational education and training. They are:

(1) The development of individual occupational adjustment from an individual user's point of view, taking self-regulation and autonomy into consideration;
(2) the safeguarding of human resources in a society, and
(3) the warranty of social share and equal opportunities.

In order to discuss, criticise and enhance these goals, they will be elaborated further:

The first goal, *individual vocational adjustment,* denotes the ability of individuals to develop relationships with their environment and create their educational pathways and life in society in a responsible and self-directed way. This refers to generic competencies like self-management skills, problem-solving skills, communication skills, and meta-cognitive skills, which emphasises that individuals are considered within the context of individual aims and efforts on the one hand, and beneficial and obstructive environmental conditions on the other hand. This perception helps to create spaces and opportunities for the development and construction of individual and occupational identities.

The second goal, *safeguarding of human resources,* subsumes every aspect of educational systems that facilitate individual's abilities to act at work and in the labour market (individual's economic user perspective) and provides workforce requirements (social demand perspective). Taken together, this refers to the suitability of vocational educational and training and the developments of occupational systems, which is, conceptualised in the concept of "Mega Trends" (Achtenhagen, Nijhof, & Raffe, 1995; Achtenhagen & Grubb, 2001; Baethge, Buß, & Lanfer, 2003). Moreover, we can distinguish between rather quantitative and qualitative aspects: From a quantitative point of view discrepancies between VET systems and occupational systems are to be avoided when possible (e.g., excessively narrow professional qualifications or over-qualifications in little-demanded or seldom available domains); from a qualitative point of view, adequate preparation for labour market requirements, which includes subject-related competencies as well as generic skills, is the main focus. The scientific discussion labels this *specialisation* versus *generalisation*, which is also a focus in the domains of lifelong learning and continuing education.

The third goal, *warranty of social share and equal opportunities,* emphasises every aspect of the relationship between vocational education and social structures. This refers to the contribution of VET to minimise dependencies between the social background and educational-, life-, and income opportunities to enhance social integration and participation of young people in processes of shaping their social and political community.

We have discussed the goals in both workshops with the result of an agreement of the goals as superior perspectives for VET. Thus, we can conclude, that a broader definition of VET, which includes, in addition to workplace-related competencies, generic competencies relating to autonomy and self-management in the labour market as well as aspects of individual biographies, lifelong learning and participation the society, is supported and accepted internationally. The participants were aware that an agreement to the suggested goals would increase the complexity of incorporated competence domains and their operationalisation (*Chapter 2*).

The need to carefully approach aspects of operationalisation and combination of different competencies, taken into consideration that competencies (subject-specific and generic) can be acquired in many different ways, was emphasised during the workshops. Moreover processes of informal learning, occurring outside institutional settings, must be taken into consideration.

2. CONCEPTS OF COMPETENCE IN VET: DEFINITIONS AND APPROACHES

One of the main research questions of the feasibility study is to determine *how* suitable measurement tools for the purpose of an international comparison of VET can be identified. The complexity of the concept of competence became apparent in the discussions with international experts in two workshops and in our review of existing literature. In agreement with many authors and workshop participants, we came to the conclusion that there is currently neither a commonly-accepted concept of competence nor an ideal framework for its measurement. Depending on the scientific orientation and cultural context there is a multitude of different approaches. Norris (1991), for example, differentiates between three main research traditions since World War II: The behaviourist tradition, which is focused on competency-based training (mainly dominant in the U.S.); the generic competence tradition, which is mainly based on management education (widely used in Britain); and the cognitive competence tradition (mostly used in linguistics). Weinert (2001) even distinguishes nine different theoretical approaches—general cognitive ability, specialised cognitive skills, a competence-performance model, a modified competence-performance model, motivated action tendencies, objective and subjective self-concepts, action competence, key competencies, and meta-competencies (ibid., pp. 6–14).

Competence measurement in the field of VET is more complex than in compulsory education. Whereas international large-scale assessments like TIMSS (Third International Mathematics and Science Study) and PISA (Programme for International Student Assessment) are limited to assessing mathematics and science performance of fourth- and eight-graders (TIMSS) or literacy, numeracy, science, and problem-solving performance of 15-year olds (PISA), a PISA-VET has to take into account individuals' performance in the workplace and the labour market as well as practical aspects (motor skills, dexterity). Moreover, international student assessment programmes like TIMSS and PISA are based upon well-grounded research traditions and internationally-validated concepts, like a world curriculum for mathematics. In comparison a PISA-VET cannot draw on comparable concepts concerning the structure and development of vocational expertise in various occupational fields. The variety and heterogeneity of occupational specialisations, even within a society, make it very difficult to reach an international agreement concerning consistent competence standards. In addition, a variety of different competence dimensions has to be considered—vocational subject-specific competencies (e.g., accounting), general subject-specific competencies (e.g., literacy and numeracy), and generic competencies (e.g., employability). In this regard, the objective of the feasibility study is to find possibilities

without simply reducing the complexity of competence dimensions. This is of importance given that, in the context of lifelong learning, one of the interesting goals is to illuminate how the relation between general and vocational subject-specific competencies changes during vocational education and training and at a later point of time in the labour market.

In the context of VET, very little research has been done to develop comparable, internationally-valid concepts of the structure and development of vocational expertise in different occupational fields and there is nothing such as to an international agreement on competence-based occupational standards or levels[4]. At the European level an ongoing progress towards the development of a standardised European Qualifications Framework (EQF) and a European Credit Transfer System for Vocational Education and Training (ECVET) can be considered a basis for further developments in this field. A PISA-VET could complement these approaches and provide solutions on the basis of well grounded scientific support.

Due to differences between national labour markets and cultures, the impact of curriculum-based objectives has to be reconsidered: The majority of most VET systems are not based upon formal qualifications (e.g., Germany) and curriculum contents function as a basis for entering the labour market. Thus, the question of how to compare curriculum contents and objectives will be one of the central aspects for the development of a PISA-VET from a political and institutional point of view taking into consideration aspects of assessment and research on learning and instruction.

In accordance with the three broad goals for VET (*Chapter 1*), a PISA-VET must focus on three competence levels and their interrelationships:

1) *General subject-specific competencies*, like reading, writing, mathematics, and problem-solving, have been discussed for a long time in the context of international comparative studies of compulsory education. They provide a basis for evaluating successful performance in different domains of life in today's society.
2) *Generic occupational competencies* are related to successful performance in the labour market. They refer to the notion of key skills, which comprise knowledge about the structures of organisations and labour markets, interacting in socially-heterogeneous groups, acting autonomously, and using tools interactively (Rychen & Salganik, 2003). The debate about generic competencies has generated a variety of different approaches and terms. In Germany, for example, they are labelled "key qualifications" (Schlüssel-

4 At a national level vocational subject-specific standards have been developed in the U.S. in the fields of economics (National Standards for Business Education, National Business Education Association, 2001; National Content Standards in Economics, National Council on Economic Education, 2000), and technology (Standards for Technological Literacy & Advanced Excellence in Technological Literacy, International Technology Education Association, 2002).

qualifikationen), whereas in the Netherlands the discussion is focused on "core competencies". However, there is no clear separation between generic, general, and occupational competencies.

3) *Vocational subject-specific competencies* refer to the collection of required individual prerequisites in order to successfully develop occupational identities and solve work-related tasks.

In the context of this report, the main focus is centred around the question how to relate the different competence levels to the fields of vocational and continuing education and training. In this regard approaches regarding and employability skills vocational subject-specific competencies at a medium-skill level are of particular interest. In this report, however, we are not aiming to provide a detailed discussion about general subject-specific competencies (PISA has provided extensive literature in this area, *Figure 5.2*). However, concepts relating to general subject-specific competencies must be adapted to the context of VET, such as vocational literacy.

For many years, national debates on generic skills[5] have been focused on different objectives. According to literature, they can be classified into three broad approaches (Smith & Comyn, 2003; Kearns, 2001): An American model which involves broad competencies in the areas of general education, personal attributes, values, learning to learn, and general workplace-related skills (e.g., communication of ideas and information, problem-solving); an Anglo-Australian model which involves a relatively-narrow and instrumental set of competencies without personal attributes and values; and a European model which incorporates broader social, as well as workplace-related issues in the context of lifelong learning.

Within the DeSeCo-project (Definition and Selection of Key Competencies), suggestions for summarising different cultural approaches have been made. According to Smith and Comyn (2003), generic skills initiatives can be grouped into three main contexts: Curriculum development, which is mainly a teaching strategy to improve schools; broad national efforts of societal renewal; and efforts for maintaining or improving the national competitiveness in an increasingly-global economy. An analysis of 12 OECD-countries revealed that certain generic skills can be found in every country. These are: Learning/lifelong learning, mother tongue literacy, communication competencies, social competencies/cooperation/teamwork, information/problem-solving/information technology-media

5 Due to the variety of terms in different countries (Australia: Key competencies; New Zealand: Essential skills; United States: Secretary's Commission on Achieving Necessary Skills (SCANS) & workplace know-how; United Kingdom: Core/common skills & employability skills; France: Transferable competencies; Germany: Key qualifications; Denmark: Process independent qualifications; Italy: Transversal competencies; Netherlands: Core competencies; Canada: Strategy for prosperity; South Africa: Critical cross field outcomes; Singapore: Critical enabling skills training (CREST); Finland: Framework for evaluating educational outcomes; Norway: Core curriculum; Switzerland: Trans-disciplinary goals) we use the term "generic skills" to refer to this type of competencies.

competencies, and numeracy/mathematical literacy. Others are present in all countries but weighted differently. These include autonomy/self-management, action orientation/taking decisions, and value education/ethic competencies. Still others are present in some countries but not in all—creativity/expression/aesthetic, foreign language/internationalisation, cultural identity, and tradition/intercultural, religion, political democracy, ecological awareness/valuing nature, and physical ability/health competencies (Smith & Comyn, 2003, p. 18). A number of empirical studies proved that generic skills are significance only within a particular occupational context, i.e. within the domain of electronic engineering (Nijhof, 1998).

The following discussion is mainly focused on illustrating the complexity of the concept of competence and possible operationalisations in the context of an international comparative study. The challenges and opportunities in relation to with regard to PISA-VET are outlined in the following chapters.

2.1 PROBLEMS WITH THE TERM "COMPETENCE"

Referring to the term *competence* only, we detected multiple and varied definitions in different countries, scientific disciplines, and various practical domains (*Figure 2.1*).

A comparison of different competence definitions shows that all of them have in common the ability to act successfully in different situations. Very often the differences are based on historical and cultural contexts resulting in terminological confusions, which was confirmed by many participants on the second workshop. Alan Brown (UK), for example, pointed out that the term competence in the UK has a more negative connotation than in Germany and that methodological competence in German has a completely different meaning from how it is used in English. Frank Coffield (UK) emphasised that from a historical point of view, the term competence is more acceptable as a scientific category in Germany than in the UK. Based on an extensive literature review, Ellström (1997) concludes that it is impossible to identify a coherent definition to accommodate the variety of ways in which the term is used (Ellström, 1997). Norris (1991) explains this dilemma as follows:

> "As tacit understandings of the word [*competence*] have been overtaken by the need to define precisely and operationalise concepts, the practical has become shrouded in theoretical confusion and the apparently simple has become profoundly complicated" (ibid., pp. 331–332).

At both workshops, the importance of an initial clarification of semantic differences and the determination of a consensual definition of competence have been the central points of discussion. A less scientific and academic definition, like for example the one from CEDEFOP (*Table 2.1*), was considered a possible solution.

Table 2.1: Examples of different competence definitions

The cognitive abilities and skills possessed by or able to be learned by individuals that enable them to solve particular problems, as well as the motivational, volitional and social readiness and capacity to utilise the solutions successfully and responsible in variable situations (Weinert, 2001, p. 27).
The proven/demonstrated and individual capacity to use know-how, skills, qualifications or knowledge in order to meet usual and changing occupational requirements (Bjornavold, 2000, p. 208).
The ability to successfully meet complex demands in a particular context through the mobilization of psychosocial prerequisites (including both cognitive and non-cognitive aspects (Rychen & Salganik, 2003, p. 43).
An ability that extends beyond the possession of knowledge and skills: it includes: i) cognitive competence involving the use of theory and concepts, as well as informal tacit knowledge gained experientially; ii) functional competence (skills or know-how), those things that a person should be able to do when they work in a given area; iii) personal competence involving knowing how to conduct oneself in a specific situation; and iv) ethical competence involving the possession of certain personal and professional values (Lisbon-to-Copenhagen-to-Maastrich Consortium Partners, 2004, p. 89).
The ability to apply knowledge, know-how, and skills in habitual and/or changing work situations (CEDEFOP, 2003, p. 39).

At the EU-level, similar issues are currently being discussed in the context of using competencies as a tool for improving transparency and flexibility. In the course of the development of an overall European Qualifications Framework (EQF), a primary objective is to develop a common set of reference levels within a two-dimensional, competence-based matrix, which allows classification of entire training programmes as well as single modules. The vertical dimension consist of levels, which are described in terms of learning outcomes and expressed in competencies in the horizontal dimension. The Maastricht Communiqué, endorsed in December 2004, underlined the need for a European Credit transfer system for VET (ECVET), which is compatible with the European Credit transfer system (ECTS) in Higher Education. The reference levels of the matrix are shared by ECTS and ECVET with the purpose to "… enable citizens to fully utilise the rich diversity of education, training and learning opportunities in Europe, to enhance communication and transparency between systems and providers, to facilitate recognition, and to promote mobility (European Commission, 2005). The structure of the—Framework is illustrated in *Table 2.2*.

Winterton, Delamare Le-Deist, and Stringfellow (2005) have presented a report with a proposal for the development of a ECVET. Based on a detailed review of existing typologies in different countries, they developed a "prototype typology of knowledge, skills and competencies (KSC)" as a common basis for all countries: "Knowledge" is captured by cognitive competencies; "Skills" are

captured by attitudes and behaviour; and "Behaviours and Attitudes" are captured by social competence (e.g., interactions with others, behaviours, attitudes, motivation, meta-cognition).[6]

Table 2.2: A possible reference level structure

Level	Cognitive Competence (Knowledge)	Functional Competence (Skills)	Social- and Meta-Competence (Behaviours and Attitudes)
Level 8			
Level 7			
Level 6			
Level 5			
Level 4			
Level 3			
Level 2			
Level 1			

Source: Winterton, Delamare Le-Deist & Stringfellow, 2005, p. 4

Regarding the number of levels, an eight-level framework with general descriptors was proposed by the QCA (Qualifications and Curriculum Authority, London), which is broadly compatible with proposals by ISCED (International Standards Classification of Education), CEDEFOP (European Centre for the Development of Vocational Training) and ISCO (International Standard Classification of Occupations). Winterton, Delamare Le-Deist, and Stringfellow (2005) proposed a level structure of increasing complexity from level 1, covering learning normally acquired during compulsory education, to level 8, covering qualifications of leading experts in highly-specialised fields with complex situations (Winterton, Delamare Le-Deist & Stringfellow, 2005, p. 42).

The current discussion is focused on an integration of sub-levels to increase flexibility and to take into account national differences between education

6 The Commission of the European Communities (2005) in their working document for a European Qualifications Framework for lifelong learning have proposed a further differentiation of learning outcomes in the horizontal dimension: "Knowledge", "skills", and "wider competences described as personal and professional outcomes". The latter is distinguished according to "autonomy and responsibility", "learning competence", "communication and social competence", and "professional and vocational competence". However, the problems of labbeling the columns in a coherent way, operationalising formal definitions on a very abstract level, and the lack of a broad basis incorporating central goals for VET, are remaining. The eight reference levels of the EQF defined by three types of learning outcomes and their progression is illustrated in Appendix A3. A general problem of classifying learning outcomes on the basis of semantic differences between verbs and adjectives, refers to an associative level which is problematic as a basis for developing measurement tools for an international large-scale of VET.

systems. For example, without sub-levels, vocational education and training qualifications in Denmark and the UK would be classified at the same level, even though the Danish training is broader and deeper than that in the UK. The problem of adequate classification also became apparent in a comparative study of the English-Welsh General National Vocational Qualifications (GNVQ) at level 3 and the German dual qualification "Industriekaufmann" (industrial clerk): On the basis of the currently-debated proposition for a new guideline COM (2002)119, which consists of a five-level classification-system of VET (EUR-LEX 1985), German VET-degrees are classified at level 2, despite Germanys calls for level 3 classification. The results of the study show, that the German qualification reaches, if not exceeds, the determined standard for the English-Welsh GNVQ Level 3. This was supported by statistically-validated empirical results (Fulst-Bleil & Ebner, 2005). In their proposal for a EQF and a ECVET, the German Central Business Association suggested a six level Framework for both, vocational and academic education, which allows a classification of "domain-specific" competence profiles (German Employers Organisation for Vocational and Further Training, 2005).

The development of the levels in the EQF was based on existing national Qualification Frameworks in different EU-member states, in particular the ones in Scottland and Ireland. The Scottish Framework comprises 12 levels, whereas the Irish consists of 10 levels (National Framework of Qualifications, 2003). Both frameworks are based on the idea of modifying their educational systems towards internationally-compatible standards. The question of how to integrate these proposals into the EQF has to be resolved in further discussions on the basis of a comparison of the respective level descriptors. Both Frameworks are designed to integrate vocational and higher education.

Winterton, Delamare Le-Deist, and Stringfellow (2005) conclude that in many EU-member states there is a tendency towards implementing learning-outcomes and competence-based VET-systems, which demonstrates growing interest in multi-dimensional frameworks. With regard to the three competence dimensions, *knowledge*, *skills*, *behaviours and attitudes*, there is a broad concordance. In some cases discrepancies are based on terminology but in other cases they are "fundamental and conceptual" (ibid., p. 44). Therefore, the challenge is "to develop a consistent and coherent typology of KSC in a context where even within countries, there is apparent diversity in the approaches to competence" (ibid., p. 40). In the UK setting in the context of NVQs, Oates (2004) pointed out five misleading assumptions of outcome-based qualifications and the way in which language functions to describe competence:

- First, a precise language can be established for the accurate communication of competencies and attainment "…approaching the precise communication level of science…";
- second, outcome-based qualifications possess high levels of validity, credibility, and utility since they are based on the content of work processes;

- third, for the purpose of modern national VET systems, it is enough to state desirable outcomes, since these provide the basis for recognition (that someone is competent), and for instruction (these are the outcomes to which the instruction should lead);
- fourth, significant benefits accrue from assessment and certification being independent of the mode, location, and duration of training, particularly when using a separat facility to accredit experienced workers and apprentices who have qualified through initial VET programmes; and
- finally, unitised qualifications allow more accurate and flexible arrangements for recognising competence since different combinations of units can be used to construct varying qualifications (Oates, 2004, p. 56).

The development of NVQs was based on these assumptions as a result of observations in very limited, specific circumstances, based on a restricted concept of human beings, excluding aspects of individual personalities.

Regarding the concept of competence Winterton, Delamare Le-Deist, and Stringfellow (2005) are facing the same problems we mentioned earlier—there are multiple and varied definitions in different countries, scientific disciplines, and practical domains. However, to use concepts of competence productively, it is necessary to have a well-elaborated and theoretically well-founded basis. Moreover, formal definitions at a very abstract level do not provide usable operationalisations (e.g., the definition provided by the Ministry of Education, Germany, 1996). Since Winterton, Delamare Le-Deist, and Stringfellow (2005) do not provide clear goals for vocational and continuing education and training, they encounter problems for classifying the variety of competence dimensions; very often they apply contradictory approaches for evaluating competencies. Particularly problematic appears to be the column, "social and meta-competence" (*Tables 2.2, 2.3*), where aspects of social competence (interactions with others) are mixed up with general abilities (behaviours, attitudes, motivation and meta-cognition), which have an impact on every competence dimension. To overcome this shortcoming we will classify general abilities into the competence dimension of self-competence, which incorporates the remaining dimensions of cognitive-, functional-, and social-competence.

The authors' problems of labelling the columns in a coherent way, indicates their lack of theoretical consistency and underlying goals for VET. Their repetitive emphasis on existing agreements at the European level does not hide this fact. We consider it fundamental to take a stringent (from an economic and educational view) perspective of competencies, which is based on a detailed examination of existing international operationalisations of knowledge dimensions and personal dispositions, as an underlying concept.

In the following paragraph, Winterton, Delamare Le-Deist, and Stringfellow's (2005) proposal will be compared with different German concepts for the definition and measurement of competence. The main problems of Winterton, Delamere Le-Deist, and Stringfellow's paper are the lack of explicit explanations

for the goals of VET, the acceptance of individuals adaptation to the workplace, and an insufficient explanation of competence dimensions.

Table 2.3: Approaches to KSC in EU member states

Country	Cognitive Competence (Knowledge)	Functional Competence (Skills)	Social and Meta-Competence (Behaviours & Attitudes)
Austria	Sachkompetenz		Personalkompetenz
		Sozialkompetenz	
Finland	Knowledge	Working methods	Core competencies
	Safety		
France	Savoir	Savoir-faire	Savoir-être
Germany	Fachkompetenz		Personalkompetenz
		Methodenkompetenz	
	Sachkompetenz	Sozialkompetenz	
Hungary	Knowledge	Application of knowledge	Professional attitudes & behaviours
Ireland	Knowledge	Know-how and skill	Competence Role, Insight, & Learning
	Competence-context		
Netherlands	Profession-specific methods		Establishing & maintaining contacts, cooperation, teamwork
			Development of an individual team
	Professional competencies for functioning in the context of work organisations		
Portugal	Competências cognitivas	Competências funcionais	Competências sociais
England & Wales	Underpinning knowledge	Functional competence	Social competence
Scotland	Knowledge & understanding	Practice (applied knowledge)	Autonomy, accountability, working with others
	Generic cognitive skills	Communication, Numeracy, IT skills	

Source: Winterton, Delamare-Le Deist & Stringfellow, 2005, pp. 45–46

In Germany, Erpenbeck and von Rosenstiel (2003a; 2003b) have published their "Handbook of competence measurement," focussing on the "identification, comprehension, and assessment of competencies in occupational, educational, and

psychological practice". From a psychological point of view, this subsumes the adjustment of:

- Subject-centred personnel diagnostic: "Which personal characteristics does a person have?";
- action-oriented diagnostic of behaviour: "How is a person performing?"; and
- diagnostic of outcomes/results: "What are the persons' individual achievements?" (ibid., 2003b, p. XXIX).

Outcomes are distinguished according to convergent, requirement-oriented and divergent, self-organised actions and work-related situations. The former are directly focused on the compliance of external requirements and objectives, whereas the latter can be characterised as creative and partly or fully open to final outcomes and objectives (p. XXVIII). These categories are classified according to:

Qualifications, which are "clearly defined areas of knowledge, skills, and abilities, people need to possess to be able to perform in a certain occupation". They are characterised as *action-orientated* and *convergent, requirement-orientated*. Typically, they can be unambiguously operationalised and accredited (p. XXIX).

Abilities, which are action-oriented "solid systems of general psycho-physiological action processes"; they can refer to *convergent, requirement-orientated* and *divergent, self-organised* action situations.

Competencies, which indicate "…dispositions of self-organized, psycho-physiological actions". Dispositions are considered developed, internal preconditions for occupational adjustments; they do not only consist of individual talents but also of outcomes of developmental processes; they are *action-oriented* and primarily focused on *divergent, self-organised* action-contexts (p. XXIX).

For aspects of measurement this distinction entails the following: Assessment of psychological characteristics "provides insights into existing disposals for self-organization and, therefore, competence." However, the actual existence of competence can be "assessed only in the context of action" (p. XXIX).

In this context, the importance and impact of knowledge inventories, as well as individual collections of skills and abilities, has to be accentuated. Within the competence-based OECD project DeScCo, Weinert (2002) proposes the following definition for competencies: "The cognitive abilities and skills possessed by or able to be learned by individuals that enable them to solve particular problems, as well as the motivational, volitional and social readiness and capacity to use the solutions successfully and responsibly in variable situations" (ibid., pp. 27-28).

Weinert (2002) mentions eight competencies according to the frequency of occurrence: Mastery of mother tongue (oral and written); mathematical knowledge; reading competence; mastery of at least one foreign language; media competence; learning to learn; social competence; and divergent critical thinking. This listing corresponds to the occidental curriculum (Dolch, 1965), which is mainly focused on the "gold standard" of education whithin the educational track from compulsory to university education; references to vocational education and

training are not included (Achtenhagen, 1998). However, an inclusion of occupational subject-related competencies seems to be possible. Weinert's elaborations on competencies are important because interdependencies of competence dimensions are emphasised and aspects of motivation, attitudes, and metacognition are incorporated.

In the context of NVQs in England and Wales, Oates (2004) complains that there has been a tendency to push VET-systems and compulsory education too rapidly from a "formation" model to a "competence" model (Oates, 2004, p. 59).

Table 2.4: Comparison of the "formation" model and the "competence" model

Formation	Competence
formative assessment to provide feedback for learners in order for them to enhance their performance	**summative assessment** to provide a transparent statement of what someone can do for certification, selection, etc.
open, unstructured information on performance which reflects the specific composition and meaning of achievement, motivation, etc. of an individual	**tightly-structured, centralised information on performance** which provides a common language for describing competence
emphasis on inputs relating to effective, structured use of learning processes and settings, with assumptions that processes of learning affect the nature of emerging competence	**emphasis on outcomes** with an emphasis on inferring competence from performance, on leaving open the mode, duration, and location of learning

Source: Oates, 2004, p. 59

A comparison of the two models clearly shows that too narrowly-defined concept of competence do not incorporate the fundamental goals of VET outlined earlier. The pejorative use of the term *competence* in English was also confirmed by the British workshop participants (e.g., Alan Brown, Frank Coffield). Therefore, we decided to draw upon a broader concept, founded in pedagogic anthropology, which relates individual attitudes and motives into the domain of self-competence.

In the context of national educational standards in Germany (Klieme et al., 2004), competencies are considered school related; they describe

> "abilities of the kind intended and assumed by the term *Bildung:* abilities which are acquired, not inherent, which are experienced in and through particular aspects of societal reality, and which are suited to shaping this reality. In addition, these are abilities that are capable of being cultivated over a lifetime, enhanced and refined in such a way that internal gradations can be measured, such as from basic to more advanced general education. They are, however, also abilities that enable a process of self-directed learning, since the aim is abilities that are acquired not only in conjunction with tasks and processes, but can be detached from the original situation, meet the challenges of the future, and are open to problems of all kinds" (ibid., pp. 59–60).

In the standards orientation towards vocational education is mentioned, but a link to the contents of vocational education, which are a constituent of compulsory education in Germany, is not provided. Improving students' performance level when completing lower secondary education would certainly have a positive impact for learning and instruction in VET. After all, a PISA-VET has to be based on the same underlying concepts of individual development as in general education.

The three depicted German approaches range from a primarily psychological point of view (Erpenbeck & von Rosenstiel, 2003a; 2003b) to more content-related aspects in the context of output-related standards at the end of lower secondary education for the subjects of German, mathematics, and first foreign language (Klieme et al., 2004). However, all three approaches refer to Weinerts definition of competencies.

Contrary to Winterton, Delamare Le-Deist, and Stringfellow's proposal, the specification of knowledge is more prominent in the German approaches. Whereas Erpenbeck and von Rosenstiel (2003a; 2003b) incorporate this aspect very little, the commission of experts (Klieme, et al., 2004) consider aspects relating to different knowledge dimensions to a greater extent. Winterton, Delamare-LeDeist, and Stringfellow (2005) distinguish between different content domains, but with little reference to occupational fields and industries. For VET, we consider it necessary to focus on output-oriented measurements of competencies for different occupational domains.

2.2 PROBLEMS OF OPERATIONALISATION AND MEASUREMENT OF COMPETENCIES

Looking at the constituents of different competence definitions, certain elements (regardless of linguistic differences) can be found internationally in the majority of proposals: The application/implementation of an individual collection of (performance) prerequisites (knowledge, skills, aptitudes, know-how, strategies, routines, emotions, attitudes, motivation, willingness, etc.) to solve certain problems/complete particular tasks or/handle certain situations (successfully) in various contexts.

At a different level of analysis, emerging layers of complexity and confusion about competencies are demonstrated by levels of occupational specificity (Oates, 2004). The variety of invisible individual conditions, which can be demonstrated in different situations, were discussed in the workshops. *Table 2.5* illustrates examples of different internal conditions as a basis for operationalisation and measurement.

Table 2.5: Possibilities for operationalisation and measurement of competencies in VET

Internal conditions/individual abilities (person)	**External performance/behaviour/action (situation)**
Declarative knowledge	
Specific procedures	
General procedures	
Strategies and skills	
Values	
Attitudes	Performance
Motivation	Solving tasks in different situations
Self-efficacy	
Locus of control	
Self-control	
Anxiety	
Big Five*	
not directly observable	observable

* Measurement model comprising five personality dimensions: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

One of the demanding challenges of an international comparative study, based on the three superior goals outlined in *Chapter 1*, is to specify selected competence domains in accordance with these goals. Particularely in the second workshop it was emphasised that making reasonable selections in relation to the three goals is more important than creating lists of elements of competence. Therefore, we can conclude that a broad approach, which exceeds pragmatic methods for personnel selection, is the favoured objective for an international large-scale assessment of VET.

Based on the distinction between individual conditions, which are not directly observable, and behaviour that is directly observable in various situations, two approaches can be identified: Competencies as a potential for application in different situations refers to internal conditions (e.g., knowledge, skills, motivation, values, meta-cognition), which are not observable to outsiders. On the other hand, competent behaviour in a specific situation can be observed by third parties. Consequently, operationalisation and measurement of competencies can refer to the

- determination of *external tasks*, denoting internationally representative situations and requirements for an occupation or occupational function; or

- determination of *internal conditions*, such as knowledge, skills, motivation, values, or meta-cognition, characterising the development of vocational expertise in a certain occupation/occupational field.

Both alternatives are discussed with regard to their impact, challenges, and opportunities for a PISA-VET. In this regard, Lang-von Wins (2004) points out that theoretical standards on a high level of abstraction are opposed to more practically-oriented pragmatic aspects.

2.2.1 Identification of competencies on the basis of external activities

This approach requires a classification of typical, internationally-comparable occupational activities and tasks in different occupational fields. Occupational situations and requirements could be determined on the basis of existing formal descriptions in different countries in terms of their communalities. For example, NVQs, which are competence-based descriptions of required outcomes. Following this approach Preiß (2004) compared the structures of competence descriptions in the occupational field of business/administration in Germany and England/Wales. In a similar vein, Fulst-Bleil and Ebner (2005) conducted a comparative study in Germany and Wales. These approaches might be a way for identifying competencies on the basis of external tasks. However, there are considerable problems involved: Firstly, there are linguistic differences between certain terms (how do we know that descriptions of occupational tasks in different countries really have the same meaning?); secondly, it is possible that educational contents are not relevant in the labour market; thirdly, different educational tracks (initial and continuing VET) can be involved; and finally, a broad underlying concept of VET is missing.

Problems of comparability are worsening as more countries become involved in the study. Assuming that there are internationally-valid common requirements for certain occupations (e.g., car mechatronic), it must be assured that national characteristics are adequately incorporated. For example, in the field of car mechatronics, although there are many commonalities at an international level, there are a number of specific characteristics at a national or even regional level.

Identifying competencies on the basis of external tasks is less time consuming than the development of concepts for internal conditions (e.g., technical literacy). The time-saving advantage was emphasised by some of the participants during the second workshop. It was suggested to first take typical, internationally-comparable tasks, which will then serve as a basis for empirical examination. In this regard, Tim Oates (UK) points out the danger of ignoring the complexity of competence and advises to infer competencies from observation of individual performance in different situations to avoid measurement errors.

One approach for identifying external tasks refers to an analysis of work and work requirements, which is distinct from occupational profiles. However, this

approach is relatively time consuming and work intensive; typically it is based on a three-step procedure (Lang-von Wins, 2004):

> First, analysis of task- and performance conditions of a working activity; second, identification and description of task dimensions; and third, analysis of the working activity on the basis of these dimensions, for identifying achievement and qualification- differentiating variables (ibid., p. 591).

This approach refers to working samples on the basis of controlled imitations of critical success situations with the purpose of observing and assessing individuals performance in work contexts (Lang-von Wins, 2004). However, this approach requires a high amount of effort for the construction of work samples. Therefore, a combination of simulation-based and biographical methods is considered a reasonable alternative (Lang-von Wins, 2004). Examples are provided by the Association for Research in Professional Development (ABWF) (2005), and the PISA report for Germany (Wirth & Klieme, 2003).

The most elaborated and concrete version for identifying competencies on the basis of external activities, refers to aggregated work activities of work functions and was proposed by Wim Nijhof (Netherlands) at the workshops. The underlying instrument is based on a labour market-related occupational database (O*NET) from the U.S. (Jeanneret et al., 2002). One of the central elements of the O*NET are Generalized Work Activities (GWA), which are "…aggregations of similar job activities/behaviours that underlie the accomplishment of major work functions" (Jeanneret et al., 2002, p. 106). A taxonomy of 42 GWAs provides the basis for assessing work activity requirements for the majority of occupational fields at an international level. The database follows the idea that work behaviour is not necessarily linked to specific tasks and techniques and, therefore, can be aggregated at a higher level of abstraction. However, not all GWAs are located at the same level of abstraction; some are very generic and others are more specific, but not too specific that they can be linked to only one occupation.

The classification of Generalized Work Activities has strong theoretical and empirical foundations. In a number of empirical studies, the underlying work structure was investigated on the basis of individual behaviour. Jeanneret (1969) investigated the hypothesis that "there is some structure underlying the domain of human work, and that this structure can be identified in terms of one or more sets of job dimensions" (ibid. 1969, p. viii). The criteria for determining whether a construct would qualify as a GWA include: i) Being broad in scope and having applicability to a wide range of occupations; ii) being based on job-analytic research; and iii) being characteristic on the underlying structure of work (Jeanneret, et al., 2002)[7].

Toolsema (2003) adopted the GWA concept in his research for the identification of competencies in Higher Education. On the basis of a comparison of different studies, he derived six competence categories—social, participative, cognitive, physical/technical, learning, and employability, and associated them

7 A detailed illustration of the GWA concept is provided in Appendix B of this report.

with GWAs based on the assumption that if a person performs a certain activity, he/she possesses the required competencies for that activity. Activities are perceived as competence indicators at a higher level of abstraction, indicating the purpose of the corresponding competencies. The instrument is suitable for measuring cognitive, affective, and meta-cognitive aspects of competencies at different levels of abstraction and, therefore, provides a basis for an international comparison of VET.

The advantage of the approach is that competencies for different occupational fields can be derived on an aggregated level on the basis of working activities, which is consistent with a more pragmatic line of reasoning: Individuals do not necessarily need all their competencies in the work place and therefore, measuring the entire collection of the corresponding competencies would not be necessary. With regard to reducing the level of complexity in a PISA-VET some participants recommended to concentrate on those competencies that are really necessary in the workplace (this aspect was mainly emphasised by Wim Nijhof). However, there are two objections: Firstly, there is a risk of taking a too narrowly-defined concept of competence which is based on work performance, and secondly, from a research methodical point of view, there is a circulation problem—how can competencies in different workplaces be determined without previously analysing work and work requirements? Moreover, it can be assumed that this approach demands time-consuming and work-intensive observations of direct performance. In this regard, indirect measurement based on simulations and tests, complemented with biographic questionnaires, might be a reasonable alternative.

For the purpose of a PISA-VET, the question whether a restriction to measurements of individuals' performance or an investment in a more time-consuming development of concepts for identifying internal conditions is the more reasonable alternative, must be examined carefully.

2.2.2 Identification of competencies on the bases of internal conditions

Measurement of competencies on the basis of individual potentials/internal conditions in large-scale assessments is based on testing methods. This approach requires conceptualisations comparable to the ones in PISA (e.g., literacy, numeracy) relating to internationally-valid models of the structure and development of vocational knowledge and skills in different occupational domains. Existing concepts need to be examined regarding their suitability for an assessment of vocational expertise and with regard to necessary modifications or further developments. Moreover, an agreement concerning the relevant competence dimensions and level structure must be made. In this regard, the "Standards for Technological Literacy," developed by the International Technology Association (ITEA) are an example for a national concept. The standards for technological literacy

are designed to determine how well students are attaining technological literacy in grade K-12; they specify what every student should know and be able to do in order to be technologically literate. There are a total of 20 standards, which fall into two types: First, what students should know and understand about the technology; and second what they should be able to do. The first (cognitive) type sets out basic knowledge about technology, whereas the second type (process) describes the abilities that students should have.

The standards are organised into five major categories: The nature of technology; technology and society; design; abilities for a technological world; and the designed world. The format of each standard follows this structure: First, the standard is expressed in sentence form; second, a narrative explains the intent of each standard; third, grade-level material is presented for grades K-2, 3-5, 6-8, and 9-12; and fourth, a narrative is presented that explains the standard for each grade level and provides suggestions of how the standards can be implemented in the laboratory-classroom. For each grade-level, benchmarks that detail the particular knowledge and skills, vignettes with examples of laboratory-classroom experiences, and illustrations of how the standards can be put into practice, are provided (International Technology Education Association, 2002).

In the field of Business Administration, comparable concepts have been developed in the United States: The National Council on Economic Education (NCEE) has produced a set of curriculum standards based on the essential principals of economics, titled "Voluntary National Content Standards in Economics". Each of the 20 standards, developed by a panel of economists and economic educators, includes a rationale for its inclusion, benchmarks indicating attainment levels for students in grades 4, 8, and 12, samples of what students can do to enhance or demonstrate their understanding in economics, and correlations of existing publications to the standards.

The standards consist of generic skills and cognitive competences. The key skills students must develop in economics include an ability to identify economic problems, alternatives, benefits, and costs; analyse the incentives at work in an economic situation; examine the consequences of changes in economic conditions and public policies; collect and organise economic evidence; and compare benefits with costs. Additionally, students should have gained several kinds of economic knowledge by the time they have finished the twelfth grade:

- First, they should understand basic economic concepts and be able to reason logically about key economic issues that affect their lives as workers, consumers, and citizens, to avoid errors that are common among persons who do not understand economics;
- second, they should know some pertinent facts about the American economy, including its size and the current rates of unemployment, inflation, and interest; and
- third, they should understand that economists hold differing views on some economic issues. This is especially true for topics such as the appropriate size

of government in a market economy, how and when the Federal Government should try to fight unemployment and inflation, and how and when the Federal Government should try to promote economic growth.

The essential propositions of economics are identified in the 20 content standards that follow: Scarcity, marginal cost/benefit; allocation of goods and services; the role of incentives; gain from trade; specialisation and trade; market-, price-, and quantity determination; the role of pricing in the market system; the role of competition; the role of economic institutions; the role of money; the role of interest rates; the role of resources in determining income; profit and the entrepreneur; growth; the role of government; using cost/benefit analysis to evaluate government programmes; macro economy, income, employment, and prices; unemployment and inflation; and monetary and fiscal policy.

These examples illustrate the complexity of developing internationally-valid competence models and measurement tools in different occupational fields.

As previously mentioned, the complexity of PISA-VET increases due to functional aspects of skills and the impact of individual behaviours in the labour market. With regard to successful performance in the labour market and lifelong learning, modifications of knowledge- and skill-structures during the development of expertise are of high importance for professional competencies and need to be taken into consideration. In this regard, aspects of informal learning, attitudes, and motives play a substantial role (*Chapter* 2.5).

2.3 COMPETENCIES BETWEEN NATIONAL SPECIFICS AND INTERNATIONAL COMPARABILITY

Another relevant dimension of competencies in an international comparative study refers to the level of aggregation. The complexity and multi-faceted nature of different levels is illustrated in the following sections on the basis of a literature review and contributions from the workshops.

The discussions in the two workshops clearly showed the difficulty in agreeing upon a level of aggregation; the more specific the definition of competencies/working tasks, the less comparability exists at an international level. After all, the participants generally agreed that an adequate approach has to differentiate between occupational fields and, at the same time, ensure international comparability; i.e., the successful performance of a salesperson mainly depends of his abilities to advertise his products, whereas the successful performance of a kindergarten teacher probably rather depends on his assertiveness skills (Kanning, 2003). Since there are additional national specifications and facets of competencies and situations, an aggregated level above the national level should be determined.

On the basis of a review of different techniques used in work analysis, Oates (2004) suggested the following distinctions in the scope or level of aggregation of descriptions of competence:

- *Generic* relates to skill components, which are not expressed in the form of activities specific to a particular setting (e.g., key skills in communication).
- *Occupational* relates to descriptions at an occupational level relating to activities with a high level of generality. The descriptions intend to apply to a wide range of settings and specific methods for completing activities (e.g., a general description of fault diagnosis in aircraft hydraulic systems).
- *Task-specific but independent of specific jobs* relates to descriptions which provide the task details (e.g., constructing lead roofing), but which do not describe the way that the task might be organised within a specific work system. This level of description is common in the development of national standards (qualifications and quality standards). Within this level, there is a hierarchy of sub-levels; for example, a standard can relate to *inert gas welding* or deal with a lower level of aggregation e.g., i) *argon welding*; ii) *tungsten inert gas welding*; iii) *oxy-acetylene welding,* etc.
- *Job-specific, enterprise-specific* relates to descriptions of the way a task is undertaken in a specific work system. Falling into this category are jobs descriptions, work analysis processes (for the purpose of pay regulation reward systems, and management control), etc.
- *Person-specific* relates to the way the task or the job is undertaken by a specific work system. This is important for training, assessment, and evaluation purposes. Undertaken by the workers or learners themselves, it can also form her/his contributions to appraisal systems and, assessment, as well as assist in preparing applications for jobs and for training/education programmes (Oates, 2004).

The multi-faceted aspect of possible levels of competence descriptions illustrates the importance of an initial fundamental clarification of the intended goals of VET. For example, using *job-specific, enterprise-specific* description of competence would be too specific to ensure international comparability of different occupational branches. Thus, for the purpose of an international large-scale assessment of competencies in VET, a manageable and informative level of comparability must be determined.

Within the different levels of competence, a variety of approaches based on external performance, as well as internal conditions, can be found in literature. For example, the job-competency approach by Spencer and Spencer (1993) is a very specific, performance-based approach from the U.S.: The person in the job is analysed without making assumptions about the characteristics needed to perform well in the job. With open-ended behavioural event interviews, it is determined which characteristics are associated with job success. Thus, in this approach competencies are associated with "underlying characteristics of people" that are causally related to effective or superior performance in a job or situation

… generalising across situations, and enduring for a reasonably long period of time" (ibid., p. 9). A similar approach for identifying competencies in the work place is the project "Joint-Service Job Performance Measurement Standards (JMP)", which is an instrument for assessing job performance of military personnel in the U.S. (Wigdor & Green, 1991). However, this type of approaches are too specific for an international comparison, since a determination of standards at a national level makes comparability at an international level more difficult.

Contrary to very specific competence descriptions, cross-occupational generic skills at a higher level of abstraction are often debated in the discourse of *generic skills*. Very often they are also labelled *enabling skills*, *key competencies*, *essential skills* or *necessary skills* (Nijhof 1998; Murray, Clermont & Binkley, 2004). Typically, generic approaches are concerned with the development of transferable general skills and competencies that are required for most jobs and are distinctive from more workplace-specific technical knowledge and skills. The proliferation of different concepts was a response to an inflation of the meanings of formally specific and narrowly-defined competencies. In Germany, the debate has been focused on *key qualifications* (Schlüsselqualifikationen; Mertens, 1974), which are linked to the concept of vocational action competence (Handlungskompetenz). The concept of key qualifications has generated about 650 different versions, which is one of the reasons why the concept of competence in VET is discussed at different levels in Germany at the moment.

At an international level, different education and labour market institutions have conducted a number of projects for the identification of employability skills. Most of these studies initially implemented a broad definition, followed by an observation of workers, supervisors, or experts, to determine common competencies for particular work activities. For example, the SCANS (Secretary's Commission on Achieving Necessary Skills) project differentiated between two categories: *Foundation skills,* which include basic academic skills (reading, writing, arithmetic), thinking skills (reasoning, problem-solving), personal qualities (responsibility, self-esteem); and *work competencies* which comprise the ability to use resources, interpersonal skills, information systems, and technology (Secretarys Commission on Achieving Necessary Skills, 1991).

Cross-occupational generic approaches have been questioned by several authors for various reasons. For example Weinert (2002) argues that there are skills/ competencies that are useful in many learning and working situations, but they are no substitute for content-specific knowledge. Others argue that individual competencies cannot be viewed as independent of a specific context (Sandberg, 2000). The constructivist approach proposes that the extent to which a definition proves to be adequate depends on the context in which it is used (Stoof, 2005). Based on the results of their study, Stasz and colleagues (1996) argue that "an analysis of skill requirements that ignores work context draws attention away from workplace characteristics and possible shortcomings in firm behaviour that affect skill utilization and performance" (ibid., p. 5). They identify variations and

similarities in "generic" competencies that occur in work practice, which confirms that individual characteristics needed for a particular job "can be understood only within a particular working context, from the perspective of individuals in the social setting" (ibid., p. 9). After all, cross-occupational competencies are not always transferable and associated with specific vocational training processes. In this report, we distinguish between vocational subject-related and cross-occupational competencies for analytical reasons.

2.4 WORKING DEFINITION FOR A CONCEPT OF COMPETENCE IN VET

To handle the multi-faceted nature of different competence dimensions and -levels, we go back to a pedagogical -anthropological approach of Roth (1971) and Reetz (1999) as a basis for developing a research model for PISA-VET. According to Roth (1971) and Reetz (1999), four competence domains can be distinguished:

- Self-competence,
- cognitive competence,
- functional competence, and
- social competence.

This broad concept is currently dominating the German discussion on competencies in VET. The different elements and the structure of the model will be illustrated later in this paragraph. However, it has to be pointed out that the four domains cannot be explicitly separated. For example, knowledge is not limited to the domain of cognitive competence and different kinds of knowledge are applied in functional and social contexts as well (knowledge is used for performing theoretical occupation-related tasks; utilising tools, equipment, and technical resources; and in interactions with others). The acquisition and application of knowledge is governed by motives and attitudes; the process of reflecting upon applied knowledge in a comprehensive way relates to "meta-cognition". Thus, adopting a two-dimensional approach to account for interrelations between the domains, ensures that the conceptualisation is coherent and provides a strong, consistent, and defensible basis for measurement and comparison of VET. The schema (*Table 2.6*) incorporates John Stevensons suggestions from the workshops. It takes into account the different competence domains, as well as their interrelations in different occupational situations.

In the framework, a broad concept of competence is combined with possible actions and activities as a basis for research and interventions. In future large-scale assessments, the approach considers apprentices, workers, and employees not only as a medium or executor of occupational activities, but as individuals with personal identities. Therefore, self-competence is determined by the characteristics of cognitive, functional, and social competence, as well as attitudes, motives, meta-cognition, and occupational-action orientation (Beck, 1995).

These aspects are particularly relevant in complex occupational tasks, such as in the occupational field of "symbol analysts" (Symbolanalytiker; Baethge, Buss & Lanfer, 2003).

Table 2.6: Matrix of interdependencies between self-, cognitive-, functional-, and social competence in different contextualised areas of performance

Individual Capacities (accessed and interpreted in different contexts)	Competence Domains (different contextualised areas of performance)		
	Self-Competence		
	Cognitive Competence	**Functional Competence**	**Social Competence**
	(theoretical/ analytical requirements) "applying concepts"	(technical/practical/ functional requirements) "using tools, equipment, and technical resources"	(interpersonal requirements) "interacting with others"
Attitudes Values Perceptions			
Incentives Motivation			
Meta-Cognitive Strategies			
Declarative Knowledge			
Procedural Knowledge			
Strategic Knowledge			

The introduction of SAP has eliminated the need for developing an understanding of the underlying contents. To a great extent accounting activities are reduced to a mere operation of screen masks with the result that VET in the field of commerce and administration is no longer comprehensible as a whole. Therefore, it has become difficult for apprentices and the majority of commercial employees to understand business and production processes. From an *economic* point of view, individuals hardly have a chance to critically assess business development and from an *educational* point of view, which is not supporting the development of individual identities. This also refers to the development of human resources; the underlying knowledge dimensions (declarative, procedural, and strategic knowledge), as well as motivation, interest, and meta-cognition are not

addressed. Thus, the development of competencies to motivate further learning in professional and private contexts is excluded.

Taking these aspects into consideration, we suggest to consider competencies mainly from a knowledge-developmental point of view, since vocational expertise in a knowledge domain/occupational field is a crucial precondition for the development of transferable skills, achievement motivation, interest, and metacognition. This perception corresponds with today's tools of management education (e.g., organisational learning, knowledge management), which are based on the assumption of active and creative employees (Tuomi, 1999).

The question of whether or not this managerial point of view, which "emphasizes the importance of learning processes, particularly in blooming companies" resulting in—"a work habit which is characterized by an excellent quality and individual focus on holistic working processes" might be called a "coincidence of economical and educational rationality" (Achtenhagen, 1990, p. VII) remains open at this point. However, this assumption could be regarded a reasonable conclusion for research projects in this area (Heid & Harteis, 2005).

In the following section, each of the fow competence domains will be discussed to illustrate aspects of measurement methods.

Self-competence

Self-competence subsumes every aspect of personal development in a narrow sense—the development of cognition, emotion, motivation and ethics (Achtenhagen, 1996, p. 27). According to Reetz (1999), self-competence refers to "the ability of morally self-determined behaviour, which includes a positive self-perception and the development of moral reasoning skills" (p. 42). In the context of PISA-VET, personal characteristics which impact vocational performance (attitudes, values, perceptions, incentives, motivations, and meta-cognitive strategies) have to be measured.

Descriptions of self-competence are impacted by assumptions about individual developments over the lifespan relating to motivation and metacognition. However, in this regard, there are immense research gaps and literature on lifelong learning systematically neglects this fact (Achtenhagen & Lempert, 2000). Achievement motivation and the corresponding behaviours develop at an age of about 11 to 12 years and meta-cognition around 14 to 15 years. The impact of motivation and meta-cognitive structures is underestimated in literature which, might explain why many investments in continuing education have failed to result in the expected outcomes. Company representatives have admitted that about 50 percent of their expenses for continuing education are wasted without resulting in any notable effect.

Empirical studies on the development of motivation during VET show that, at least in the field of commerce and administration in Germany, a decrease in the level of interest in learning in schools *and* workplaces occurs. Towards the end of

educational programmes, school- related data remain low, whereas workplace-related data increase (Prenzel, Kramer & Drechsel, 2001).

After all, we can conclude that it is worthwhile to obtain information about motivational and meta-cognitive structures of apprentices. In this regard, applying multiple measuring points and retrospective methods will be particularly helpful to illuminate individual competence developments. For the measurement of motivation, we need a combination of items consisting of habitual and situational components. For example, items related to motivational preferences (Deci & Ryan, 1985), self-related cognition (Little, Oettingen, & Baltes, 1995; Bandura, 1989) or effort and perseverance (Kuhl, 1984; 1998). A measurement of the state *and* trait components of motivation is the favoured option (e.g., Winther, 2005). With regard to meta-cognition, there are concepts, which consider problems relatively independently from contents. Well-established concepts for the description and evaluation of meta-cognitive structures can be found in the context of bachelor and master degree qualifications (e.g., questionnaire for the use of the Motivated Strategies for Learning (MSLQ); Pintrich, Smith, Garcia, & McKeachie, 1991), and in the area of VET (e.g., Bendorf, 2005). Diverse vocational tasks and requirements are the basis for drawing conclusions about learning- and processing strategies and factors of success.

Measurements of motivation and meta-cognition are essential for gaining insights into integrated competence developments and the corresponding cognitive and procedural structures. The PISA study showed that it is meaningful and necessary to measure aspects of attitudes, values, perceptions, and motivation. Therefore, comparable instruments should be applied in a PISA-VET. Provided that the short version of the Big Five personality dimensions (Rammstedt, 2004) proves sufficiently reliable, it might be a possible measurement tool for a PISA-VET. A high level of validity is particularly important taking into consideration the expected heterogenity of the sample.

Cognitive competence

Recent empirical studies showed that a differentiation of knowledge according to declarative, procedural, and strategic knowledge, is plausible and feasible (e.g., Bransford, Brown, & Cocking, 2000)[8]. In the domain of *declarative knowledge* (factual knowledge), new concepts, which are related to work requirements, have to be developed. Particularly with regard to the increasing complexity and networked structures in work contexts, declarative knowledge cannot be developed in linear, abstract, and fragmented ways.

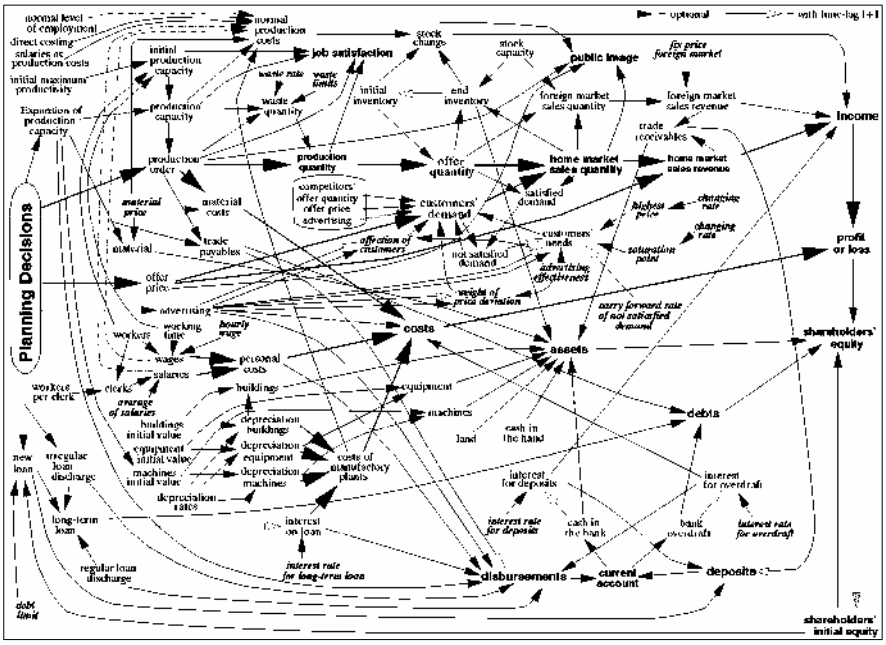
8 Shavelson, Ruiz-Primo, and Wiley (2005) suggest a fourth category of *schematic knowledge* (knowing why). Detailed analysis in the context of science instructions shows, however, that there are immense overlapping areas in the three remaining knowledge categories. With regard to the level of performance, they distinguish between *extent* (how much?), *structure* (how is it organized?), and *others* (how efficient? how precise? how automatic?).

Figure 2.1 the complex and networked structures of organisational processes: A concept for initial instruction in the field of business and administration in Germany, based on the idea of "networked" thinking in management education (Gomez & Probst, 1987). For a PISA-VET in the field of business and administration, test items must be constructed on the basis of understanding networked structures. For example, Shavelson, Ruiz-Primo, and Wiley (2005) have generated proposals for corresponding test items and methods (Achtenhagen & Baumert, 1999). Effective actions at work regarding organisational aspects and accounting procedures require the corresponding knowledge structures. In this regard, simulations might be a reasonable solution. The depicted knowledge net provides a basis for the development of virtual firm simulations.

The domain of *procedural knowledge* subsumes the application of knowledge, that is, how to operate with knowledge nets and the corresponding elements.

Finally, the development of *strategic knowledge* is considered one of the central goals for VET. The focus is on behavioural conditions and problem solving, which refers to applying deep declarative and procedural knowledge, taking into account possible primary effects and unintended side effects. Thus, decisions are made with a full awareness of possible negative consequences.

Figure 2.1: Knowledge structure for initial instruction in the field of business administration

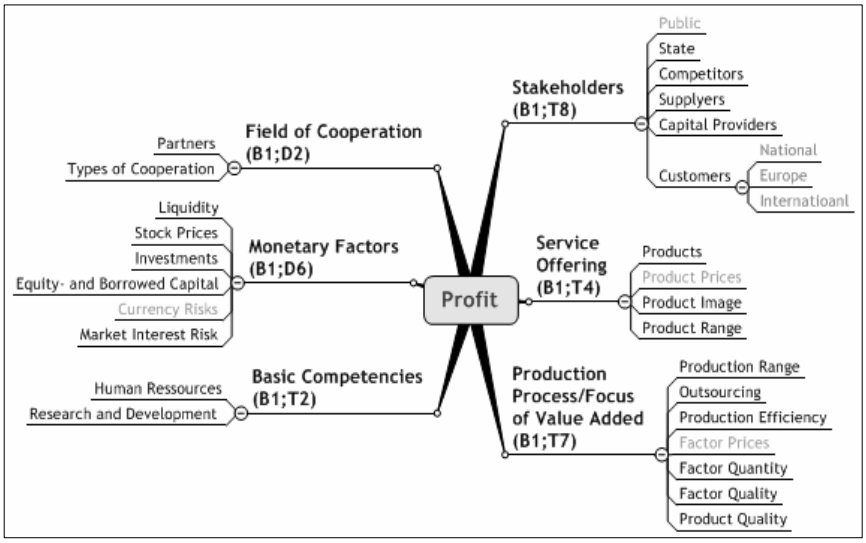


Source: Preiß, 1992, p. 62

Other concepts principally follow this idea but classify the components differently. The reformulation of Blooms taxonomy by Anderson and Krathwohl (2001) explicitly incorporates different knowledge dimensions and metacognition is added as an additional category. The integration of declarative and procedural knowledge is one of the advantages of the reformulated concept since both knowledge dimensions can be applied to solve particular tasks in specific situations. One-sided learning and deep understanding are balanced this way. Measurement tools must be able to assess whether applying concepts, using tools, and interacting with others in different situations are performed competently.

In the context of a PISA-VET, the *development* of competencies plays a substantial role. For example, measuring the development of declarative, procedural, and strategic knowledge in the field of commerce and business administration in Germany is illustrated in *Figure 2.2*.

Figure 2.2: Classification system for the development of knowledge dimensions



Source: DFG-project AC 35-24.1/2, 2005

For the scenario of a profit collapse in a company, one of the tasks refers to a short essay. To measure individuals' competencies, the answers could be classified according to breath (B) and depth (D). In this way, knowledge development can be measured at different points in time. A detailed discussion about measurement models is provided later in this report.

We can conclude that cognitive- and functional competence play a substantial role in a PISA-VET. This implicates domain-specific agreements regarding the goals and contents of VET and the development of suitable items for research.

Functional competence

For the clarification of the relevant underlying occupational subject-related competencies in different educational tracks and occupational tasks, an example of empirical research consisting of significant occupational tasks in German occupations is introduced. On the basis of task observations in the workplace, the Sociological Research Institute (SOFI) has conducted a study, which was supplemented by interviews with experts in the corresponding occupational fields (Baethge & Baethge-Kinsky, 2006)[9]. Skill-lists have been developed according to

9 Subjects of the study: *What* (knowledge, skills, attitudes), *when* (duration, time, and frequency of specific working processes), *how* (quality, level, complexity). A distinction was made between broad domains of internal conditions (knowledge, skills, and attitudes) and specific combinations of knowledge, skills, and attitudes necessary for successfully managing particular occupational situations.

significant activities of performance in different occupational situations, which were classified according to different competence domains. The following occupational fields were investigated:

- Metal and electrical skilled workers (industry),
- Bankers,
- ICT skilled workers, and
- Nursing staff.

A comparison of the different occupational profiles showed the following:

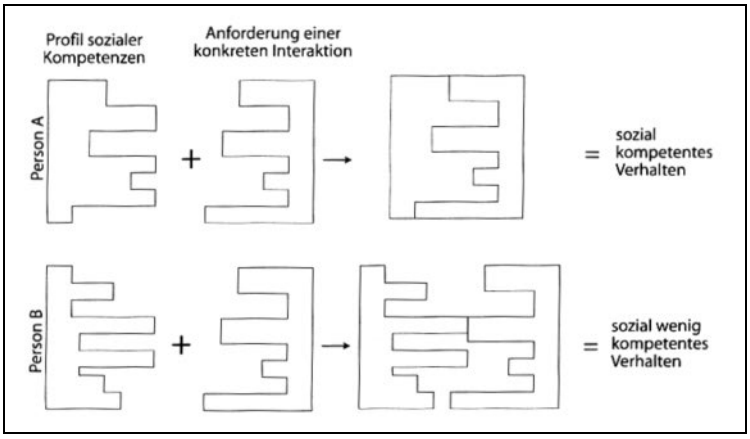
1) Subject-related knowledge is a fundamental constituent of every occupational field. However, there are more knowledge differences *within* occupational sectors than *between* them; knowledge is differentiated according to declarative and procedural knowledge. For example, skilled metal workers demonstrate deep knowledge of materials and processes of metal operations, whereas skilled electricity workers apply highly-developed knowledge of the constitution and functioning of line- and operational-systems and the corresponding electrical, electronic, pneumatic, hydraulic, and software -components.
2) Knowledge domains play a substantial role in every occupational field, with regard to processes of interaction, analytical and reasoning skills, meta-cognitive skills, and organisational participation. Working self-directed within limited time, knowledge decline, time-consuming contact (peers, customers, patients), and susceptibility to failure of organisational processes, are common phenomena in every occupational field across all levels. However, every knowledge dimension has domains-specific characteristic, which cannot be classified as general occupational competencies. For example, in the domain of social competencies there is a substantial difference between the structure of teamwork in industrial, healthcare, or sales industry contexts.
3) According to the occupational domain and the corresponding tasks (e.g., production/maintenance, distribution, care), a sense of perception relating to vocational expertise has an additional impact. For example, a technically skilled worker knows his facilities by touching or hearing, a banker has developed a sense for his customers needs, and a nurse can observe the patients state of health.
4) Sensory-motor skills refer to manual aspects of performance such as speed and accuracy of physical movements or dexterity. They are particularly relevant for skilled workers in the occupational domains of industry (installing, fixing) and care (medicating wounds, giving injections, transferring patients between beds) (Achtenhagen & Grubb, 2001).

The method of work samples (external measurement) is particularly relevant for the domains of functional- and social competence. For example, measurement methods based on the concept of completed actions („vollständige Handlung") could be one possibility.

Social competence

Social competence refers to individual preconditions necessary for successfully behave and interact with others. According to the differentiation between internal conditions and external performance, Kanning (2003) distinguishes two aspects of social competence: The non-observable, individual collection of internal conditions as a precondition to act successfully in social contexts (*social competence*), and observable, socially competent behaviour in social situations (*socially competent behaviour*). Social competence is context-dependent and varies according to different occupational fields (ibid., p. 14). At a different level of aggregation, *general social competencies* without a specific situational context (e.g., adopting a different perspective) can be identified. In contrast, *specific social competencies* denote individuals with particular learning experiences (e.g., elderly caregiver, salesperson). Since both types of social competence develop simultaneously, specific competencies are considered differentiated, experienced-based, general social competencies. General and specific social competencies contribute towards the development of socially competent behaviour in a specific situation. Kanning (2003) illustrates this as follows (*Figure 2.9*):

Figure 2.3: Socially-competent behaviour in relation to the requirements of different situations



Source: Kanning (2003).

Figure 2.3 shows that the composition of internal conditions of one person (A) can be clearly distinguished from those of the second person (B). In a specific social situation, the second person (B) performs worse than the first person (A). In contrast, if the situation corresponded more to the internal conditions of the second person (B), the performance of the first person (A) would probably show worse results than the performance of the second person.

Kanning (2003, p. 21) distinguishes three domains of social competence:

- *Perceptive-cognitive domain:* Self-attention (direct, indirect), perception of others, adoption of perspectives, belief of control (internal, external), making decisions easily, knowledge;
- *motivational-emotional domain:* Emotional stability, pro-sociality, plurality of values; and
- *behavioural domain:* Extraversion, assertiveness, ability to act, communication style (asking for and providing support, appraisal, exertion of influence, expressiveness, listening), conflict behaviour (realising ones own interests, considering other peoples interests), self-control (control of behaviour in social contexts, self-presentation).

Weber (2005) has developed an integrated instrument for the measurement of *mindful identity negotiation.* The instrument, which is based on cultural-historical social learning theory, is an integrated communication model, consisting of 17 reliable items indicating inter-cultural communication skills (ibid., pp. 251–252)—presenting, evoking, facework, supporting others personal identity needs, self-reflection-awareness, self-reflection-questioning, personal identity, and relational identity. The instrument was developed for the evaluation of intercultural learning. The concept of inter-culturality refers to different nationalities, including different hierarchical levels in organisations and cooperation between different occupational fields. The concept provides a basis for the development of measurement tools in a PISA-VET.

2.5 PROPOSAL FOR DATA COLLECTION

The two depicted basic approaches for the identification of competencies, based on external performance or internal conditions, correspond with the following measurement approaches:

(1) Measurement of competencies during work (performance), or
(2) measurement of competencies on the basis of tests (which can include work samples).

There are three aspects against applying the first approach for the purpose of identifying competencies in a PISA-VET:

- Measurements of performance of specific working tasks entails an agreement of performance levels in different occupations with regard to occupational tasks. Moreover, the tasks must be considered relevant in every participating country. The preceding illustrations clearly showed how difficult it is to agree upon typical occupational tasks/bundles of activities across countries.
- Performance based measurement is very time-consuming, in particular with regard to reaching a high level of validity, reliability, and objectivity. For example, the working samples applied in the World Skills Competition com-

prise up to 20 hours, which is feasible given the low number of participants in the contest.

- The superior goals of VET are not incorporated sufficiently. Aspects of work-related qualifications are included, whereas facets of personal development and societal participation are excluded.

Thus, for theoretical and methodical reasons, the measurement of work performance is not sufficient for a PISA-VET. However, this approach might have valid results with regard to occupational subject- specific competencies.

The measurement approach is consistent with the approaches introduced by Erpenbeck and von Rosentstiel (2003a; 2003b) and the Association for Research in Professional Development (2005). They consistently emphasise that the measurement of performance in working processes is too complex and time-consuming and occupational tasks or bundles of activities have to be identified in a very comprehensive and time-consuming analysis to ensure representativeness. Nevertheless, the problem that rarely performed but very important competencies are excluded in this approach remains. Thus, measurement of competencies during work is not recommended in the context of a PISA-VET; simulations and complex measurement methods, as well as biographical records, are the favoured options.

During the workshops, Generalized Work Activities (GWA) were recommended as measurement tool for an international comparative study of vocational competencies; I was emphasised that the tool operates at an aggregated level, and therefore, ensures comparability between occupational fields *and* working activities in different countries. Moreover, GWAs could be measured within a limited period of time.

With regard to the level of aggregation, it is difficult to estimate whether this method is occupation specific enough and sufficiently detailed for drawing inferences from performance to competence and vice versa. In the context of a PISA-VET, the instrument might be suitable for the purpose of triangulations. We consider it, proplematic to apply GWA's as the only measurement tool since it is not adapted to VET issues. Toolsema and Nijhof (2003) have used GWAs to measure competencies in the context of Higher Education. Transferring this tool to VET contexts might be problematic with regard to reliable outcomes.

The second aspect involves both, advantages and disadvantages:

- As in the case of performance measurement, typical occupational tasks or bundles of activities across countries, must be identified. Existing qualifications (e.g., NVQs), as well as concepts for situation-specific assessments, might provide a basis for the development of measurement instruments. Thus, assessment of performance in the domains of cognitive-, functional-, social-, and self-competence seems to be feasible on the basis of adequate tests.
- The advantage of measuring competencies on the basis of tests is it's incorporation in a broad concept of competence:

- The empty cells (*Figure 2.10*) can be filled in to illustrate the various kinds of complex tasks. In addition to the ability to use knowledge in different situations, the underlying knowledge structure can be determined.
- On the basis of measurements for motivation and meta-cognition, correlations between cognitive-, functional, and self-competence can be identified; this in turn allows inferences to be made about internal conditions for lifelong learning.

The complexity and breadth of test items could be adapted according to the intended measurement time. On the basis of multiple measuring points over time, quantitative, as well as qualitative developments could be identified. Moreover, the test items are applicable to retrospective techniques. *Table 2.7* illustrates the approach. The domain of self-competence is pictured above the remaining competence domains to emphasise its superior relation to cognitive-, functional-, and social-competence, as well as the different attitudes, values, perceptions, incentives, motivation, and meta-cognitive strategies. Biographical reports could be adopted alternatively or additionally in every cell. This illustrates that the approach takes into account the overarching goals for VET. However, the final combination of measurement tools depends on the given measurement time.

Conclusion

In a PISA-VET we suggest measuring internal conditions of performance. The corresponding occupational test items must be developed on a national and international basis. Additionally, triangulations based on GWAs are recommended for monitoring purposes. The next paragraph illustrates how adequate test items might be constructed and applied. The concept of completed actions is outlined as an example for measuring cognitive competence in the occupational fields of car mechatronic, nurse, and banker. Moreover, the concept of the learning task (Lernaufgabe), the purpose of actions (Handlungsfunktionen), problem solving, and computer-based simulations are discussed with regard to didactical conceptualisations. Finally, different measurement models are illustrated.

The concept of completed action

The concept of completed action (e.g., Aebli, 1987) or related concepts would be the basis for item development when observation methods are excluded. The model provides information about processes of internalisation and the development of vocational expertise in work situations (*Table 2.7*). The concept was tested in the context of vocational examinations in Germany (Federal Institute for Vocational Education and Training, 1999).

Table 2.7: Concept for measuring vocational subject-specific and cross-occupational competencies (including references for measurement tools)

<table>
<tr><td rowspan="4">Individual Capacities

(accessed and interpreted in different contexts)</td><td colspan="3">Competence Domains
(different contextualised areas of performance)</td></tr>
<tr><td colspan="3">Self-Competence</td></tr>
<tr><td>Cognitive Competence</td><td>Functional Competence</td><td>Social Competence</td></tr>
<tr><td>(theoretical/analytical requirements)

“applying concepts”</td><td>(technical/practical/ functional requirements)

“using tools, equipments and technical resources”</td><td>(interpersonal requirements)

“interacting with others”</td></tr>
<tr><td>Attitudes
Values
Perceptions</td><td colspan="3">Self-efficacy for predicting achievements
(4 items, applied in the PISA survey: Kunter, et al., 2002, p. 168)</td></tr>
<tr><td>Incentives
Motivation</td><td colspan="3">Meta-cognitive strategy: effort and perservance
(4 items, applied in the PISA survey; Kunter, et al., 2002, p. 171)

Interest and motivation
(18 items, Prenzel, et al., 2001; short version applied in a national project funded by the German Research Foundation, DFG)

Meta-cognitive strategy: self-related cognition (verbal)
(3 items, applied in the PISA survey; Kunter, et al., 2002, p. 169)</td></tr>
<tr><td>Metacognitive Strategies</td><td colspan="3">LIST-Questionnaire (inventory of measuring learning strategies)
(22 items, Wild & Schiefele, 1994; short version applied in a national project funded by the German Research Foundation, DFG, α = 0,74-0,81)</td></tr>
<tr><td>Declarative Knowledge</td><td colspan="2" rowspan="2">Complex task for simultaneous measurement of cognitive and functional competence</td><td rowspan="2">Questionnaire for mindful identity negociation

(17 items, Weber, 2005
α = 0,89;

(consists of items for declarative, procedural, and strategic knowledge in the domain for social competence)</td></tr>
<tr><td>Procedural Knowledge</td></tr>
<tr><td>Strategic Knowledge</td><td colspan="2">According to the occupational field:

complex task for measuring cognitive competence
complex task for measuring functional competence</td></tr>
</table>

Car mechatronic

Based on the concept of completed actions, working activities were used as a basis for item generation in the occupational field of a car mechatronics (Wolter

& Petersen, 2002). Following the idea that car accidents occur in every country and the assumption that the corresponding working tasks of buying and fixing accident-damaged vehicles are based on similar competencies, standardised tests have been developed. The tasks provide an opportunity for measuring different competencies in the field of a car mechatronics.

The development of complex tasks based on this model allows the measurement of declarative, procedural, and strategic knowledge in the domains of cognitive and functional competence. The different knowledge domains could be identified by applying concepts and using tools in different contextualised areas of performance. In this regard, computer-based simulations are another option. Additionally, items for measuring interactions with customers (social competence) have to be developed, since bringing a damaged car to a garage represents a critical social situation.

Table 2.8: Elements of the concept of completed actions

Understanding the working task/ determination of goals	– understanding the kind of working task – recognising the demanded product type and quality – if necessary, determining own goals – recognising and understanding overarching goals (e.g., economy)

Analysis of initial situation	– selecting, acquiring, and analysing information
	  – Which information is needed, available, redundant?
	  – Where, how, and who can information be accessed?
	  – How can the information be integrated?
	– recognising which individuals (with the corresponding rights, duties, competencies) have to be integrated into processes of action
	– selection of working tools (in a broad sense: e.g., ICT programmes, payment facilities)
	– recognising external conditions (e.g., time limitations, law regulations)

Determination of further steps	– recognising which steps have to be conducted and in which order – recognising gaps, contradictions and redundant steps, as well as their consequences – possibly compare different action strategies
Execution of further steps	– execution of sub-steps
Monitoring outcomes of action processes	– examine the quality of outcomes – recognising if and to what extent a goal has been reached – recognising errors – detecting error causes – recognising error success

Source: Federal Institute for Vocational Education and Training, 1999

Table 2.9: Buying and repairing accident-damaged cars

Working task	Learning Situation	Unit of Instruction	Time	Central Question
Fixing Accident-Damaged Cars	Examination of vehicle *Didactical focus: System analysis*	Damage estimation	20 h	– What are the functions of the car body part? – What is the composition of the car body part? – What are security features of modern car body parts? – How to perform professional damage examination of car body parts of accident cars?
	Fixing damaged body parts	Analysis of need for repair or total part replacement	60 h	– How to measure car body parts? – What are the criteria of partly and sector repair? – What has to be kept in mind with regard to guarding body parts and painting of new parts? – How to measure undercarriages? – What has to be kept in mind with regard to wheel rims and wheels? – Which types of axis and suspensions exist in modern vehicles? – What is the meaning of wheel position sizes? – How to drive a vehicle? – What are the technical possibilities of undercarriage stabilisation?
	Repairing undercarriage	Wheel rim, wheels, axis, size of wheel position, steering		

Source: Wolter & Petersen, 2002

On an international level, the occupational field of car mechatronic appears to be particularly convenient, since a European occupational profile, consisting of typical working tasks, already exists (*Table 2.9*):

Nursing

A number of international educational programmes (e.g., Department of Administrative Services, Oregon) show that every competence domain in our model is included. *Figure 2.4* illustrates the relation between cognitive-, functional-, and social-competence and the corresponding knowledge domains. The three depicted content areas provide a basis for developing test items. With regard to interactions with patients, simulation-based measurement methods are available.

Figure 2.4: Working tasks of the European occupation profile—Car Mechatronic

1. Standard service
2. Wear repair
3. Vehicle care
4. Administrative service tasks
5. Standard extensions and additional installations
6. General inspection including summer, winter, and holiday checks
7. Administrative service tasks (emissions test, periodical inspection)
8. Repair of damages (chassis, body, steering, etc.)
9. Trouble shooting and repair
10. Repair of aggregates
11. Extension and additional installations
12. Special diagnosis and repair
13. Repair of collision damages
14. Claims
15. Checking systems and alternatives

Source: Rauner & Spöttl, 2002, p. 144

Figure 2.5: Nursing duties and responsibilities

Patient Examination

Typical tasks: examines heart, lungs and respiratory system using stethoscope to identify heart murmurs and determine the nature and extent of other disorders; checks for masses and enlargements in major organs and glands, such as for evidence of enlarged liver, enlarged thyroid, ovarian and breast cysts and tumours; examines teeth, throat, ears, eyes, skin, and other parts of the body for signs of malformation or disease; conducts microscopic examination of samples to identify infections and blood cell counts; draws blood, takes smears, and orders laboratory studies; interprets results of laboratory studies.

Patient Treatment

Typical tasks: develops and carries out complete patient treatment plan as primary provider of health care services; treats infections, sprains, and simple fractures; sutures wounds not involving major blood vessels, nerves or tendons; suggests food supplements, non-prescription medicine, and prescribes antibiotics and other drugs to the extent authorized; refers cases to physicians and health specialists for conditions outside this employee's ability to diagnose or treat according to established protocol; confers directly with one or more physicians regarding diagnosis and care-management problems.

Training

Typical tasks: instructs nurses and health specialists in techniques to expand their diagnostic, counselling, and referral skills; confirms and extends assessments made by clinic registered nurses; authorizes and oversees treatment given under that employee's supervision; explains normal growth and development, abnormalities, diseases, diagnostic methods and treatment to patients and their relatives; follows up with parents, relatives, schools, and health resource organizations to obtain specialized help and or financial assistance, and to promote understanding.

Source: State of Oregon, Department of Administrative Services, Human Resource Service Division, 2005

Bank training

The following example may help to illuminate the structure of complex tasks: On the basis of short essays, different tasks have to be solved. This shows the integration of the different knowledge domains.

Figure 2.6: Complex task for initial instruction in the domain of business and administration

VW korrigiert Gewinnziel nach unten

Mehr als 500 Millionen Euro unter Plan / Neues Oberklassemodell auf Eis gelegt

rit. HAMBURG, 20. Juli. Der Volkswagen-Konzern wird sein Ergebnisziel auch in diesem Jahr verfehlen. Nach Informationen dieser Zeitung rechnet Europas größter Automobilhersteller für 2004 nur noch mit einem operativen Ergebnis vor Sondereinflüssen von 1,5 bis 2 Milliarden Euro. Ursprünglich wollte Volkswagen das Vorjahresergebnis von 2,5 Milliarden Euro übertreffen. Die revidierte Prognose soll am Freitag im Halbjahresbericht veröffentlicht werden, heißt es. Ein VW-Sprecher wollte diese Informationen nicht kommentieren. Den Informationen zufolge ist im ersten Halbjahr – trotz der stabilen Erträge aus der Finanzdienstleistungssparte – mit einem kräftigen Gewinneinbruch zu rechnen. Als Konsequenz aus der abermals verschlechterten Ertragslage hat der Vorstand die Pläne für den Bau des „C1" vorerst auf Eis gelegt. Mit dem Modell, das 2007 in Serie gehen sollte, wollte VW die Lücke zwischen Passat und Phaeton schließen.

Trotz der Gewinnrevision wird VW am Freitag voraussichtlich keine Verschärfung des im Frühjahr eingeleiteten Restrukturierungskurses ankündigen. Damals wurde das Sparziel für 2004 und 2005 auf insgesamt 4 Milliarden Euro verdoppelt. Die jetzige Zurückhaltung hängt womöglich damit zusammen, daß man vor den im September startenden Tarifverhandlungen keine Drohkulisse errichten will. Gleichwohl steigt der Druck auf die Gewerkschaften, bei den Personalkosten kompromißbereiter zu sein. VW will die Personalkosten bis 2011 um 30 Prozent senken.

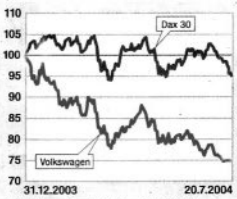
Schon in den ersten drei Monaten des Jahres war das operative Konzernergebnis um 46 Prozent auf 329 Millionen Euro eingebrochen. Seither hat sich das von Konjunktur- und Konsumschwäche geprägte Marktumfeld kaum verbessert. Der Verband der Automobilindustrie erwartet allenfalls noch einen stagnierenden Inlandsabsatz. Auf die europaweit zögerliche Nachfrage reagieren alle Hersteller mit zum Teil kräftigen Preisnachlässen; das hinterläßt deutliche Spuren in den Bilanzen von Ford und General Motors (mitsamt der Tochtergesellschaft Opel). Auch die französischen Hersteller Renault und PSA Peugeot Citroën sowie die italienische Fiat bekommen den Preisdruck zu spüren. Unter den Massenherstellern schwimmt Toyota indes gegen den Strom. Dank der hohen Produktivität eilt der größte japanische Autobauer bei Absatz und Ergebnis von Rekord zu Rekord.

VW hat die Kaufanreize im Verlauf dieses Jahres stetig erhöht. Das gilt vor allem für den neuen Golf: Nach der kostenlosen Zugabe einer Klimaanlage gibt es seit Juni zusätzlich eine Händlerprämie von 928 Euro für die Inzahlungnahme eines Gebrauchtwagens. Dies stabilisiert den Absatz, verengt aber die Gewinnmargen – und zwar nicht nur beim Golf. Denn um die Differenzierung der Konzernmodelle untereinander aufrechtzuerhalten, müssen auch die Preise der übrigen Modelle wie VW Polo, Škoda Octavia oder Seat Ibiza (direkt oder indirekt) gesenkt werden.

Zudem gerät VW auf wichtigen Auslandsmärkten unter Druck. In Amerika macht den Wolfsburgern die Dollar-Schwäche zu schaffen: Bei Durchschnittskursen zwischen 1,20 und 1,25 Dollar je Euro dürfte das operative Ergebnis im Gesamtjahr um rund 1,2 Milliarden Euro belastet werden. Auch toben in Amerika, wo der VW-Absatz unter dem Modellwechsel bei Jetta und Passat leidet, noch härtere Rabattschlachten als in Europa. All dies bescherte allein im ersten Quartal in Nordamerika einen Verlust von 235 Millionen Euro. Der chinesische Markt, auf dem VW in der Vergangenheit blendend verdiente, hat stark an Dynamik verloren. Dort versuchen die Konkurrenten – allen voran General Motors –, VW mit Kampfpreisen die Marktführerschaft zu entreißen. Die Wolfsburger müssen wohl mit einem geringeren Ergebnisbeitrag aus China rechnen.

Toyota erhöht Absatzziel, Seite 16.
Ford profitiert von seiner Finanzsparte, Seite 16.

Volkswagen im Vergleich
Kurs 31. Dezember 2003=Index 100



Quelle: Thomson Financial Datastream (20.7.: 16 Uhr)

Source: Frankfurter Allgemeine Zeitung, 2004

Item for declarative knowledge (reproduction—cognitive competence):
What are the reasons for the profit collapse of Volkswagen?
You are a shareholder of Volkswagen AG. On the 31st of December, 2003 the share price was 17.500 Euro. The following calculations should be based on the information provided in the graphic.

Volkswagen im Vergleich
Kurs 31. Dezember 2003=Index 100

Quelle: Thomson Financial Datastream (20.7.: 16 Uhr)

Source: Thomson Financial Datastream, 2004

Items for procedural knowledge (reproduction—cognitive competence):

What is the current value of the Volkswagen share on the 20th of July, 2004?

What would have been the value if you had invested in the DAX 30 share instead of the Volkswagen share on the 20th of July, 2004?

Items for strategic knowledge (recognition—cognitive competence):

You are annoyed about your loss. Which actions could you (as a shareholder) demand from Volkswagen AG to increase Volkswagen stock prices?

An integration of items for measuring interactions with customers is possible. In this regard computer-based methods provide a feasible way for measuring aspects of procedural knowledge and social competence.[10]

Problem solving and computer-based simulations

To measure performance in VET and labour market contexts, computer-based measurement tools have to be taken into consideration. The German national option of PISA (2000), measured the analytical and dynamic problem solving competencies of 15-year-old students using paper-and-pencil tests in addition to computer-based assessments. Problem solving was defined as "a cognitive process that includes goal-oriented thinking and action in situations where no routine solution procedure is available" (Wirth & Klieme, 2003, p. 330). This broad definition comprises different domains of problem solving processes:

- Problems can vary from limited tasks to complex projects;
- problems can concern different contextualised areas of practice with different levels of authenticity (Achtenhagen & Weber, 2003); and
- test formats can range from open to closed tasks in well- or non-defined problem spaces based on single modules or networked elements.

The listed aspects determine the complexity of a problem: *Figure 2.6* illustrates an example in the field of Business and Administration. Computer-based simulations are suitable for investigating complex tasks since the variety of dynamic (dependent) variables can be illustrated and individual performance can be measured. Individuals have the opportunity to explore simulated situations interactively. Moreover, computer-based simulations allow the development and testing of hypotheses on correlations, knowledge generation, and system monitoring. Within the assessment framework for PISA/Germany, two tests for measuring dynamic aspects of problem solving were developed—search and motion through

10 In addition to the classification of tasks according to levels of complexity, the corresponding task requirements must be specified. This concerns the didactical relevance of tasks. The concept of subject-related didactics is considered in a broad sense since it subsumes workplace- and labour market-related purposes. Regarding a PISA-VET, this entails the development of six different concepts of subject-related didactics corresponding to the respective occupational field. In this context, it has to be examined whether or not there are certain common elements in all categories. Oates (2004) refers to this level as *task-specific but independent of specific jobs* to indicate the transition to occupation-specific descriptions. The GWA instrument is located at a *task-specific but independent of specific jobs*-level providing 42 constructs for evaluating job activity requirements for the majority of occupations in the world.

finite problem space (Finite State Automation), and search and motion through an infinite problem space (Simulation Game). The Finite State Automation test assesses whether the problem solver finds a systematic and efficient way to explore the system, acquires the necessary knowledge about the system, and is able to use this knowledge to control the machines. The Simulation Game assesses whether the problem solver finds a systematic and efficient way to explore the simulated systems and the relationship between them, acquires the necessary knowledge during this exploration, and is able to use this knowledge to pursue multiple, ill-defined goals simultaneously (Wirth & Klieme, 2003). The extent of exploration activities is considered a decisive factor for knowledge acquisition (Kunter et al., 2002; Achtenhagen, Preiß & Weber, 2005).

Research on problem solving attributes the following advantages to computer-based assessment (Wirth 2001; Baumert, et al., 2002; Wirth & Klieme, 2003):

- Individuals are required to continuously search and process external information and to receive feedback for effects of their own system manipulations;
- authentically designed problem situations; and
- discrete states of "correct" and "incorrect" are expanded to in-between-states, consisting of processes between the initial and final states, which provides a basis for measuring outcomes and processes of problem solving. Moreover, indicators for strategies (e.g., problem solving strategies) can be analysed.

The German national option of the PISA 2000 survey was based on several paradigms of experimental research on problem solving for the use in large-scale assessment contexts. Finally, four tests were selected, representing different problem types, contexts, and test formats. Two paper-and-pencil tests were used to assess analytical aspects, and two computer-based tests were applied to assess dynamic aspects of problem solving. The combination of different problem solving instruments shows that analytical and dynamic aspects must be distinguished, given that dynamic aspects show discriminate validity against reasoning and school-related literacies, while analytical problem solving competence is strongly correlated with reasoning (Wirth & Klieme, 2003).

In summary, we can conclude that research on problem solving provides adequate measurement tools for large-scale assessments. However, even though computer-based assessment methods show satisfactory reliabilities it correlates significantly with computer experience. In this regard, gender differences in student outcomes do not have a significant impact. For a PISA-VET, these results are important, since computer-based learning and instructional environments have been applied in German micro-structural VET contexts (Achtenhagen & John; 1992, Achtenhagen, 2004). Contrary to research on problem solving, these environments are more associated with subject-related vocational contexts. For the context of VET the development of large-scale assessment tools should be based on these concepts.

2.6 MEASUREMENT METHODOLOGY AND SCALING: CONSIDERATIONS REGARDING MEASUREMENT MODELS[11]

In the preceding paragraphs, different measurement tools were illustrated with regard to the complexity of developing test items for every competence dimension. Taking into consideration the limited measurement time, valid and time saving assessment tools are required. In work contexts, working samples show highly valid results but can be virtually excluded; immense developmental efforts in every occupational field are exceeding the purpose of an international comparison of VET. The relationship between aspects of education and competence, measurement methods, and efforts in different occupational fields can be illustrated as follows (also see *Table 2.10*).

- *Observation-based methods (working samples*), which are applied in nearly every occupational field for measuring competencies, are very time-consuming. The time required for measuring just one technical content area (German Standards for Technical Literacy, published by the Association of German Engineers, 2004) amounts to 3 hours. Measurement in other fields requires even more time (e.g., nursing, complex task in the field of skilled electrical work)
- *Paper-and-pencil-based working samples* usually refer to a given scenario with different tasks, such as short essays, requirement specifications, or technical sketches. The advantage of this method is that it can be performed within a manageable period of time. For example, the written task in the field of skilled electrical work in Germany provides a basis for a feasible measurement of complex cognitive and functional competencies within a limited period of time.
- The less time consuming method is the *multiple-choice task*. In Germany, it is used, for example in the context of skilled worker examinations (PAL/AKA). It is applied for measuring declarative knowledge in limited content areas. However, for measuring procedural and strategic aspects of knowledge, this method is not sufficient. Therefore, a combination of a paper-and-pencil working scenario followed by multiple-choice tasks and open questions seems to be a feasible option.

In the context of a PISA-VET with limited time for measurement, adequate methods for replacing work samples based on observation have to be developed. However, at the moment there are at least two remaining questions:

- Is it possible to measure different competence dimensions on the basis of one single assessment method? In theory this seems to be possible: For example, an essay might be assessed by evaluating various dimensions, including

11 The following comments are based on a discussion with Professor Jürgen Rost (IPN Leibniz Institute for Science Education, University of Kiel) on the 7th of September 2005 in Goettingen.

length, complexity, number of arguments, etc., and the answers might be combined (multidimensional items).

- Should the dimensions be measured separately or simultaneously on the basis of different assessment methods, such as a combination of different item batteries and essays? The latter is favoured to prevent symptoms of fatigue.

After all, existing research on different assessment methods will certainly provide a solid basis for developing adequate measurement tools. As an example, assessment of essays on the basis of multiple-choice questions and questions requiring open-ended answers (OECD, 2001; Altiner et al., 2002; Scholz, 2003) could be used as a tool for reliable assessment of competencies in VET. Finally, there is research on computer-based tests for simulating technical equipment malfunctions as a basis for troubleshooting and fault diagnosis.[12]

Adequate measurement tools should be able to identify single scores, and at the same time, overall scores for every competence dimension. In principal this is possible independent of the underlying assessment method. Nevertheless, modelling testing tools with regard to single *and* cross-sectional competencies is problematic: Firstly, developing tasks with distinguishable levels of complexity is very difficult; and secondly, every assessment method has to refer to one single reference system (Rost, 2004).

The question of how to relate cross-sectional measurement to competence development over time (longitudinal measurement) will be particularly challenging. In this regard, the variety of illustrated tasks provides a solid basis for rapid development of adequate tasks and measurement tools.

12 There is no empirical research on integrated learning tasks (Federal Institute for Vocational Education and Training) available yet.

Table 2.10: Aspects of education/competence, measurement methods, and efforts in different occupational fields

Competence Domain/ Occupational Field	**Aspects of Education/Competence**	**Measurement Methods**	**Effort**
Standards for Technical Literacy/ Content area: Work and Production (Association of German Engineers)	Applying and assessing methods of planning, construction, production, operation, maintenance.	Work samples based on observation and written tasks	3 hours
	Knowing the main characteristics of materials and determining their suitability for the intended products.		
	Considering organisational principles of the work place and shop.		
	Possessing basic skills for using tools, equipment, and technical resources.		
Nursing Education (Open University of the Netherlands)	Care for ill, disabled, and dying patients.	Work sample based on observation (patient consultations) and written tasks (identification of symptoms, care and treatment)	5 days
	Preventive measures at different levels for reducing health risks.		
	Informing patients and affiliated for promoting a healthy life style.		
Metal and Electrical Occupations (Industry) – 1st part (Borch, et al.)	Analysing documents, planning work processes, disposing tools and equipments.	- Work sample (complex task) partly consisting of written tasks - Multiple-choice test (subject-related knowledge)	12 hours
	Installing, wiring and connecting parts/units.		
	Analysing safety of facilities and equipments and checking electrical safety programmes.		
	Analysing electrical systems, checking functions, identifying errors, adjusting and measuring performance data.		
	Putting into operation, transferring and explaining products, documenting orders.		

Metal and Electrical Occupations (Industry) – 2nd Part (Borch, et al.)	Self-organised work in the domains of organisational processes and quality management (organisational structure).	Documented work sample	18 to 30 hours
	Technical problem analysis; developing solutions taking into consideration technical, economical and organisational factors; selecting user-friendly specifications, components, and methods.	Written task (change of a system based on customer requirements)	2 hours
	Analysing technical documents; selecting methods for measurement, assessment, and diagnostics; analysing functional interrelations; determining cause of errors.	Written task (analysis of functions and systems)	2 hours
	Analysing economical and societal inter-dependencies at work and in the labour market.	Test (economic and civic education)	1 hour
Particularly in Technical and Industrial Fields (Federal Institute for VET)	Understanding and analysing the concept of integrated tasks (different examples).	Learning task	2 hours to 5 days

Sources: Adapted and modified from the Association of German Engineers (VDI), 2004: Standards for Technical Literacy; Borch, et al., 2004: Industrielle Elektroberufe; Smits, et al., 2005: Performance Assessment in Nursing Education; Federal Institute for Vocational Education and Training (BIBB), 2005.

Pretest: Generating and testing proficiency levels

The development of test instruments is not part of this feasibility study, but a later phase of pilot testing. With regard to the aims of the study, however, adequate testing theories, in addition to heuristic testing models, are essential to validate different tests. A detailed examination of different test theories and -models suggests using Item-Response-Theory (IRT) together with the corresponding Mixed Rasch Model (Rost, 2004):

1. Item response theory (IRT) is an alternative to classical test theory. IRT are stochastic models for responses to individual test items. The probabilities of these responses are modelled as a function of separate parameters for the item and for the person. It is a method recommended in cases of:
   - New tests being developed;
   - large amounts of data being missing;
   - developmental measurements (repeated measures using the same person) being part of the design; and
   - conclusions regarding partial competencies being derived on the basis of sub-tasks.

   With regard to panel studies and the repetition of measurement (performing the same test with the same person), the impact of structural and quantitative changes on individuals' competence developments must be determined. In this regard Item Response Theory provides a basis for explaining these aspects.
2. The Mixed Rasch Model, one specific version of the Rasch model, is based on the idea that the Rasch model does not hold true for a given data set although it could be expected to hold true for theoretical considerations. In this case, the reason may be that the sample consists of more than one subgroup (class) of subjects in the sense of a mixture-distribution model. Within each of these classes the Rasch model is valid. This means that the items are measured unidimensionally within each class, but also that the classes differ with respect to rank, order, and the level of difficulty of the various items. Therefore, a different latent dimension of ability is being measured in each class. The Mixed Rasch Model allows data analysis based on pre-tests to:
   - Select those items out of a given number of items necessary for optimising the test regarding unidimensionality and classify them into unique (clearly separable) dimensions;
   - identify person groups that are excluded for Rasch model scaling on their competencies (overlapping dimensions); and
   - distinguish competencies of groups or individuals regarding different answering structures (Rost, 2004).

The development of Rasch scales allows the determination of scores for single competence dimensions as well as for cross sectional-competencies. There are different types of Rasch scales: The PISA survey (OECD, 2001; Scholz, 2003) was mainly based on dichotomous Rasch scales (task success, task failure).

Recent developments have generated polytom (ordinal) Rasch models, which proved that existing tests, for example the Big Five personality test (five factor model), can be considered Rasch scales (Rost, et al., 1999). In principal, dichotomous scales can be developed on the basis of scenarios with open answers, and then, binary coded according to different aspects (number of answers, number of correct or false responses, number of different arguments, structure of the arguments, etc.). The difficult part is to preserve local random independency of different items, which entails that the tasks must be developed in a way such that it is possible to solve one task successfully even if the remaining tasks were performed without success.

The complexity of test improvement cannot be precisely determined at the moment. We can assume that for every occupational field, an item pool consisting of 20 dichotomous items in every competence dimension, must be developed and tested. The items must be embedded in complex tasks. In this regard, about 12 items in every dimension are satisfactory for meeting the objectives of the measurement model. In every country about 100 test persons in every occupational field will be needed to minimize measurement errors.[13]

13 Given the heterogeneity of persons and tasks that are involved in the assessment, we are satisfied with these statements. The Education Testing Service (ETS) suggests a comparable number of items for testing computer literacy.

3. INSTITUTIONAL AND INDIVIDUAL FACTORS INFLUENCING QUALITY IN VET

The measurement of competencies in initial VET must be the focus of a PISA-VET. However, international educational research is in agreement regarding the relevance of institutional and individual factors for the development of individual competencies (European Union, 2001). In this regard, differences in the students' competence-profiles cannot be described solely with reference to learning preconditions and individuals' dispositions; the educational organisation in its social, cultural, and economic context has to be taken into account (Baumert & Schümer, 2001). Thus, in addition to measuring vocational subject-related competencies, a PISA-VET should be able to explain the impact and weight of outcome factors relating to the utilisation of acquired competencies in the work place, the labour market, and for individual career aspirations. Institutional and individual context factors help to explain these interdependencies. Thus, a PISA-VET will systematically incorporate fundamental institutional and individual variables of initial VET into the survey design. Finally, a PISA VET will help to identify critical contextual factors for developing and using competencies but will vary in its approach according to differing national education- and labour market structures.

Based on these considerations, an adequate research design has to account for a systematic identification and analysis of the differences in VET processes. In agreement with international education experts, we decided upon a multi-level approach—an analysis of system-, school-, and instruction characteristics as well as their influences on the development and use of competencies, taking into account interactions between individual and social factors. The framework is based on the theoretical model of the PISA survey (Scheerens, 1990). The operationalisation of influencing factors is based upon the same model and is conceptualised according to input-, process-, and output variables at different levels of analysis (Baumert, et al., 2001; OECD, 2003; Baethge, et al., 2003). With regard to institutional conditions, there is a distinction between systemic conditions, provided by the respective national educational system, and specific conditions of different educational institutions in the represented countries (Scheerens & Bosker, 1997). The impact of different factors regarding a PISA-VET will be discussed in the following paragraphs.

A set of indicators, consisting of significant information about the quality of VET, has been developed on the basis of a comparison of institutional and individual factors in comparative-education statistics. This set of indicators provides a basis for the classification and analysis of existing institutional and individual influencing factors referring to the quality of VET.

To identify which indicators are appropriate for evaluating the quality of educational systems, the types of indicators and comparisons which are currently used were the focused: Firstly, we analysed the methods and purposes for describing micro- and macro-structural factors of educational processes; and secondly, these indicators were examined according to their suitability for an international comparison of VET. This allowed a comparison of concepts for a PISA-VET with existing international educational assessments and the identification of specific methodological problems. The subsequent chapter highlights the relevant institutional and individual factors influencing the quality in VET. Finally, on the basis of existing international data, we identified which data are missing and must be generated for the purpose of a PISA-VET.

3.1 CONTEXT CONDITIONS IN INTERNATIONAL COMPARATIVE STUDIES OF EDUCATIONAL SYSTEMS AND -PROCESSES

In the context of recent international developments, interest in national educational systems and policies has increased. At the European level this tendency has been promoted by recent studies, such as the OECD publication "Education at a Glance" (OECD, 2003) or the measurement of competencies of 15-year olds in compulsory education (PISA) at an international level. Improving the quality of national VET-systems in an integrated European process has become an important part of the work to achieve the Lisbon goals (BLK, 2002; European Union, 2001; European Commission, 2005). Different aspects of education policy are compared in an international context for drawing inferences concerning national education systems. Particular emphasis is put on aspects of transparency with regard to improving the understanding of VET systems, objectives, and outcomes (Corti, 2000).

The following paragraph summarises and analyses the objectives and methods of the reviewed literature with regard to the following two aspects:

- Whether they contribute towards identifying the relevant context factors for developing and utilising competencies in initial VET, and
- whether they provide convenient data for an international comparison of VET systems and processes.

International data collections provided by UNESCO, OECD, and Eurostat (UOE)

The "UOE Data Collection" is administered jointly by the United Nations Educational, Scientific, and Cultural Organisation Institute for Statistics (UNESCO/ UIS), the Organisation for Economic Co-operation and Development (OECD), and the Statistical Office of the European Union (EUROSTAT). The objective is to provide internationally-comparable data on key aspects of educational systems, specifically on the participation and completion of educational program-

mes, as well as the cost and type of resources dedicated to education. The data collection mainly consists of quantitative data in the areas of finance, human resources and education access, participation, and outcomes. The "UOE Data Collection" is an integrated part of many projects on education statistics, such as the "World Education Indicators" (WEI) administered by UNESCO, the "International Indicators of Education Systems" (INES), provided by OECD, the "Key Data on Vocational Training in the European Union," prepared by EUROSTAT, and others (BLK, 2002).

The most comprehensive set of comparable indicators of education is provided by the 2003 edition of "Education at a Glance" (OECD, 2003). It consists of information on the output of educational institutions and the impact of learning; financial and human resources invested in education; access to education; participation and progression and learning environment and organisation of schools. Comparisons of qualification levels are mainly based on the ISCED classification system and very little activity has been focused on the VET sector; that is, most of the statistical collections do not provide a separate category for VET (Van den Berghe, 1997; NCVER, 2004). To anticipate the risk of a double count during the separation between school- and VET-based qualification, a significant amount of indicators is merely classified at an "upper secondary" level, and very little information is provided at ISCED level III (A–C) or vocational upper secondary education.

The "Global Education Digest" is the central report of the UNESCO Institute for Statistics (UNESCO, 2004). It provides an additional analysis of school life expectancy, progress towards education goals, and a survey of national compulsory education standards. However, data are provided at a highly aggregated, upper secondary level without information on the quality of VET.

Easily accessible information on the relative performance of national VET-systems are provided by Eurostat's publication, "Key Data on Vocational Training in the European Union". It consists of a series of reports on the structure, modalities, and socio-economic background of students in different VET systems in EU-member states. The information is mainly based on administrative data and additional surveys, primarily descended from the "European Labour Force Survey" (Kogan & Müller, 2003; European Union, 2001).

Even though the illustrated reports on education statistics do not provide much feasible data for an international comparison of VET, they represent a solid basis for identifying the relevant contextual factors of learning conditions in schools. Thus, they contribute towards reaching an agreement with regard to an international comparison of education (OECD, 2003). However, the main purpose of the "UOE Data Collection" refers to policy-related, influencing variables.

UOE data comprise: *Access to education, participation, and progression-rates, socio-economic background, qualifications, financial and human resources invested in education, learning environment and organisation of schools, labour force participation, earnings and educational attainment,* and *economic ratios.*

European Labour Force Survey

The "European Labour Force Survey" (ELFS) is a large household sample survey providing individual variables for the analysis of general labour market characteristics, such as employment, unemployment, and industrial sectors. The main individual factors, such as age, gender, marital status, or nationality contribute to a set of standardised European data on education participation in co-operation with Eurostat. With regard to identifying VET graduates, precise information about age is available, contrary to educational qualifications, which are not so precisely documented. To conclude, although the "European Labour Force Survey" data are not suitable for a comparison of VET, they allow the determination of individual factors relevant for analysing different strategies of young people's school-to-work transitions and subsequent educational developments in the labour market.

ELFS data comprise: *Social background, participation in educational programs, first significant job after completing vocational training,* and *occupational position.*

Continuing Vocational Training Survey (CVTS)

The Department for Education and Employment in the UK is responsible for the European Community Continuing Vocational Training Survey (CVTS). CVTS covers employers with ten or more employees and was conducted across all the 12 EU member states. A follow-up survey, CVTS2, took place in 2000/01. CVTS2 collected similar data as before but covering 25 European states (Grünewald, et al., 2003). The survey was designed to complete existing information on firm-based training. In 2006 CVTS will be extended to cover data on initial training in addition to continuing training (Federal Institute for Vocational Education and Training, 2005). Thus, in the future, CVTS will provide data on firm-based training. However, up to now, provided information has been limited to firm-based continuing training.

CVTS data comprises: *Existing establishments, types and providers of continuing education, participation rates in courses, hours of attendance, financing and costs, firm size,* and *professionalism of continuing education.*

CEDEFOP monographs

The CEDEFOP (European Centre for the Development of Vocational Training) monographs, provide a summary of facts on general and vocational education in EU member states. The reports consist of descriptive characterisations of general information based on different sources (CEDEFOP monographs, 1999-2002). However, a theoretically well-founded, comparable illustration of indicators in different countries is not available.

CEDEFOP monographs comprise: *A description of systemic context factors (social, demographic, and economic contexts of the respective VET-system), historical background of national VET systems, a description of initial VET systems,* and *references to system regulations and financing.*

In conclusion, the illustrated surveys provide an empirical basis for comparing structures of educational systems. However, information on the quality of VET and the developed competencies is not provided: Educational qualifications are considered indicators for successful learning and individual competencies. In the following paragraphs, international comparative studies on competence measurement will be analysed regarding different institutional factors on the quality of learning conditions.

The PISA study

The "Program for International Student Assessment" (PISA) was developed by the OECD to assess the reading, mathematics, and science literacy of 15-year olds in the participating countries. PISA represents the most comprehensive international survey to date for assessing how well students are prepared for life beyond the classroom. The study's focus is on measuring knowledge and skills applied to problems in a real-life contexts. PISA is conducted every three years, with a primary focus on one competence area each cycle (PISA 2000—reading literacy, mathematics in 2003, and science literacy in 2006).

In the PISA framework, only learning environment context factors in schools are measured using their own measurement instruments. Descriptions of national educational systems are mainly based on data collections provided by the OECD "International Indicators of Education Systems" (INES) project, which is supported by international and national documents and secondary data collections (OECD, 2003; Baumert, et al., 1999). In Germany, in addition to international measurements, questionnaires for school principals, students, and parents were implemented: The questionnaire for school principals primarily consisted of questions regarding quality assurance and regional co-operation (e.g., social capital of schools), whereas the student questionnaire related to context factors relevant for personal development, such as peer relationships and media competence. In contrast, the international questionnaire was focused on the resources of school personnel and on educational materials (Baumert, et al., 1999).

PISA context factors comprise: *resources of school personnel and educational materials.* (In Germany, these factors also include: *school organisation, supply of learning opportunities, quality assurance, social capital of schools (regional co-operation), instructional quality, school climate, socio-economic background, migration, family structure, upbringing and domestic support, peer-group integration,* and *leisure-time activities).*

Adult Literacy and Life-skills (ALL) survey and International Adult Literacy Survey (IALS)

The Adult Literacy and Life skills (ALL) Survey is an international comparative study to provide participating countries with information about the skills of their adult population. ALL measured the literacy and numeracy skills of a nationally-representative sample of 16- to 65-year-olds from six participating countries.

ALL builds on the foundation of earlier studies on adult literacy. Chief among these earlier studies is the International Adult Literacy Survey (IALS), which was conducted in three phases (1994, 1996, and 1998) in 20 nations. On a pilot basis, ALL also measured adults' problem-solving skills and gathered information on their familiarity with information and communication technologies. Thus, the adult competencies measured in ALL and IALS can be compared to student competencies measured in PISA. These studies provide a lot of information regarding interrelations between socio-economic background variables and other factors, such as continuing education behaviour. Compared to PISA, influences on education policy are limited due to the lack of incorporated institutional context factors (BLK, 2002).

ALL and IALS context factors comprise: *occupational status, income, average working time, highest completed level of education, parent's education, nationality, participation in continuing adult education, type of educational program,* and *civic participation.*

Programme for the International Assessment for Adult Competencies (PIAAC)

The Programme for the International Assessment for Adult Competencies (PIAAC) has put forward the ambitious aim of developing a strategy to address the supply and demand of competencies that would i) identify and measure differences between individuals and countries in competencies believed to underlie both personal and societal success, ii) assess the impact of these competencies on social and economic outcomes at individual and aggregate levels, iii) gauge the performance of education and training systems in generating required competencies, and iv) help to clarify the policy levers that could contribute to enhancing competencies (Green, 2004; OECD, 2004).

In order to reach these objectives, the following aspects regarding the conditions of developing and utilising competencies will be measured: *Information on national labour markets, socio-economic and demographic factors, participation in formal and informal firm-based and external VET programmes, characteristics of work places/firms,* and *information on social capital (e.g., civic participation, health).*

Since PIAAC is still in the planning stage, detailed information on the measurement of context factors is not available at present. The first cycle of data collection is anticipated being available in 2008-9, with reporting in 2010 (Green, et al., 2004). At the moment, measurement tools are being developed.

Table 3.1 summarises the illustrated international comparative studies in education with regard to whether they describe the relevant micro- and macro-structural context factors of educational processes, and whether indicators and outcomes have a practical value for an international comparative study of VET:

- At an individual level, the focus is on the measurement of socio-economic background variables; if necessary, supplemented by the current employment

and occupational status. This information will be linked to competencies developed in VET, or occupational/personal contexts after completion.
- Information on financial-, personnel-, and material resources, utilised by educational institutions, will be linked to competence levels and qualifications.

Conclusion

With regard to information for comparing institutional arrangements and competencies in VET, the studies proved to be fundamentally incomplete, which indicates desiderata of research:

There is a lack of harmonised data regarding the complexity or breadth of single VET programmes, such as the relationship between practical and theoretical parts, demand and supply, overall budget, quality standards, and transitions from education to work. This is mainly due to the fact that data are provided at a highly-aggregated level, referring to primary-, secondary-, and tertiary levels of education. This also explains why specific characteristics of initial VET are hardly taken into consideration (NCVER, 2004). The international experts at the workshops were in agreement with this estimation.

Additional sources focus on the measurement and presentation of structural data, which do not provide a basis for illuminating qualitative processes of VET: Research designs based on input/output-models address learning environments and -institutions as mere "black-boxes" (Scheerens, 2004); there is a lack of data referring to aspects such as instructional quality and methods of instruction and learning. In this regard, the context factors provided by PISA are exceptional since they allow data analysis to be made at a lower aggregation level.

Finally, the reports provide little information on the type of competencies young people have acquired after completing VET. Competencies are regarded in terms of certificates of formal education, and not in terms of the underlying structures of knowledge and skills. Thus, measurement concepts of adult competencies provide very little information regarding the effects of different educational structures on competence developments. In this regard, a comparative study of VET could provide internationally-comparable data on existing interrelationships between the competencies *actually* developed in VET, the corresponding context factors of educational systems, and the utilisation of competencies in the labour market. Moreover, it could provide reliable assessment of institutional efficiency, which would be beneficial for educational policy issues. This aspect was mainly emphasised by Pascaline Descy (CEDEFOP) during the first workshop. She pointed out that a comparison of VET is nearly impossible on the basis of existing statistical data in different European countries; national educational systems are very diverse and correspondingly, a classification is very difficult (CEDEFOP, 1998).

Table 3.1: International comparative studies in education

Source	Content	Main Institutional and Individual Factors Covered	Statistical Unit	Countries	Strong and Weak Points
UNESCO OECD Eurostat (UOE)-Database; INES	Collection of demographic, social, financial, and organisational indicators, which cover almost the entire formal educational system	Access to education, participation, and progression - rates, socio-economic background, qualifications, financial and human resources invested in education, learning environment and organisation of schools, labour force participation, earnings and educational attainment, and economic ratios	Programme, Household, Individual	OECD Countries	Internationally-comparable data about participation, completion, and expenses and resources for education; Information about VET only available on a high aggregation level
European Labour Force Survey (ELFS)	Data on individuals' participation in different programmes of continuing education	Social background, participation in VET programmes, completed VET programme, first significant job after completing vocational training, employment status, and occupational position	Household, Individual	EU member states	Tries to cover the resident population. Connection with International Socio-Economic Index of Occupational Status (ISEI) and other information (transition from school to work) is possible; graduation level as an indicator for skills and knowledge; the only factor that can be relied upon with certainty is information on age groups
Continuing Vocational Training Survey (CVTS)	Data on continuing education in firms	Existing establishments, types, and providers of continuing education, participation rates in courses, hours of attendance, financing and costs, firm size, and professionalism of continuing education (structural context factors)	Company	EU member states	Data derived solely from enterprises—only activities for the improvement of workplace performance are taken into account Information about initial VET will be integrated in 2006
CEDEFOP monographs	Data and facts about general and vocational education in EU -member countries	Social, demographic, and economic context of different VET-systems, structure of initial VET-systems, system regulation, and financing	System, Programme	EU member states	Basic information on context factors in national educational systems
PISA	Survey of policy-oriented international indicators of the competencies of 15-year old students. Covers the domains of reading, mathematics, and scientific literacy	Resources of school personnel and educational materials, quality assurance, instructional quality, school climate, socio-economic background and cultural capital of the family, leisure-time activities, and peer-group integration	Individual	32 countries, incl. 28 OECD Countries	Internationally-comparable data on the skills and knowledge at the end of compulsory school; considers only the 15-year olds
ALL / IALS	ALL seeks to profile the skills of adults in the domains of prose and document literacy, numeracy, problem solving, team-work, and ICT IALS: seeks to creation of comparable literacy profiles among the adult population	Employment- and occupational status, income, average working time, highest completed level of education, parents' education, nationality, participation in continuing adult education, type of educational program, civic participation	Individual	Several OECD Countries	International comparable data on literacy (IALS), numeracy problem solving, team-work, and media competence (ALL)
PIAAC	Internationally-comparative study for the measurement of adult competencies.	In particular, participation in formal and informal VET programmes, social capital	Individual	Several OECD Countries	International comparable data of competencies and learning biographies of adults; no indicators available yet.

3.2 INDICATORS FOR DESCRIBING INSTITUTIONAL AND INDIVIDUAL CONDITIONS OF VET: AN INTEGRATED MODEL OF VET EFFECTIVENESS

Until now indicators of educational quality provided in international comparative studies were examined. In this chapter, single indicators will be analysed systematically and summarised in a model for comparing competence development and utilisation at an international level.

International comparative studies on education systems consistently assume that both, institutional and individual conditions affect the quality of learning processes and their outcomes. Learning is considered an integrated process of personality development in interaction with the environment (Lempert, 1979). This comprehensive view prevents one-sided concepts, focusing either on individual or environmental factors of learning. In educational and training processes, institutional and individual conditions continuously interact. Therefore, the impact of each factor can be determined only by relating them to the others.

Research on educational effectiveness has generated approaches, such as the economic education-production function, input/output studies of education, and others (Scheerens, 2004; OECD, 2003). These approaches are based on the assumption that educational institutions have the function of "producing" qualification achievement. The following equation is an example for a well-known education production function (Weiß & Preuschoff, 2004): A student's knowledge at a particular time (**K**it) is a function of accumulated influencing factors relating to family background (**F**it), peer-group (**P**it), school (**S**it), and individual competencies (**C**i):

$$\mathbf{K}it = f(\mathbf{F}it, \mathbf{P}it, \mathbf{S}it, \mathbf{C}i)$$

The aim of an analysis on the basis of an education-production function is to gain information about the nature and specific explanation of each independent variable regarding its criterion variance (output -variable, dependent -variable) as well as the influence of different variable interrelations.

From a policy perspective, the main interest in education-production functions is to obtain information about more efficient resource allocation. However, the results of different meta-analyses do not confirm significant correlations between expenditures on education and students' performance (Hanusek, 1989; 1997): "With over three decades of analysis, new studies have reinforced earlier conclusions … there is no strong or consistent relationship between variations in school resources and student performance" (ibid., 1997, p. 141). With regard to American surveys, Weiß and Preuschoff (2004) point out that increasing expenditures per pupil *is* associated with higher student achievement. However, a precise determination of optimal financial resources to obtain more favourable results is not possible.

The outcomes of education-production functions indicate fundamental theoretical shortcomings. The main objection refers to the assumption that school resources are direct by influencing competence development. It is obvious that distal variables of instruction cannot directly impact students' achievements; they rather have a mediated, indirect influence (e.g., on learning opportunities, quality of instruction). Recent studies, investigating the impact of educational expenditures assigned for different purposes, confirm this assumption (Weiß & Preuschoff, 2004).

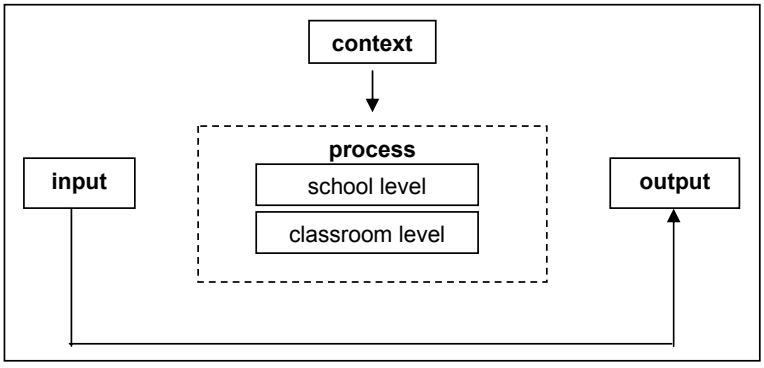
The results of PISA and earlier surveys on learning and instruction support that the utilisation of education resources positively correlates with students' learning achievement (Treiber, 1982).

Consequently, a PISA-VET should focus on illuminating the processes in VET and not only on revealing the impact of relevant input and output-variables, and on "breaking up" the "black box" of learning places, environments, and processes. The need for this type of conceptualisation was also emphasised by the international experts at the two workshops. The results of illuminating correlations between context factors and competencies will help in gaining insight into VET-output and encourage the implementation of different educational policy actions.

There was consensus among the experts that input-, process-, and output-variables and their interdependencies must be conceptualised at four different levels:

- Macro (system)- level of national institutions and organisations;
- meso-level of educational institutions (schools, firms);
- micro-level of processes of interaction in educational institutions (in the classroom, workplace); and
- individual level (apprentices, students).

Figure 3.1: A Basic Systems Model for the Functioning of Education



Source: Scheerens, 1990

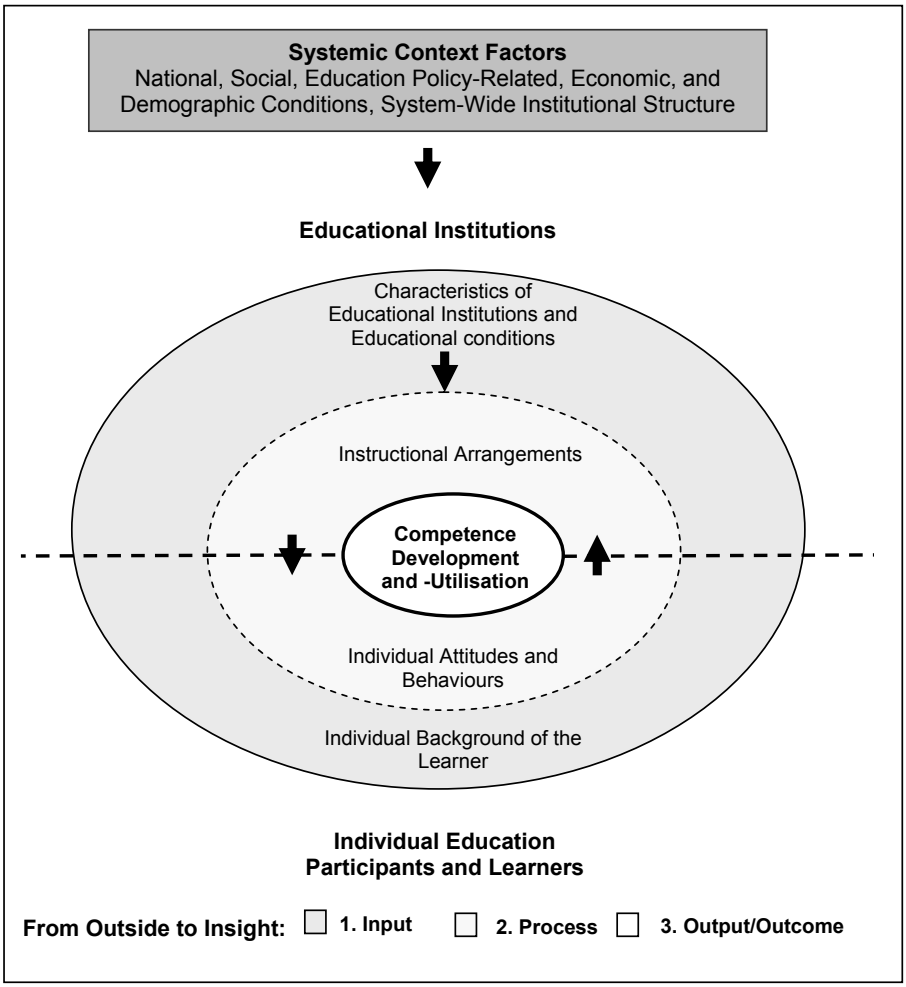
Creating a benchmarking for VET systems on the basis of system factors cannot be recommended due to the internal heterogeneity within and between countries (CEDEFOP, 1998). For example, the German VET-system is so differentiated that it can no longer be regarded as *one* VET-system. System characteristics, such as regulation and financing modalities, legal and administrative guidelines, etc., merely provide system-wide information for determining and explaining specific attributes of educational sectors and the corresponding institutional levels (schools, firms, instruction, workplace).

For characterising individual and institutional conditions of competence development and utilisation, we distinguish between the following four aspects:

- Institutional conditions and requirements of the educational system as whole (context level, police regulations);
- context factors and requirements at the level of educational institutions and individual background variables of learners (input level);
- processes of institutionalised education relating to the level of operationalising material (technical equipment, instructional materials) personal (quailfication of teachers, knowledge and attitudes of apprentices), and symbolic resources (e.g., curricula) in didactical settings;
- output of educational processes at the level of certificates and their utilisation in the workplace, labour market, and personal biography (outcome or output level).

The concept for a PISA-VET which incorporates the depicted correlations between institutional and individual conditions for developing and utilising competencies, can be illustrated as follows (*Figure 3.2*):

Figure 3.2: Correlations between institutional and individual conditions for the development and utilisation of competencies



In the following chapter, we analyse institutional and individual factors at input-, process-, and output-levels to identify the factors with the greatest impact for explaining differences in individuals competence developments and utilisation, as well as the corresponding indicators for measurement. The most important indicators of the illustrated international studies are related to the input/process/output -model (*Figure 3.2*), supplemented with the missing aspects relating to initial VET.

3.2.1 The educational/occupational system

The need for overall information on education- and employment-systems is based on the assumption that a comparison of different initial VET programmes in the context of a PISA-VET can only be explained against the background of the corresponding social, education policy-related, economic, and demographic conditions of the respective country. As mentioned earlier, this type of information does not represent a basis for creating a benchmarking for national VET-systems; rather, it helps to illuminate information at a lower aggregation level and allows interpretation in macro-structural contexts.

At a macro-level of the education- and employment-systems, each viewed as a whole, we distinguish between the following categories:

The educational/occupational systems

1. Systemic Context Factors
2. Systemic Institutional Constitution
   - Coordination and Steering
   - Standards and Norms
   - Financing

Systemic context factors

The mapping of systemic context factors indicates that VET-processes are embedded in other social sub-systems. In this regard, the social, cultural, economic, and political conditions are having a direct influence on VET-systems and must be identified. For example, the actual economic growth rate in a given society (possibly differentiated according to regions), the status of the public budget and fiscal policy, demographic development and the basic structures of social welfare, cultural factors (e.g., family structures), the general expectations for education in a society, and the attitudes towards migration must be considered (Lassnigg, 2003).

Systemic institutional constitutions

For describing processes at the macro-level of education- and employment-systems, we propose three indicators: Coordination and steering of VET-systems, standards and norms (including curricula), and financing of training.

Coordination and steering

One important function of institutional coordination is to ensure adaptability of educational systems to individual and social needs and labour market require-

ments (Descy & Tessaring, 2001). The corresponding institutional factors provide information regarding the type of institutional steering and coordination that optimally ensures this adaptability.

One central aspect of the analysis refers to characterising the coordination system and its various functions of supply and demand. This refers to the relationship between *internal* actors of educational systems and the most important *external* stakeholders of the economy and VET representatives. In this regard, Descy and Tessaring (2002) suggest three types of national settings for relating education and work to one another—close, loose, and varied settings (*Table 3.2*).

Table 3.2: National settings of relating education to work

Type	**Setting**	**Country**
I: Close	Close relationship between educational system and labour market, including a tracked system of education and a qualification structure which has direct relevance to occupational entry	Austria, Czech Republic, Denmark, Germany, Netherlands, Hungary
II: Loose	Loose relationship between educational system and labour market, with a flexible match between qualifications and occupations or jobs, allowing for predominant school-based, broad vocational education and subsequent on-the-job training	Australia, Canada, Japan, U.S.
III: Varied	Varied relationship between educational system and labour market, with close matching confined to apprenticeship or specialised VET and loose matching related to predominant full-time education and qualification frameworks across all sections	England, Estonia, Finland, France, Greece, Norway, Portugal, Scotland, Spain, Sweden

Source: Decy and Tesaring, 2002

Greinert (1997) differentiates between three models of policy steering: The liberal-market economy model, the etatistic-bureaucratic model, and the dual-corporatist model. The dual VET-system in Germany is an example of a corporatist-controlled model, with the social parties mainly determining the qualitative standards for firm-based training and pre-structuring standards for school-based training. France can be regarded an example for the state-controlled, school-based model, with the ministry of education determining the standards, whereas social parties merely have a consultative function in the process of structuring vocational qualifications. Finally, the "National Vocational Qualifications" in England represent the market-driven model with the state recognising a large the number of national qualification standards provided by industry to increase transparency of educational programmes (Koch & Reuling, 1998).

Differentiations of governance types represent variations or combinations of the three basic models depicted above. Lassnigg (2000) provides a heuristic framework, which starts by breaking down the scope of coordination processes in a stylised manner along the following lines:

- Types of *players* involved (e.g., individuals, communities, organisations, levels of government, corporative actors, international bodies and organisations);
- in the affected *sectors* (e.g., sectors of education, employment, labour market, civil society);
- at different levels of society (micro, meso, macro);
- who handle the various coordination and steering *tasks* (e.g., coordination of demand and supply for the education and training market, coordination of the teaching/learning process, coordination of demand and supply on the labour market (see *Figure 3.3*), and
- through numerous possible *interactions and mechanisms* (bureaucracy, market, associations, networks).

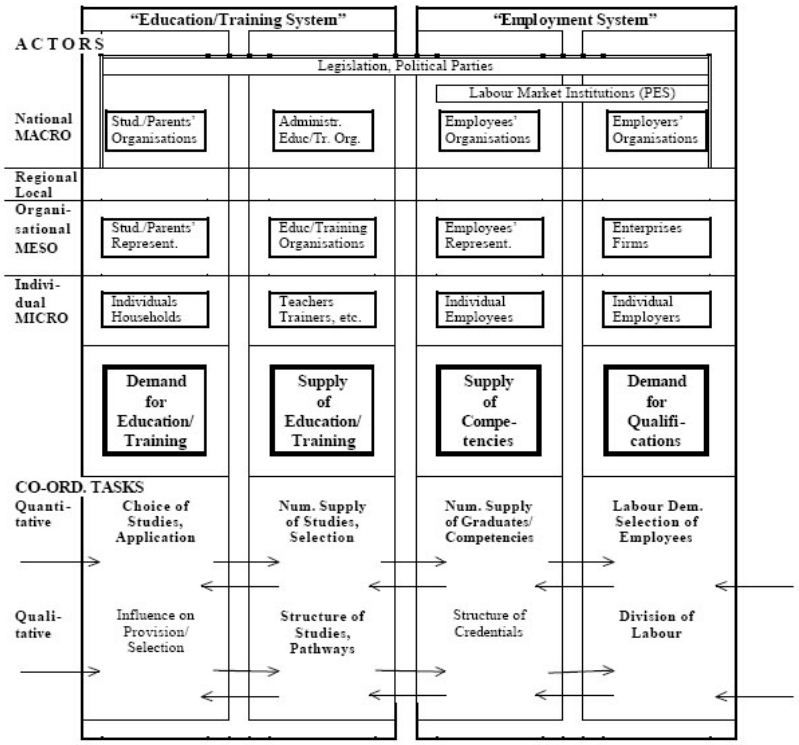
Coordination procedures refers to interactions and relationships between the players, which aim at striking a balance between the two poles of the coordination system: The demand for training and the demand for manpower. In schematic form, the coordination procedures can be distinguished in terms of a quantitative and qualitative structural dimension, so that certain chain reactions are set off, triggered in particular by the two vertices—in *quantitative* terms we are talking about transitions, often in the form of decisions regarding selection or capacity. In this regard, macro-structural data of institutional outputs provide references for successful coordination and steering of VET-systems, such as completion and dropout rates[14] in initial VET programmes, and organisation of transitions to the labour market, such as youth unemployment rates. Moreover, governance types provide a solid basis for the selection of macro-structural classifycation of countries involved in a PISA-VET. In *qualitative* terms this refers to aspects of establishing and changing the structure and profile of training courses or jobs, or "shaping" the profile, as it is known in vocational sociology.

Figure 3.3 gives a stylised picture of the different kinds of actors and mechanisms in the coordination system. The education system and the employment system, each of them including actors at the demand and the supply side (those types of actors constituting the columns in the scheme), are distinguished from one another. The different types of actors can also be distinguished at various levels of aggregation (constituting the rows in the scheme): From the individual (micro) level to the organisational and regional (meso) levels to the national and

14 Completion rates represent the present output of the educational system—that is, the percenttage rate of students, who have successfully completed initial VET in relation to the total population of young people of a typical age for completing initial VET. Completion rates are mainly output-indicators for whether or not educational systems are able to supply the quailfied workers required by the labour market (OECD, 2003). Dropout rates indicate disparities between institutional offers and individual requirements and capacities, such as between occupational offers and individual wishes, existing VET-programmes and individual wishes, or VET-programmes and individual education conditions or preconditions. This refers to problems of transition form compulsory education to VET.

increasingly supranational (macro) levels. The framework provides a basis for international comparative studies.

Figure 3.3: Stylized actors and co-ordination tasks



Source: Lassnigg, 2000

Standards and norms

Institution-based mandatory standards for developing education/training quality can consist of *input standards* for single quality indicators, such as standards regarding organisation and contents/curricula of education/training (see below); action conditions of teaching staff, such as required qualifications, continuing education, collective agreements on working conditions, and behavioural norms; allocation of resources; and/or *output standards* for learning outcomes and examination requirements. Public quality control at a systemic level must be differentiated from internal institutional methods of quality assurance (not illustrated further in this report).

The central function of public quality control is to provide confidence in the quality of VET:

> "Young people have to be sure that the provided VET programmes correspond at least with public quality norms, employers have to trust in the significance of VET qualifications, which is in turn an essential precondition for efficient initial labour markets. Moreover, educational institutions must be confident that formal qualifications and the corresponding authorisations are associated with the appropriate competencies". (Koch & Reuling, 1998, p.7).

With regard to a systematic understanding of the complex and multi-faceted array of VET contents, Lassnigg (2000) distinguishes the following levels:

- First, the way in which the various element of qualification and contents of learning are conceptualised: The most important differentiation refers to traditional distinctions between school-related and VET-related elements of qualifications (competencies of general education versus competencies of vocational education; see *Chapter 2*). Moreover, the differentiation between general and vocational contents can be based on a decoupling model of general and vocational education or an integrated model (Lassnigg, 2000). A third differentiation refers to the relation of theoretical and practical subjects in VET. The outcomes of various weightings can be measured at an institutional level.
- Second, the structure of complex and comprehensive education and training pathways in terms of possible sequencing in flexible, modularised units.
- Third, the relationship between educational processes (curricula) and qualifications (certificates) and the corresponding standards, which can be aligned or separated from each other.
- Fourth, the link between vertical and horizontal educational tracks within educational systems, and their relationship to compulsory- and higher education.

Particular emphasis is put on the standardisation of instruction, which refers to differences in the qualification requirements of the teaching staff and continuing professional development. In this regard, it is assumed that teachers' initial and continuing professional development corresponds to their quality of instruction (OECD, 2003).

Financing

Educational expenditures can be considered either as a valuation of education by suppliers, or as a financial burden for different groups of actors in VET. Expenditure on education per student relative to GDP per capita is a spending measure that takes OECD countries' relative wealth into account. Since education is universal at lower levels, spending on education per student relative to GDP per capita at lower educational levels can be interpreted as resources spent on young people relative to a country's ability to pay. At higher levels of education, this measure is affected by the combination of factors of national income, spending, and enrolment rates. Differences in educational expenditure, however, cannot be solely ascribed to differences in material resources; it is also related to differences in the relative levels of income. In this regard, teachers' salaries constitute the main portion of the total costs (OECD, 2003). The allocation of financial resources indirectly influences the quality of the learning environments and

learning conditions of VET. Therefore, financial input must be differentiated according to types of expenses and the institutions paying the net costs.

3.2.2 Educational institutions

At the level of educational institutions, indicators relating to the structural and organisational conditions of schools and firms as well as to learning and instructions are more or less pre-formed (Kunter, et al., 2002)

Educational institutions and providers of educational services

Input

Characteristics and Organisation of Educational Institutions and Providers of Educational Services

- Utilisation of Resources
- Co-operation of Education Service Providers
- Quality Control

Process

Instructional Setting and the Learning Environment

- Learning Conditions in Schools and Workplaces
- Learning Climate

Characteristics and organisation of educational institutions and providers of educational services

This paragraph illustrates the conditions for instruction at the level of educational institutions. At this level, the required information must be collected on the basis of interviews with school principals or persons in charge of firm-based training.

Utilisation of resources

The allocation of resources refers to the disposition and type of allocation of resources, such as personnel expenditures and material costs of schools and/or firms for VET-purposes. With regard to the allocation of *personnel resources,* the ratio of students to teaching stuff is considered an indicator for indirect measurement of instructional quality (Baumert, et al., 1999). Moreover, material resources play a substantial role in knowledge development, particularly with regard to the use of recent information and communication technologies. Very often the

ratio of students to computers is used as an indicator for measuring the provision to the availability of information and communication technology.[15]

Cooperation of educational service providers

Although varying somewhat, depending on the national system addressed, VET is typically organised as a cooperation between different institutions such as schools, firms, and cross-occupational institutions. The extent and type of cooperation can be considered an indicator of the variety and differentiation of training programmes and whether or not there exists a successfully coordinated implementation of content-related requirements between firms/cross-occupational institutions and respective schools (Kunter, et al., 2002). A close collaboration between VET institutions is based on the assumption that VET-programmes, which function for transitions between different educational institutions, require well-coordinated, didactic and methodical concepts. However, co-operations are very often problematic due to the conflicting objectives of different institutions. From a historical perspective, VET in Germany has always been within the proposed-objectives of vocational qualification and occupational integration/utilisation.

Quality control

Quality control refers to the tools for quality development that schools have at their disposal within their given field of responsibility. One possible differentiation refers to formative and summative assessment and types of evaluation, such as multiple-choice tests and work samples.

Instructional setting and the learning environment

Factors at the macro- and meso-levels are mediated by proximal process factors and are therefore, indirect on VET-processes. They refer to pedagogical and operational tools implemented by educational institutions.

Learning Conditions in Schools and Workplaces

The instructional setting refers to the tools necessary for quality development of students' educational attainment, which schools and firms have at their disposal within their given field of responsibility. Instructional settings denote different pedagogical and organisational settings.

The quality of learning in schools refers to the relationship between implemented pedagogical and organisational settings (e.g., self-organised learning, project-based learning, or teamwork) and to the results of competence measurement. In addition, the type of institution (e.g., library) as well as supplementary activities provided by schools (e.g., special needs education), must be identified and analysed with regard to their utilisation by students. In this regard, it can be assumed that an intensive use of such learning opportunities correlates with high learning achievements.

15 Successful integration, however, requires more than merely providing appropriate resources; the use of recent technologies must be part of the curricula, and consequently, the teaching staff must be trained, to be able to effectively use computers for instruction purposes.

The quality of training in firms refers to characterisations of firm-based learning and instruction differentiated according to working tasks, working environments, and pedagogical and didactical settings (Onstenk, 2003). With regard to the pertinence of learning in workplaces for vocational education and training purposes, Tramm (1992) distinguishes between the following quality dimensions:

- Opportunities for content-related experiences: What kind of content-related experiences, such as definitions, theories, and problems, are accessed when performing the working tasks?
- Complexity of working tasks: To what level of proficiency is the working tasks related? (e.g., aspects of complexity, pre-structuring, problem-orientation, personal responsibility).
- Opportunities for task-related, socially-interactive discussion: What type of opportunities for task-related, socially-interactive discussions with employees or experts are provided?

On the basis of an analysis of learning processes in the workplace, Blockhuis (see Nieuwenhuis, 2004) as well as Franke and Kleinschmitt (1987) present clusters of variables impacting the development of vocational competencies:

- *Problem-orientation and complexity* refers to the cognitive demands of work (complexity of cognitive operations, unfamiliar tasks, number of pre-structured solution steps);
- *variation*—learners/workers should be involved in different aspects of the work process in order to acquire a feeling for the big picture, which is a precondition for being able to effectively transfer knowledge;
- *participation*—to enhance critical reflective work behaviour; it is important that workers/learners perceive that they can participate in workplace decisions, and
- *support*—scaffolding the learning process by foremen or expert colleagues is an important aspect of the learning potential for students at the workplace.

Likewise, Baethge, and Baethge-Kinsky (2004) distinguish between four dimensions of work organisation having a positive impact on learning:

- Opportunities for development at work;
- holistic working tasks;
- communication and social interaction; and
- opportunities for participation in workplace decisions.

After all, the main impact on learning can be ascribed to opportunities for competence development in firms and working tasks based on holistic concepts, which are, in turn, interlinked with the extent of communication and cooperation in workplaces. Opportunities for participation, on the other hand, are less correlated with learning processes. However, differences between these dimensions are rather marginal and therefore, all of them should be taken into consideration.

Learning climate

The learning climate in schools/firms refers to the learning climate at the classroom level and to instructional quality (Kunter, et al., 2002). Many class-

room observations have provided evidence for influences of general school environmental factors on instructional quality in the classroom, and the effects of teacher behaviour and classroom practice on the overall performance of schools (Kunter, et al., 2002, p. 255). Following the PISA study, the following factors are examined: Contentedness with school/firm, contentedness with the instructional quality in specific subjects/at the workplace, teacher-student relationships student-student relationships, and relations to colleagues at work.

3.2.3 The individual participants in learning activities

Within the overall framework of indicators (*Figure 3.2*) at the level of educational participants and learners, we distinguished between individual life- and learning conditions (input), attitudes and behaviours (process), and the quality of individual learning results (outcome). In addition to the measurement of competencies, learners individual background variables will be evaluated using questionnaires. The following paragraph illustrates each sub-category of individual context factors[16].

Individual participation in learning activities

Input
Learners Life- and Learning Conditions
- Socio-Economic Status of Family
- Cultural Capital of Family
- Educational and Occupational Career

Process
Educational Aspiration and -Behaviour
- Information Behaviour and Learning Time
- Educational Aspiration

Output/ Outcome
Acquisition and Utilisation of Competencies
- Qualification Level
- Transition to an Adequate Job
- Occupational Mobility

16 One of the main sources for sub-categories was the German PISA report provided by Kunter, et al., 2002: „PISA – 2000: Dokumentation der Erhebungsinstrumente", published by Max Planck Institute for Human Development.

Individual background variables (input)

This sub-category mainly refers to indicators of correlations between family background and students performance. Educational outcome is influenced by the family background in many different and complex ways.

Learners life- and learning-conditions

The socio-economic status of families has proved to be an important variable for explaining variances in student achievement (Baumert, et al., 2001). PISA has incorporated indicators of the family background into the research design based on Bourdieu and Coleman's concept of social capital (Baumert & Schümer, 2002): A student's social background is determined by the socio-economic status of the parents. However, research showed that this indicator alone is insufficient for explaining students social background:

> "Firstly, economical, political, and social developments of the last decades have changed the structure and function of families, and secondly, due to high mobility between countries, cultural differences within countries have increased" (Kunter, et al., 2002, p. 225).

With regard to these developments and their impact on successful students' performance, aspects of cultural- and social capital of students/apprentices and their parents must be added. The social and cultural capital of students/apprentices are considered as independent concepts. The increasing influence of peer groups and media at a very early age results in parent- and family-independent environments for young people, which have a substantial impact on educational outcomes.

Socio-economic status of the family

The PISA results showed that the parental occupational status, as a measure of socio-economic status, can influence students' aspirations and attitudes (Baumert, et al., 1999). PISA captures aspect of students home backgrounds through information on parents occupations and the activities associated with those occupations in a way that is internationally comparable. The resulting *socio-economic index of occupational status* measures the attributes of occupations that convert a person's education into income; the higher the value on the index, the higher the occupational status of a student's parents. The "International Socio-Economic Index (ISEI)" has been generated on the basis of educational, occupational, and income-related data on 74.000 employees from 16 countries, and was successfully administered in different international comparative studies (Ganzeboom & Treimann, 2003).

Cultural capital of the family

Research consistently shows that the parents level of education and the language spoken at home are significant sources for explaining disparities between students performance (Leschinsky & Mayer, 1999). Bourdieu explains the differences in students achievements with his cultural capital thesis:

> The *high culture* of a society's dominant socioeconomic classes (i.e., its upper class and upper middle class) plays a major role with respect to the reproduction and legitimising of socioeconomic inequality. High culture is constituted by competitive attempts to attain social distinction and sublimated pleasure via activities (e.g., attendance at art exhibits, classical

> plays, and classical music concerts) that involve a high level of abstracttion, intellectualism, and refinement. (Bourdieu, 1983).

In the PISA study, cultural resources subsume symbolic instruments of power, such as artwork or literature, institutionalised forms of potential power, such as educational certificates or titles, and particularely schemata of perception, thinking, and acting in internalised processes of socialisation. The PISA results showed that students' access to possessions and activities related to "classical" culture, such as literature and art, corresponds with educational success. To assess the relationships between cultural possessions and performance, students were asked to indicate whether they had classic literature, books of poetry, and works of art (e.g., paintings) in their homes (Kunter, et al., 2002). For the purpose of a PISA-VET, the parental occupation and their support of the student's development of personal identity must be taken into consideration.

In addition, parental participation in continuing education must be incurporated in a PISA-VET, since a high impact of this factor on students' perceptions and behaviours during and after completing VET can be assumed. The cultural capital of students' family backgrounds represents an important indicator for educational careers and competence developments. In this regard the German PISA consortium emphasised that aspects of cultural capital have hardly ever been quantitatively measured to date (Baumert, et al., 2001).

To identify indicators of students' cultural capital, information regarding the place of birth, home language, familiarity with the migration country, as well as parental education, must be identified. (Kunter, et al., 2002). The difficulties educationally disadvantaged students' with ethnic minority backgrounds and/-or the children of migrants are facing is substantially influenced by the circumstances from which they have come; nationality alone cannot be considered a sufficient criterion for distinguishing between different forms of family socialisation any longer. Therefore, this indicator must be differentiated further according to the students' *place of birth of students, their parents native language, time of school attendance abroad,* and *time/-age of migration* (Federal Ministry of Education and Research, 2006).

In the PISA questionnaire, the cultural capital of students' families was identified with questions regarding possessions of *items related to classical culture*, such as internet access or calculators, and *activities related to classical culture*, such as music, artworks, or literature. The reliability of indicators, however, depends on whether the selection of cultural items and activities corresponds with the respective learning environments. In this regard, reliable indicators for PISA may not be reliable for a PISA-VET; for example, aspects of technical goods and practical issues of everyday life (e.g., self-repair) as well as daily communication issues, may represent more reliable information for a comparison of VET[17].

17 This indicates that aspects of social and cultural capital merge.

Educational and occupational career

With regard to the purpose of measuring students'/-apprentices' competencies, we can assume that they not only represent the outcomes of VET institutions and -processes, but also relate to former school and school-external factors influencing individual educational histories (Lang & Pointinger, 2002).
Indicators referring to individual education histories comprise, in addition to school-based qualifications, certificates of continuing education, competencies of informal learning acquired in work- and social-contexts, and in self-organised learning activities. Issues relating to informal learning have gained importance in educational policy discussions over the past years (Straka, 2003; Baethge & Baethge-Kinsky, 2004)[18].

The international experts at the two workshops also confirmed the impact of aspects of informal learning. In particular with regard to little formalised VET-systems in many countries, informal learning has a significant influence on the development of general and vocational competencies.

Individual spare-time activities and the corresponding integration in social networks represent another important factor: *Analysing spare-time activities of apprentices* refers to various lifestyle concepts, which are related to diverse perceptions of educational requirements and cognitive endeavours (Kunter, et al., 2002). Measurement of spare-time activities, including media use, provides insight into the relationships between spare-time activities and learning achievement. In this regard, a positive correlation between self-determined, creative spare-time activities and high learning achievements can be assumed. In addition, social relationships in peer groups are indicators for the impact of spare-time activities and mutual peer support for learning achievements.

Individual attitudes and behaviours (process)

This indicator refers to the identification of individual learner's characteristics. The relevant learner characteristics for VET (motivation, meta-cognition, and self-efficacy) are discussed in *Chapter 2*. In this paragraph, learner characteristics are supplemented with indicators regarding information behaviour, average learning time, and educational aspiration.

Information behaviour and learning time

Two more indicators are relevant for characterising individual learning processes: Individual information behaviour and learning activities outside schools and/or workplaces. Individual information behaviour provides insights into the scope of knowledge, which is the basis for career choice during the transition from school to VET. In addition, information about the type of informational resources used for knowledge acquisition are provided. The learning time used for internal and external school learning activities is regarded as an additional

18 In the past, international organisations, such as the OECD, have increasingly focused on comparative studies for measuring relationships between formal and informal education; for example, studies for the measurement of adult competencies.

important indicator for learning achievements in schools (Treiber, 1982). It refers to student motivation and individual tendencies for investing time in learning activities. Learning time could be operationalised according to the *average amount of time invested in learning activities* (Fisher & Berliner, 1985).

Educational aspiration

This term refers to students'/apprentices' believes about their future jobs. The beliefs about the relevance of education to future jobs is one of the strongest one's predictors of educational aspiration. Therefore, it plays a substantial role in the choice of career pathways (Baumert, et al., 2001). Students' and/or apprentices' career aspirations are regarded as individual projects in terms of action plans for realising individual objectives. In this regard, two aspects must be considered: First, to what extent students/apprentices stick to their career objectives; and second, their expectation for reaching these objectives (Kunter, et al., 2002).

Acquisition and utilisation of competencies (output/outcome)

In addition to measuring vocational competencies, a PISA-VET must provide information on possibilities for utilising vocational competencies developed in VET in workplaces and labour markets, and with regard to individual educational career developments. Individual outputs, however, are not solely based on VET-processes and personal characteristics; they are rather determined by aspects of the respective occupational field and by organisation-specific conditions. This refers to opportunities during the transition from VET to work, and to the assessment of vocational competencies in the labour market. Thus, the following variables can only be evaluated on the basis of a detailed examination and documentation of systemic context factors (see above).

Qualification level

The qualification level indicates the *formal level of vocational qualifications,* in addition to the *actually* acquired competencies in VET, corresponding to vocational valuations within educational systems, and the corresponding labour market requirements. For a differentiated analysis, the qualification level must be supplemented by information about possible horizontal and vertical combinations of different educational systems in terms of access authorisations.

Transition to adequate jobs

This refers to the quality of career entry after completing VET and the frequency and duration of occupational mobility activities. The quality of career entry can be considered from different perspectives: For example, *educational adequacy* refers to the degree of concordance between the current employment of entrants and their acquired competencies in initial VET. Since the objective is to identify to what extent competencies acquired in VET can be utilised in the labour market, settings of education-work relationships are too narrow; very often they are focused only on workplace requirements (Büchel & Neubäumer, 2001).

The ratio between the acquired VET-competencies and a currently-performed occupation refers to the illustrated relationship between VET-systems and labour

markets indicating the extent to which structures of VET-programmes correspond to occupational structures. In addition, employment status and occupational status (vertical dimension and income) of entrants must be identified. In this regard, data provided by Eurostat refer exclusively to ISCED levels of educational qualifications without differentiating according to general and vocational education (*Chapter 4*).

Occupational mobility

The globalisation of markets and the technological revolution bring about drastic and rapid changes in the relative importance of individual sectors and occupations existing in a particular economy. This means that more and more people will have to adapt to changes of jobs or careers, almost certainly requiring different skills during their working life. Occupational mobility is characterised by two aspects: The perceived discontinuity in terms of occupational-, organisational-, and employment transitions, which represents a challenge for developing individual competencies, and the individual concern for dealing with occupational mobility. Both aspects, perceived discontinuity and mobility activities, must be assessed in a PISA-VET.

3.3 DATA AVAILABILITY

At a national and an international level, empirical research, as a basis for the measurement of indicators and/or identification of a rough typology of educational levels (upper secondary VET/ISCED 3A-C), is insufficiently available, and even at a higher aggregation level, appropriate data are not provided for every country. This is due to the fact that VET (in Europe and OECD countries) is much more diverse compared to other areas of education, and that research in VET is less developed than in other educational areas. Yet the variety of VET-definitions in various countries tends to be fundamentally problematic for cross-country comparisons. This also explains the existing differences between national statistics and terminologies in various countries (*Chapter 4*).

Data are missing particularly in Anglo-Saxon countries, such as Australia or Great Britain. This might be due to the modularised VET systems in these countries, which do not allow a clear separation between general and vocational tracks of education and the respective, mediating institutions. Additional problems result from large variations in the duration of post-secondary educational programmes and the corresponding difficulty of identifying a typical completion age for VET (NCVER, 2003).

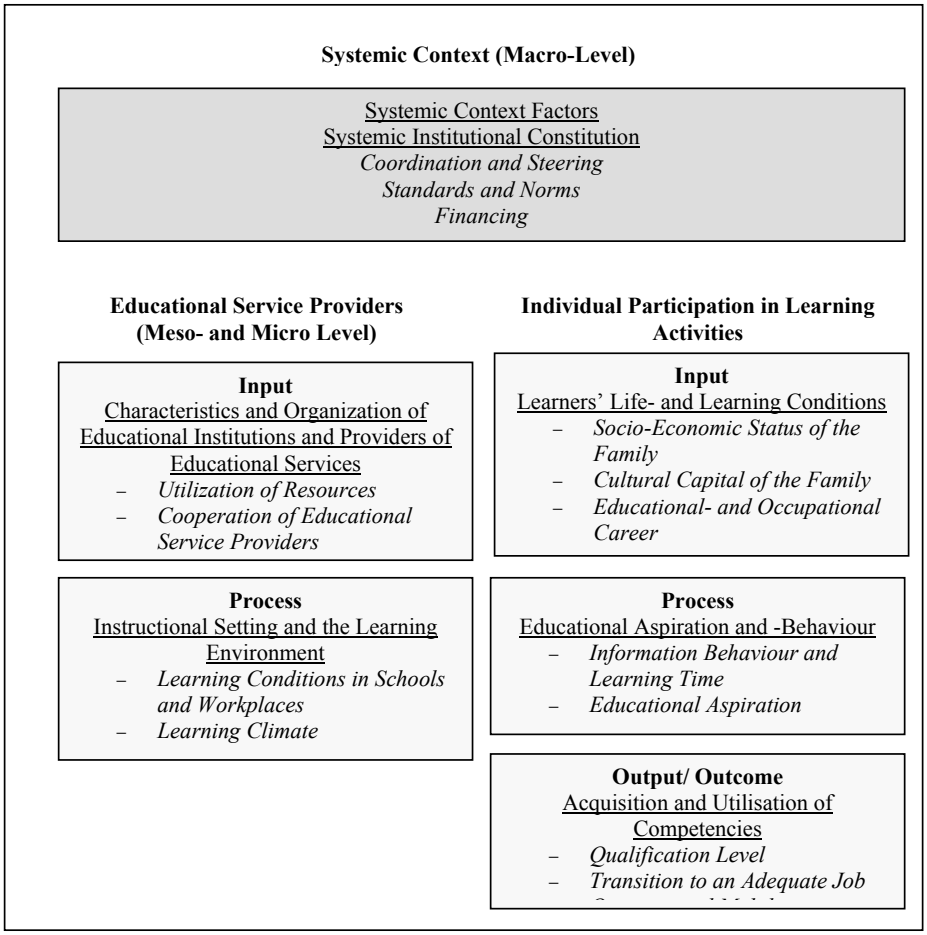
Therefore, we suggest developing measurement instruments for the illustrated indicators, ourselves. The corresponding problems and possible solutions with this approach for sampling are discussed in *Chapter 4*.

The presented bundle of indicators, the educational and employment systems, taken each as a whole, educational institutions, and individual educational participants, must be measured in different ways:

- Context factors must be identified on the basis of secondary data analysis, possibly complemented by interviews with educational experts in the respective countries.
- Institutional input conditions must be identified through interviews with school princepals or persons in charge of firm-based training.
- Institutional process conditions must be measured on the basis of institution analysis and apprentice surveys. This refers to collecting individual information on class size, school/work climate, or instructional programmes, as examples. However, information from apprentices only represents subjective perceptions. On the other hand, research has proven that it is possible to identify reliable information with an impact on learning and training behaviour by evaluating data of entire classes or single education institutions, rather than that from individuals (Kunter, et al., 2002). Thus, although apprentices' questionnaires only provide individual data, they are significant as aggregates at the school or educational institution level.
- Individual conditions must also be identified on the basis of apprentice questionnaires in the context of competence measurement.

Figure 3.4 provides an overview of the most relevant institutional and individual context factors to be identified in a PISA-VET:

Figure 3.4: Measurement plan for institutional and individual context factors



4. THE PROBLEM OF VERTICAL AND HORIZONTAL COMPARABILITY OF VOCATIONAL EDUCATION AND TRAINING STRUCTURES: RESEARCH DESIGN AND SAMPLE CONSTRUCTION

Compared to school-based comparative studies (e.g., PISA, TIMMS), a comparison of VET is different regarding institutional, as well as content-related aspects. Compulsory education can refer to a common curriculum and comparable age groups despite the differences in micro-structural aspects of learning. In contrast, VET is very divergent with regard to contents, given the large diversity of occupational fields and the institutional heterogeneity within and between societies.

Accordingly, the feasibility study must resolve a number of different problems. The complexity of a possible large-scale assessment can be characterised as a two-fold problem of comparison:

- *Vertical* comparability refers to the educational level and can be determined by institution (e.g., tertiary level (academic tack), secondary level) or by duration (e.g., number of years) and/or age.
- *Horizontal* comparability refers to differences in curricula and occupational fields and the problem of defining comparable vocational tracks.

For the purpose of linking the results of individual competence assessment to the micro- and macro-structural context factors, both aspects of comparability are essential and must be taken into account. Moreover, aspects of occupational contents, proficiency levels, and age structure have an impact on *every* competence dimension that will be measured. Research in developmental psychology shows that generic competencies, such meta-cognition, varies according to age (Hasselhorn, 2000).

From a methodological point of view, a comparison of national samples must be based on about the same educational level and content (vertical and horizontal comparability). Vertical comparability is problematic since it refers to national traditional backgrounds of educational systems. Horizontal comparability, on the other hand, refers to national, specific, occupational profiles and the corresponding curricula within the existing structures of labour organisation (Heikkinnen, 2001). In the following chapters, problems of comparability and the corresponding consequences for sampling will be discussed. Finally, possible solutions will be suggested.

4.1 VERTICAL AND HORIZONTAL COMPARABILITY AND THE PROBLEM OF DEFINING INITIAL VOCATIONAL EDUCATION AND TRAINING

4.1.1.Vertical comparability

The research objective of the feasibility study refers to an international comparative study of initial vocational education and training. This subsumes that the focus of the comparison is on the quality of *formal* learning processes in the domain of *initial* VET. This includes institutionalised educational processes, which are

- normally completed within upper secondary education (ISCED 3a–c);
- mainly focused on the development of employability skills and labour market competencies; and
- related to qualifications for entering the labour market.

From an international as well as a European point of view, VET can be characterised at a systemic level as highly dynamic, which makes the idea of *one* entity with a consistent structure very difficult. This refers to the entire VET, as well as to what is called initial VET. Many institutional definitions of VET neglect important aspects of initial education, such as during compulsory education (Descy & Tessaring, 2002). This refers to aspects of demarcation between compulsory education, vocational preparation programs, and continuing education.

In Italy for example, vocational preparation programmes do not belong to the VET-system. Countries with little-differentiated systems of initial VET (e.g., Australia) very often do not provide clearly-separable educational tracks. In some countries (e.g., France, Finland, Italy), some initial VET-programmes are considered independent programmes for preventing dropouts or providing vocational orientation (Beathge, et. al., 2003). These developments can be considered one of the main institutional changes in national educational systems in the past years (Descy & Tessaring, 2002).

With regard to the research design, an exclusion of these programs is not possible since they are independent constituents of initial VET-programmes in many countries. Therefore, it can be assumed that a considerable percentage of one age-cohort is enrolled in these programs, developing the same medium-level competencies as in, for example, the German dual system.

With regard to the research design, the problem of defining initial VET—classifying different programs for vocational preparation or transition from compulsory education to VET—is connected to the problem of disadvantaged students; this subsumes all aspects of successfully entering and completing educational programmes (e.g., social background, gender, cultural background) currently discussed in national and international VET research. It was left open whether or not these criteria (e.g., young people without vocational qualifica-

tions, women/young people with immigration backgrounds) should be included for stratification in the sample.

The problem was initially discussed at the first international workshop (October 18–20, 2004). With regard to the lack of data in this field and the complexity of the comparative study, the majority of international experts suggested to omit this aspect in the research design. This refers to the main problem of sample construction: For every participating country, a relevant extract of VET-programmes, which is comparable regarding competence levels and occupational fields, has to be identified. Data analysis must deal with the problem of insufficient availability of updated and internationally-comparable, classified data, as well as with culturally-divergent definitions of educational qualifications and occupational tasks.

From an international point of view, the identification of comparable VET-programmes is based on very few and little updated data, mainly provided by the following European institutions: OECD, UEO, or EUROSTAT/CEDEFOP. The most beneficial sources for identifying quantitatively-relevant programmes of initial VET were published by EUROSTAT and CEDEFOP—a handbook of different educational programs in different European countries—based on data from 1995-1996 and updated in 2000. It consists of abstracts containing information about different educational programmes (title), the type of programme (vocational preparation, general education, vocational education and training), ISCED level, and typical entrance age. Moreover, information regarding the type of labour market qualification (generic, subject specific) and the number of participants is provided. Unfortunately, the same information is not available for every country and data on the number of beginner participants entering educationnal programmes are entirely missing.

Based upon data analysis and discussions with international experts, we came to the conclusion that ISCED is not a very suitable classification for identifying initial VET-programmes addressed to a medium level of proficiency, since it relates to levels of general education, rather than VET.

According to ISCED, the corresponding educational programmes in different countries are located at levels 3 and 4 (with exceptions found on level 5). Provided that ISCED level 3 represents the lower limit for VET programmes in the sample, certain vocational preparation and short-term programmes[19] would be located below these categories.

The selected occupational profiles, which were presented at the second workshop, proved to be insufficiently representative to account for existing differences and changes in national educational concepts and relevant occupational fields. For example, in the U.S., not only two-year community colleges, but also vocational high schools used to be the main institutions for developing vocational

19 Germany: Basic Vocational Education (Berufsgrundbildungsjahr, einjährige Berufsfachschule); Denmark: Basic Vocational Education (EGU); Switzerland: „Anlehre"; Hungary: Secondary VET Schools; USA: Vocational High Schools; UK: NVQ-level 1 or 2.

competencies at a medium level (Münch, 1999). However, educational experts from the U.S confirmed that there has been a change, at least in certain occupational fields (Banking, Information Technology), towards four-year colleges. Comparable developments have been reported for France; vocational education programmes in the fields of Banking and Information Technology are located within the sector of general education ("Baccalaureat Professionel/Technologue").

Another problem in this regard refers to fact, that formal definitions of occupational fields do not provide a basis for identifying qualitatively-comparable processes of VET, since institutional contexts can vary within formally-identical educational tracks: Firstly, national curricula, in addition to methods of instructtion (e.g., different combinations of theoretical and practical training subjects in schools and firms), can vary due to modularised and institutional structures, without the impact on content structures and competence levels being assessable. This is particularly the case in countries with loose connections between education system and labour market (i.e., between qualification and occupational tasks).

This problem is particularly prevalent the in the U.S., Australia, and the UK. Moreover, in the field of full-time, school-based educational programmes, loose connections between the educational system and the labour market are problematic in the UK[20], Finland, France, Greece, Norway, Sweden, and Spain (Lasonen & Manning, 2001; Hannan, et al., 1997). However, even the results of recent studies can prove wrong in the light of up-to-date developments; educational experts in the Netherlands confirmed that the closed connections between qualifications and work-place requirements in their country, proved to be unstable in a modularised system.

Secondly, with respect to the variance of age for entering initial VET-programmes, considerable differences *between* and *within* countries can be identified. In the majority of countries, compulsory education is completed at the age of 16, which corresponds with the minimum age for entering initial VET programmes. Germany (18 years) and the Netherlands (17 years) are exceptions to that rule by being somewhat higher, whereas other European countries (e.g., Poland, Czech Republic, Portugal, and Belgium), have lower than average entry age (*Table 4.1*).

According to statements made at the second workshop, the age range for entering initial VET programmes has gradually increased. This has likewise been a recent tendency in Germany; according to data provided by the Federal Institute for Vocational Education and Training (BIBB), the average age for entering a dual VET programme was 19 in 2003. The span of quantitatively-weighted age co-

20 A recent comparative study of competencies developed in the German programme of „Industriekaufmann" and the English/Welsh Advanced Business qualification (GNVQ, level 3), showed that some competencies of the English participants were at a lower level than those of the Germans (Fulst-Bleil & Ebner 2005).

horts ranged from below 16 up to 20-year olds (http://bibb.skygate.de/Z/B/30/99600000.pdf). The Danish experts have reported related tendencies in their country. Moreover, some VET programmes require higher entrance ages than others—nursing education as an example, requires a minimum age of 18 to begin in most countries.

Table 4.1: Typical ages at the end of compulsory education, as well as after completing educational programmes at ISCED level 3 (percentage)

	Completing Compulsory Education	**Completing ISCED Level 3**
Australia	15	17
Czech Republic	15	17–18
Denmark	16	19–20
Germany	18	19
Great Britain	16	*
Spain	16	17
France	16	17–21
Netherlands	17	17–20
Norway	16	16–19
Poland	15	18–20
Portugal	15	18–20
Sweden	16	19
Switzerland	16	17–20
U.S.	16	not specified

* not declared

Source: OECD, 2004

Finally, more and more VET programmes allow for external assessment outside of standardised (regarding content and time), formal VET-programmes, which are only included to some extent in educational statistics. In many European and non-European countries, there is a large number of entrants at a lower secondary level (ISCED level 2 and below)—that is, young people entering the labour market without general or vocational qualifications at an upper secondary level.

In 1997, the rate of entrants into the European labour market (EU-member countries) on ISED levels 0–2 amounted to 32 percent (*Table 4.2*). Countries such as Finland, Austria, Germany, France, and Sweden contributed to the lowest percentages (14-19 percent), whereas Great Britain, Spain, Italy, and Portugal represented the highest rates (38–62 percent). The American workshop-participants reported that in the U.S., direct transitions from compulsory education to the labour market is a widespread phenomenon.

Table 4.2: Proportion of entrants in the European labour market in 1997, according to ISCED levels (percentage)

Country	ISCED 0-2	ISCED 3	ISCED 5-7	Total
Belgium	21	35	44	100
Denmark	20	52	28	100
Germany	17	56	27	100
Greece	25	54	21	100
Spain	45	15	39	100
France	18	44	38	100
Ireland	25	38	37	100
Italy	50	42	7	100
Luxemburg	30	35	35	100
Netherlands	26	40	34	100
Austria	15	76	9	100
Portugal	62	17	21	100
Finland	14	55	31	100
Sweden	19	62	19	100
Great Britain	38	36	26	100
EU-15 Total	**32**	**42**	**44**	**100**

Source: European Union, 2001, p. 49

Despite the lack of representative data, we can conclude that the mentioned rates[21] only denote the minimum level of the whole extent of young people entering the labour market without any further qualification after compulsory education. Since there are no representative and comparable recent data provided by official statistics, we have estimated the degree of institutionalisation of initial VET programmes in the EU member states. As shown in *Table 4.2* there are immense discrepancies between the countries. For example, in Greece, Ireland, Spain, and Portugal (and possibly Italy), entering initial VET programmes immediately after compulsory education is playing a marginal role (around 20 percent). On the other hand, in countries like Belgium, Germany, the Netherlands, Austria, the Czech Republic, and Hungary, it is the predominant track (between 60–78 percent).

The main reason for collecting data on the transition from compulsory education to the labour market is that, particularly in Anglo-Saxon countries (U.S., Great Britain, Australia), initial VET is an integrated part of post-secondary (higher) education or work processes in firms. Very often the workplace has the

21 The rate refers to entrants into the labour market, not to unemployed looking for a job. The rate of unemployed without professional qualification was above average in the year 2000 in almost every EU-member state (except Portugal; Fischer & Siebern-Thomas, 2004). Therefore, we can conclude that this group is rather underrepresented among entrants.

function of vocational orientation and provides a financial basis for educational processes in the future. In this regard, informal learning processes *on the job* are very often the basis for formal learning processes *off the job*.

The American participants at the second workshop emphasised that in a number of occupational fields (e.g., commerce), a very low level of competencies is required for many tasks. These tasks/jobs have a supportive function for young people (16–18 years) in vocational orientation processes during the transition from school to work. It is very common to enter the labour market after completing compulsory education and work for some time before entering an educational programme; in the field of healthcare, for example, it is very common to start working as a cleaning woman in a hospital, continue with an educational programme for nurses at a very low level and work for some time before entering an educational programme at a higher level.

The workshop participants from Great Britain confirmed this tendency for their country; the function of a number of non-professional jobs in the UK is simply to provide a financial basis for living and time for vocational orientation (including participation in continuing educational programmes) after completing compulsory education—"the function of the job is to be a student". Moreover, vocational education and training systems are planning to assess and certify informally-acquired competencies without participation in formal educational programmes.

John Stevenson (Australia) pointed out, that the main preparation for VET-programmes in Australia is taking place in post-secondary institutions, such as TAFE Colleges. However, an increasing percentage of participants is older than 18 and working in a job. As in Great Britain, formally and informally-acquired competencies are considered equal.

Table 4.3: Education participation of 17-year olds in selected countries, according to institutions (in percentage) – Estimation*

Country	No Education Participation	Upper Secondary Level (General Education)	Upper Secondary Level (Initial VET)**	Total
Belgium	0	30	70	100
Denmark	17	39	44	100
Germany	7	34	59	100
Greece	30	42	28	100
Spain	18	51	31	100
France	11	39	50	100
Ireland	28	72 (52)	0 (20)	100
Italy	21	58 (28)***	21 (51)	100
Netherlands	17	26	57	100
Austria	22	22	56	100
Portugal	30	50	20	100
Finland	6	40	54	100
Sweden	4t	48	48	100
Great Britain	26	21	53***	100
Norway	7	39	54	100
Czech Republic	2	19	79	100
Hungary	15	74	11	100

* The estimation is based on the following specification: Proportion of upper secondary-level students/apprentices in different educational institutions and general education participation rate of 17–year olds. For this purpose, data from the European institutions, OECD and EU, have been linked to each other. In the majority of countries, the participation rate of 17–year olds can be regarded as representative for completing compulsory education and for participation in initial VET programmes (no continuing educational programmes).

** Without VET-preparation programmes.

*** In Italy, it is debated whether technical colleges/polytechnics ("Institute Technici") belong to institutions of general or vocational education.

**** Including VET-preparation programmes.

Source: Own estimations based on OECD, 2004

4.2.1 Horizontal comparability

At the starting point of this feasibility study, a convergence towards the most important occupational fields appeared to be a convincing solution from a German perspective. However, a comparison of employability structures from an international perspective, taking into consideration the ongoing dynamic Euro-

pean developments, is more complex with regard to the following aspects (*Table 4.4*): The service sector is certainly expanding most rapidly in the EU (+12 percent increase in the number of jobs between 1997 and 2002) followed by the industrial sectors (manufacturing and construction: increase of +3 percent). The primary sector decreased by 13 percent. Despite the high proportions of agricultural structures relative to the total of occupational fields, particularly in the new EU member states (Romania about 33 percent, Poland 12 percent), recent dynamic developments entail restricting the comparison to the service and manufacturing sectors.

The construction sector (metal industry), with an employment rate of 8 percent in Europe, represents an even larger, and on balance, still growing proportion of occupational fields. Comparable developments can be observed in the high technology and manufacturing industries. Accordingly, one manufacturing/industrial occupational field of industry and construction should be included in the sample. Within the service sector, wholesale and retail sales (including maintenance and car-repair), represent one of the largest traditional occupational fields in Europe (employment rate of 15 percent), which notably increased between 1997 and 2002. In the fields of healthcare, employment rates have amounted to 10 percent (with some countries showing even higher variance of the overall employment quota—Portugal 5 percent, Denmark 18 percent). Finally, the occupational field of high technology in the service sector represents a small but rapidly-growing occupational field with an average employment rate of 4 percent.

After all we can conclude, that the most important industrial/technical, commercial/administrative, and personal occupational fields have to be included in the sample.

This allows a first approximation towards a selection of occupational fields:

- Industrial/technical occupations in industry and trade;
- commercial and commercial/administrative occupations in commerce and other services;
- healthcare occupations in the field of personal services;
- information and communication technology in the field of information/ technical services.

However, there are still problems remaining: Firstly, there is neither an internationally-consistent competence-based concept of the structure and classification of initial VET programmes/occupational fields, nor a reporting to rely on (Lisbon-to-Copenhagen-to-Maastrich Consortium Partners, 2004). This does not imply that classification systems are completely missing, but what is needed are basic data in a usable format. At an international level alongside ISCO 88 (for occupational fields), an occupational classification of discrete fields of training has been developed on behalf of CEDEFOP (Anderson & Olsson, 1999). However, neither of the two classifications Furtems has been accepted as part of the ongoing reporting at the European level.

A reporting comparable to the German vocational education and training report (Berufsbildungsbericht) is virtually unavailable in the remaining countries (with a few exceptions, such as Australia). Moreover, there is a lack of adequate, updated registered data. Another problem refers to the variety of broad and limited concepts of VET with very detailed or undefined contents of VET-programmes due to different cultural traditions:

- Compared to the German nursing education, the 2-year health care programmes in the U.S. include both: Qualifications for an assistant nursing degree (LVN), as well as for medical technician occupations (RN) (Center for Education Statistics (NCES), 1998; US Bureau of Labor Statistics, 2005).
- With regard to commercial administrative occupations (office positions), in some countries, there is no equivalent VET programme in this field (e.g., Norway). However, even where such a programme exist, very often there is little overlap between the existing programmes (e.g., in the U.S.), which only include communal administrative occupations, contrary to other countries where they do not exist at all.
- In Australia, VET programmes are classified according to occupational tasks, which is, however, not equivalent to the contents of Training Package[22] modules.

Finally, although there are internationally-comparable occupational tasks/fields of training, very often they only represent a small part of the entire occupational field, or even none at all. In drawing a conclusion from the literature analysis and the discussions in the workshops, we can highlight three such occupational fields—car repair, banking, and registered nursing.

22 A Training Package is the grouping together of the training components designed to assist in achieving the competencies for a specific industry. Units of competency are packaged together which, when combined at various levels, can form qualifications (for example, Certificate, Diploma and the like) (National Training Information Service, 2005).

Table 4.4: Employment structure according to occupational and economic sectors in selected countries, 2002 (percentage)

Country **Occupational/Economical Sector**	**DK**	**G**	**E**	**F**	**NL**	**P**	**S**	**UK**	**CZ**	**RO**	**NO**	**CH**	**US**	Sectoral Development of Employment in the EU (1997 to 2002) (Percentage)
Agriculture, Forestry, and Fishery Sectors	4	2	6	4	3	12	2	1	5	32	4	4	3	- 13
Manufacturing Industry	22	28	29	22	20	34	23	19	40	32	22	21	21	+ 3
High-Technology Sectors of Manufacturing Industry	6	12	5	7	5	4	7	7	9	6	*			*
Construction Industry	7	8	12	7	7	13	6	7	9	5	*			*
Service Sector	74	70	65	74	77	54	75	80	55	36	74	76	76	12
Wholesale and Retail Sales, Maintenance and Car Repair Sectors	15	14	16	13	16	15	12	15	13	10	*			*
Service and Healthcare Sector	18	10	6	11	15	5	19	11	6	4	*			*
High Technology in the Service Sector	5	3	3	4	4	2	5	5	3	2	*			*
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	

Source: European Union, 2001

* not declared

4.2 CONSEQUENCES FOR THE SAMPLE CONSTRUCTION

Taking into consideration the
- entire variance of initial VET programmes,
- full variety (regarding level and age) of institutional arrangements, and
- considerable differences in processes of transition from compulsory education to the labour market (e.g., combinations of formal and informal periods of training and working)

would overstrain the design of the intended comparative study.

We are aware of the fundamental problem of heterogeneity in VET which entails that a representative sample, covering entire VET programmes and occupational fields, cannot be the objective. With regard to concepts for measurement and observation, pragmatic and methodical considerations (development efforts for measurement methods, limited participation in atypical VET programmes) strongly suggest using a stratified sample.

The aim is to identify a relevant, in terms of the most important employability segments and institutional arrangements, as well as comparable (medium competence level, similar contents and age) extract of initial VET out of the entire group of processes related to vocational education and labour market transitions.

With regard to the before-mentioned problems of international sampling, we can conclude that they are resolvable; however, not within a single homogenous sample and not without tradeoffs in weighting comparison criteria. National differences of institutional interrelations between education and employment systems on the one hand, and the impact of non-standardised vocational biographies, on the other hand, requires consideration of the following aspects:
- Incorporating non-standardised vocational biographies is only feasible on the basis of retrospective measurement of educational experiences within a cross-sectional study. The sampling has to be based on the employed population—with criteria including age and occupational tasks according to level and breadth—to reconstruct vocational biographies. Representativeness in this approach is limited due to institutional aspects of education; for example, the "fuzziness" of retrospective measurements of level and content of different VET-programmes.
- Provided that VET-programmes, -levels, and -institutions highly correlate with occupational fields and levels (e.g., in Germany based on the „Berufsprinzip"), and at the same time, cover most occupational tasks at a medium level, vertical (based on occupational tasks), or horizontal (based on VET programmes) sampling does not seem to make a big difference to the outcome. However, even in Germany, this correlation does not apply in every occupational field, and in Anglo-Saxon countries there are very few segments of this type. With regard to a panel study starting from educational institutions, a sampling based on the structure of occupational tasks is recommended. Nevertheless, for pragmatic reasons, a sample based on the structures of VET

programmes has to be taken into consideration. The differences between the two options are due to the remaining efforts in preparing the sampling and evaluating the weight of each sample in different segments. Both options are connected to the construction of vocational sub-samples.

First option: Broad sample based on occupational tasks

This method is based on the occupational tasks to be performed at a particular age (e.g., 24 years). The sample is defined on the basis of the „European Labour Force Survey".

The ISCO 88 classification (International Standard Classification of Occupations) provides a basis for approaching a comparable sample within the most important occupational fields (industrial and manufacturing, commercial and administrative, social and healthcare, information and technology). ISCO 88 has been implemented within the "European Labour Force Survey" and indicates not only the level, but also the contents of occupational tasks. In this way, aggregates of homogeneous occupational contents and levels, which are representative (at least in Germany) of a considerable number of individuals in the same age, could be identified. (*Table 4.5*): With the exception of skilled data processing workers (constituting a proportion of 1 percent of information technology occupations), the sample represents the most important occupational field of employees, comprising 47 percent of all young adults in this age.

The information value of ISCO data can be considered a solid basis for the level and content of occupational tasks, even though occupational fields are not always explicitly distinguishable. Thus, an adjustment procedure for the purpose of a representative identification of bundles of tasks must be conducted. At the second workshop, the results of the discussion provided a basis for a classification of comparable sub-samples.

Table 4.5: Selected occupational tasks, ISCO codes, and employment rates in different occupational fields at the age of 15 to 25 years

Segment	**Occupational Field**	**ISCO Code of Specific Occupational Tasks**	**Employment Rate in Different Occupational Fields (15 to 25 Years) (in %)**
Industrial/Technical Occupations (Industry and Trade)	Metal and Electrical Work	72	15,0
Commercial/Administrative Occupations	Commercial and Office Work	522, 523, 3416, 3419, 411, 343	9,7 11,9
Social, Education and Healthcare Occupations	Nursing	322, 323	8,8
Information Technology Occupation	Skilled Data Processing Workers	312	1,0
Total			100,0

The adjustment procedure is based on an analysis of work requirements according to the main functions—production, distribution, personal services—and focus of contents. The developed job profiles could be used in an iterative procedure for sampling job profiles provide reference points for identifying national tasks corresponding to the respective profiles. This allows adequate classification of the tasks that are related to adjoined levels or different bundles of tasks according to ISCO 88. Problems of classification help to identify clearly-distinguishable occupational profiles at an international level as a basis for developing measurement tools.

The bundles of tasks could be combined with national data on educational background variables derived from national labour force surveys, to identify the (quantitatively) most relevant VET-programmes and levels in the corresponding occupational field. The identified VET-programmes are the foundation for institution-based sampling. However, this method is particularly problematic with regard to Anglo-Saxon countries since broad sampling includes a variety of institutional arrangements that have to be accounted for. A stratified sample would have to account for different institutions and regional differences, as well as for the identification of homogeneous age groups.

Second option: Narrow sample based on widespread and well-defined VET-programmes

This method is based on VET-programmes with comparable curriculum contents preparing young people between 16 and 20 years to enter the labour market. This method consists of several steps: Firstly, we developed two schemata—one consisting of a pre-selection of designated VET-programmes (*Table 4.5*) and another based on EUROSTAT/CEDEFOP data, supplemented by information on relevant VET-programmes[23] from non-European countries. Secondly, we discussed the selection with international education experts at the workshops, and on the basis of this discourse, we developed a proposal for a sample, which incorporates the main occupational fields, and simultaneously provides a basis for comparing level and age (*Table 4.6*).

23 Programme refers to separable institutional fields of VET determined by their affiliation with a particular sub-system of VET (e.g., dual system), characteristics of possible qualifications, and curricula.

Table 4.6: VET programme, corresponding ISCED levels, age range at entrance

Occupational Field	VET Programme	ISCED Levels (Internationally)	Age Range at Entrance
Metal and Electrical Occupations	Car Mechatronic	3	16–20
	Mechanic (industry)	3	16–20
	Electrician (industry)	3	16–20
Commercial/ Administrative Occupations	Banker	3–5	17–20
Healthcare Occupations	Nursing	4–5	18–20
Information and Communication Technology Occupations	Skilled Computer Scientist	3–5	17–20

The field of trade is represented by the European occupational profile of car-mechatronics, of which a similarly-structured profile can be found in the U.S. and Australia. The field of industrial metal occupations is represented by electricians and mechanics (industry), although the latter may possibly be replaced by tool makers at some point)[24]. Contrary to Germany, vocational qualifications of bankers, skilled computer scientists, and nurses are formally located at a higher proficiency level in most other countries. However, with regard to contents of curricula they appear comparable.

The proposed sampling does not provide a solution for every problem, but the discussions at the second workshop showed that remaining uncertainties with regard to the comparability of levels could be solved relatively easy (e.g., nursing). Feedback from national research groups and their statistic agencies will help to resolve some of the remaining problems and improve comparability of an international sample.

Despite all the remaining uncertainties regarding an appropriate age range for entering a VET-programme, there are possibilities for identifying an age-homogeneous overall sample. The relevant age would be 18 to 19 years,—a measuring point at the beginning of VET in this age cohort (e.g., 18 year old beginners).

Due to the specific starting points of each method (broad occupational tasks versus single VET-programme curricula for comparatively-adjusted employment), the probability of identifying additional relevant VET-programmes, covering the same occupational tasks, is relatively low. Therefore, the selection

24 The discussion during the second workshop clearly illustrated the problem of identifying comparable occupational profiles of mechanics (industry) in different countries. In contrast, the more narrowly-defined occupational profile of tool makers (Konstruktionsmechaniker) can be identified, in Anglo-Saxon countries, among others.

method of the second option needs to be checked on the basis of a task analysis of educational backgrounds.

Implementation of the study—sampling, measurement, and observation
The two methods for sampling illustrated above are in principle open with respect to the organisation of the sample and the study. However, basic principles of measurement and analysis limit possible research designs. The limitation is based on correlations between micro- and macro-structural analyses, measurement of vocational competencies, analysis of institutions, and process orientation (causally relating different circumstances in time, in terms of a cause/effect analysis of competence outcomes).

Various aspects regarding the selection between the methods of *household-based* vs. *institution-based* representative sampling, in addition to *cross-sectional* (including retrospective elements) versus a *longitudinal study* (panel), have been discussed extensively in recent literature, as well as in the context of international comparative studies (Kühnel expertise; Schnell, Hill, & Esser, 1989; National Center for Education Statistics, 1998; COM, DELSA, & EDU, 2004; Baumert, et al., 2001). The differences can be characterised as follows:

- *Sampling*—in a *household-based* sample, the target population is selected in a two-step procedure: First, the relevant individuals in a particular VET programme, performing a particular occupational task, are identified. Second, the individuals chosen on the basis of questionnaires are classified according to institutions, whether educational institutions or firms. In an *institution-based* sample, the same two-step procedure is conducted, but present a rather controversial issue; firstly, the relevant institutions are identified, and secondly, individuals are classified according to the VET programmes and occupational fields. If institutions/firms are randomly selected, we get a representative sample of the target population (Kühnel expertise, Annex D).
- *Organisation of the study*—in cross-sectional studies (ex-post-facto designs) (Schnell, Hill & Esser, 1989) competencies are measured at different measuring points, such as at completing VET or after entering the labour market, in different samples together with the corresponding contextual factors. To identify possible causes and effects of measurement outcomes, the approach has to be combined with retrospective methods. At an *individual level,* this would imply the inclusion of socio-cultural and economic background variables as well as biographic incidents. At an *institutional level,* this refers to cross-sectional studies at different measuring points (trend studies in several waves, age-cohort design), in addition to simultaneous analysis of different samples representing various age cohorts in VET-programmes and labour markets; for example, a quasi panel or an institutional cohort design. In contrast, a real panel study in a narrow sense, would measure competencies in sub-samples at the beginning and the end of VET, as well as after the transition to the labour market. Contextual factors would be measured simul-

taneously on the basis of interviews and measurements from VET-institutions or firms) (Kühnel expertise, Annex D).

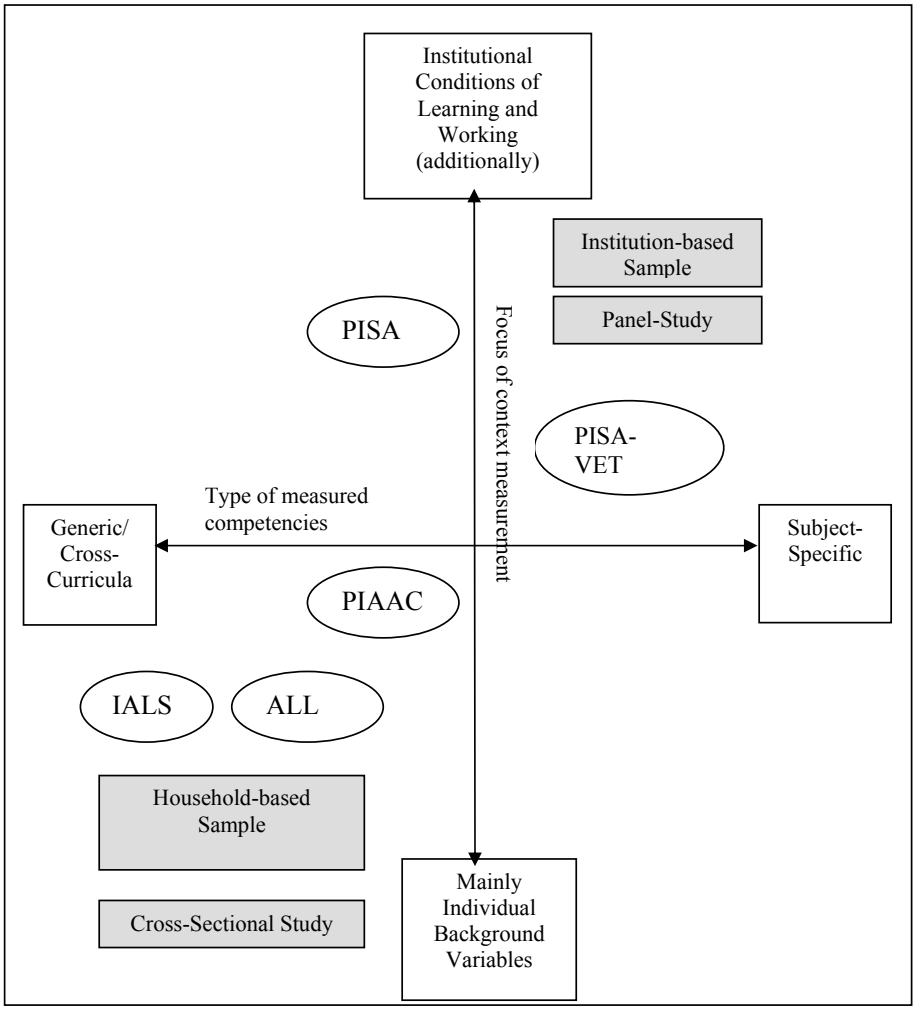
Making a decision regarding an appropriate method is not easy for several reasons: Whereas some approaches have been tested many times in international comparative studies, others have no international results on which to rely on at all. The majority of international comparative studies are combinations of household-based samplings (except PISA) and retrospective cross-sectional designs (e.g., IALS[25], ALL[26] , PIAAC[27]; National Center for Education Statistics, 1998; COM, DELSA & EDU, 2004; Baumert, et al., 2001). This combination is recommended particularly for the purpose of these studies; household-based sampling allows identification of individuals outside typical educational programmes. Moreover, subsequent cross-sectional studies are less complex than repeated measurement (Kühnel expertise, Annex D):

An analysis of the studies reveals that from a methodical point of view, they are limited in their suitability for international comparisons of VET (National Center for Education Statistics, 1998; Baumert, et al., 1999; COM, DELSA & EDU, 2004) because the entire organisation of the study depends upon the kind of incorporated competencies (and the corresponding measurement tools), and the impact of institutional context factors (*Figure 4.7*).

After all, we can conclude, that an institution-based sampling, consisting of different occupational sub-samples and organised as a panel-study, could be a feasible concept for a PISA-VET measurement. Implementing a household-based sampling would imply very high efforts due to the high amount of individuals, not being part of the target population (occupational sub-samples), who would be included. The identification of members of the target population as such would involve extensive efforts. For pragmatic, as well as methodical reasons, institution-based sampling in a panel is the favoured option. It is much easier to identify individuals of the target population in institutions together with the corresponding competencies and contextual factors at different measurement points (e.g., transition from compulsory education to VET and to the labour market). The illustrated approach provides a basis for reliable analysis of short- and long-term causes and effects (Kühnel expertise; Schnell, Hill & Esser, 1989).

25 International Adult Literacy Survey
26 Adult Literacy and Life Skills Survey
27 Programme for the International Assessment of Adult Competencies

Figure 4.7: Sampling in international comparative studies (according to the type of measured competencies and focus of contextual factors)



With regard to a real panel study, the following adverse aspects must be taken into consideration: Repeated measurement of the same population or institutions (depending on the number of measuring points and time intervals) is more complex than cross-sectional measurement. Moreover, panel mortality due to dropouts at a later point in time is a serious problem. However, this can be diminished due to systematic maintenance and an increase of the original sample size. These precautions will certainly not help to reduce the expected sample, specific dropouts of the target population, the loss of very mobile young people,

or the transition to occupational fields not included in the research design[28], which distorts the measurement results. Additional control groups could help to observe this effect (compare Kühnel expertise, Annex D).

After all, the following questions remain open at this point—how can representative occupational sub-samples be determined and what are the consequences for the sample size?

The respective sub-samples will be chosen on the basis of a two-step procedure: First, all relevant educational institutions and possibly firms will be identified; and second, the corresponding participants will be determined. A random sample of institutions and/or firms will assure representativeness of the subsamples. However, using cluster sampling based on entire school grades or groups of individuals in one firm runs the risk of bias with respect to intra-class correlations. In this regard, increasing the sample number on the basis of a certain adjustment coefficient helps to balance representativeness (expertise Kühnel, Annex D).

The exact sample size in each country based on the illustrated approach (subsamples of different VET-programs, institution-based sampling, panel construction) cannot be precisely determined at this point. Nevertheless, a model calculation for Germany will be presented. The calculation is based on the following assumptions:

- An estimated sample size for the third measuring point (simple random sampling with 5 percent accuracy);
- a panel mortality of 50 percent (the average size within an observation period of 6 years); and
- an adjustment coefficient for using clustered samples of 2.9 for schools and 1.8 for firms.

Based on these considerations, the following proposal for sampling can be presented (*Table 4.7*):

28 This refers to the problem of a panel for measuring the development of occupational competencies; some people will transfer into a different occupational field after entering initial VET programmes. However, regardless of considering these individual losses, or measuring their competencies with a different concept, the sample is decreasing.

Table 4.7: Proposal for sampling (model calculation)

	Size of Calculative Basis Sample (Simple Random Sampling, Error Margin 5 percent)	**Enlargement on the Basis of Estimated Panel Mortality (50 percent)**	**Enlargement on the Basis of an Adjustment Coefficient for Clustering (1.08/2.9)**	**Sample Size Panel Study**
Car Mechatronic	400	+ 400	+ 64	= 864
Mechanic (Industry)	400	+ 400	+ 64	= 864
Electrician (Industry)	400	+ 400	+ 64	= 864
Banker	400	+ 400	+ 64	= 864
Nurse	400	+ 400	+ 1520	= 2320
Skilled Computer Scientist	400	+ 400	+ 64	= 864
Total of VET Programmes				**= 6640**

For each of the six VET-programmes a sample of 400 individuals is required for the third measuring point. With regard to the aimed panel duration and expected mortality rate, the samples have to be increased up to 800 individuals. Since intra-class correlation in schools can be assumed to be higher than in firms, a higher adjustment margin in the field of nursing can be expected, which results in a sample of 2.320 individuals. The remaining VET-programmes for car mechatronics, mechanics (industry), electricians (industry), bankers, and skilled computer scientists require a smaller sample of 864 individuals total. Altogether, this proposal amounts to 6.640 individuals (*Table 4.7*) for measurement in the first wave.

5. RESEARCH DESIGN OF PISA-VET: ALTERNATIVE MODELS

5.1 INITIAL SITUATION

The development of possibilities for conducting an international comparative study for VET, as well as an examination of the corresponding conditions and problems of implementation, has been the task of the feasibility study. The objective of a PISA-VET is to form an assessment of competencies developed in VET and to provide an analysis of the micro- and macro-structural context factors influencing the quality of VET. A comparative measurement of competencies could illuminate those outcomes of educational processes, which are not sufficiently represented solely by formal certifications. This allows a systematic analysis of the strengths and weaknesses of different VET-systems. In order to measure competence development and to link the results to the micro- and macro-structural factors, the design should be of dynamic nature.

Contrary to school-based, large-scale assessments (e.g., PISA), a PISA-VET cannot be limited to measuring basic competencies of general education. Assessing competencies in VET entails measuring subject-specific vocational competencies. This is the most difficult aspect and represents the challenge of the feasibility study.

Based on an analysis of the state of the art and the discussions with international experts in the two workshops, three central examination steps have been conducted:

First, an examination of concepts for competence measurement and the development of proposals for VET (*Chapter 2*);

second, an analysis of the micro- and macro-structural conditions, influencing the quality of educational processes, as reference points for outcomes of competence measurement (*Chapter 3*); and

third, an illumination of the problems of international comparative research in VET, and the identification of systemic factors for the research design and sample construction (*Chapter 4*).

After all we came to the conclusion, that an international comparison of VET based on a large-scale assessment is very complex but from a scientific point of view it is feasible. The workshop participants confirmed the importance of a PISA-VET but emphasised the required effort of corresponding international research activities to develop the design. This concerns not only the scientific aspect of developing methods for competence measurement and schemata for an analysis of institutional and individual context factors influencing educational

processes, but also behaviour and utilisation. The development of methodologies in this field has political implications with regard to the definition of educational objectives and -programmes, which in turn entails consensual agreement on the political, economic, and cultural principles of VET.

In the following paragraph, the results of the preceding discussions will be summarised and condensed in two proposals for a PISA-VET research design. The plural term "proposals" expresses that, for pragmatic reasons, there is not only *one* possible solution; differences between the revenue and expenses of a comparative study, as well as the most efficient ways of achieving the objectives with reasonable resources and within a calculable period of time, have to be balanced.

The requirements for any research design are not determined solely be the nature of large-scale assessments and the relation between competencies and the corresponding micro- and macro-structural context factors; measurement has to account for the quantitatively most relevant occupational fields and the most important ways to attain competence development in different countries. On the basis of these considerations, we will introduce two models for an international PISA-VET as a basis for further modification:

- First, a real panel, longitudinal study with multiple measuring points; and
- second, a cross-sectional study.

Both options will be discussed with regard to the objectives and pragmatic aspects of a comparative study.

5.2 RESEARCH DESIGN OF A LONGITUDINAL STUDY

For theoretical aspects of measurement in the context of an international comparison of VET, a longitudinal study, with a minimum of three panel waves, can be considered the "gold standard" of comparative large-scale assessments in VET (*Table 5.1*).

Table 5.1: Panel construction – structure of single waves

	1st Wave	2nd Wave	3rd Wave
Sampling	Institution-based Sampling	Panel	Panel
Measurement of Competencies	Initial Competencies, Motivation	Competencies Developed through Educational Processes	Expansion of Competencies After Completing VET, Application of Competencies at Work, in the Labour Market, and in Personal Biographies
Measurement of Institutional Data	Systemic Context Factors, Structure of Educational Institution	Organisational Characteristics of Educational Institution, Educational Conditions, and Instructional Arrangements, Learning Climate	Sector- and Occupation-Related Conditions, Regulation Forms of Transition to or within Labour Markets
Measurement of Individual Data	Life- and Learning Conditions, Information- and Decision Behaviour During Vocational Choice, Educational Aspiration, Learning- and Working Biography	Attended VET-Programmes, Content and Location of Learning, Type of Working Experience, Perception of Educational Conditions and –Situations, Learning Time, Future Orientation	Qualification (Certified), Occupational Status, Income, Performed Tasks, Continuing Education Activities, Occupational Mobility

Research design

The first wave should take place shortly after entering a VET-programme to measure the initial competence level of students.

The second wave should take place shortly before completing a VET-programme and after approximately 2.5 years (provided the study is of a 3-year duration) to measure competence development during VET.

The third wave should take place three years after completing a VET-programme and after the transition into the labour market to evaluate whether the individual is working in a job, participates in a program of continuing education, or is unemployed. Measurement is focused on the sustainability of competencies, in addition to their expansion and usability in the domains of work and everyday life.

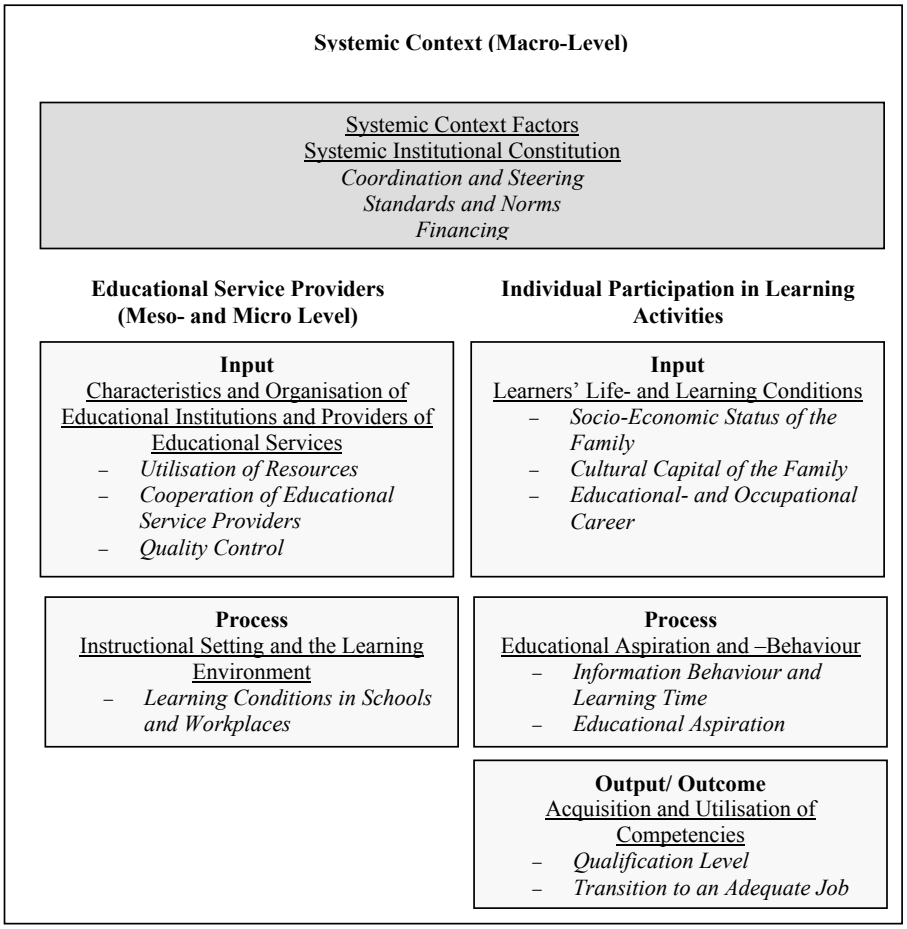
Measurement programme

- Measurement of competencies in generic, cross-occupational, and subject-related vocational dimensions (*Table 5.2*)
- Institutional and individual contexts of learning (*Figure 5.1*)

Table 5.2: Measurement plan for competencies in VET

	Competence Domains (different contextualised areas of performance)		
	Self-Competence		
	Cognitive Competence	**Functional Competence**	**Social Competence**
Individual Capacities (accessed and interpreted in different contexts)	(theoretical/analytical requirements) "applying concepts"	(technical/practical/ functional requirements) "using tools, equipments and technical resources"	(interpersonal requirements) "interacting with others"
Attitudes Values Perceptions	self-efficacy		
Incentives Motivation	effort and perservance/interest and motivation/ self-related cognition		
Meta-Cognitive Strategies	learning strategies		
Declarative Knowledge Procedural Knowledge	complex task for simultaneous measurement of cognitive and functional competence		questionnaire for mindful identity negotiation
Strategic Knowledge	according to the occupational field complex task for measuring cognitive competence	complex task for measuring functional competence	(consists of items for declarative, procedural and strategic knowledge in the domain for social competence)
Competencies of general education (corresponding to PISA)			

Figure 5.1: Measurement plan for institutional and individual context factors



The sample

- Occupational fields (*Table 5.3*)
- Sampling design for Germany (*Table 5.4*)

Table 5.3: Occupational fields, included VET-programmes, ISCED levels, and entrance age

Occupational Field	**VET -Program**	**ISCED levels (international)**	**Range of Entrance Age**
Metal and Electrical Occupations	Car Mechatronic	3 3 3	16–20 16–20 16–20
Commercial/ Administrative Occupations	Banker	3–5	17–20
Healthcare Occupations	Nurse	4–5	18–20
Information and Communication Technology Occupations	Skilled Computer Scientist	3–5	17–20

Table 5.4: Sampling design for Germany

	Size of Estimated Basis Sample (Simple Random Sampling, Margin of Error 5 Percent)	**Enlargement Based on Estimated Panel Mortality (50 Percent)**	**Enlargement Based on an Adjustment Coefficient for Clustering (1.08/2.9)**	**Sample Size Panel Study (1st wave)**
Car Mechatronic	400	+ 400	+ 64	= 864
Mechanic (Industry)	400	+ 400	+ 64	= 864
Electrician (Industry)	400	+ 400	+ 64	= 864
Banker	400	+ 400	+ 64	= 864
Nurse	400	+ 400	+ 1520	= 2320
Skilled Computer Scientist	400	+ 400	+ 64	= 864
Total of VET programmes				**= 6640**

Analysis of institutions
We suggest supplementing the assessment of individuals with expert interviews and observations to measure institutional context factors of VET. In an institution-based sample, this is necessary for relating different results of competence-measurement and contextual factors of schools and firms to each other. Consequently, the sample design is based on aspects and indicators of the main study and consists of the same schools, firms, and industry-wide institutions.

Design variants
Concerning the main option of a panel design with three waves, two possible variants might help to generate the outcomes in a shorter period of time. However, both methods involve constraints regarding the scope of conclusions/outcomes:

- The first method consists of a panel with two measuring points: The first one at the beginning and the second one at the end of a VET-programme. This approach focuses on measuring competence development during VET, including inferences regarding institutional and individual context factors. However, the aspect of competence utilisation in the labour market, in the workplace, and with regard to students' aspired educational objectives after completing VET, is missing.
- The second method also comprises two measuring points: The first one after completing a VET programme and the second one three years later. This approach is focused on the measurement of educational outcomes in the labour market, the workplace, and with regard to students' aspired educational objectives. Even though aspects of competence development during VET are missing in this approach, a classification of competencies according to institutional arrangements after completing VET is still possible. Thus, both methods could provide comprehensive results within a short period of time.

Strengths and weaknesses of the model
With regard to the objectives of a PISA-VET, the strengths of a panel design with three measuring points can be summarised as follows:

- Compared to school-based comparative studies, it allows measurement of competence *developments* and classification with regard to institutional context factors and vocational didactic approaches;
- inclusion of a third measuring point, approximately three years after completing VET, allows detailed predictions to be made regarding the sustainability and utilisation of competencies in the labour market, the workplace, and with regard to students' aspired educational objectives (e.g., career pathways, continuing education programmes); and
- international bench-marks at the level of educational institutions and occupational fields can be created, based on the linkage between competence developments and educational institutions.

Despite these strengths there are fundamental weaknesses:

- Implementing a panel study is very costly due to the required maintenance and duration. However, important results with regard to available competencies in different dimensions are accessible immediately after the first wave has taken place;
- realising the advantages of a panel study is only feasible within institution-based samples, which is connected to the problem of vertical comparability (i.e., determining age cohorts according to VET-participation and completion); and
- young people outside formal VET-programmes and drop-outs would not be included in an institution-based sample.

Both options would reduce measurement complexity by about three years. Depending if the main research objectives is more focused on developing competence during VET or on the utilisation of competencies after completing VET, the decision would favour the 1st or the 2nd option.

5.3 RESEARCH DESIGN OF A CROSS-SECTIONAL STUDY

From a pragmatic point of view, a cross-sectional study, initially having one measuring point in a homogeneous age cohort, represents a remarkable alternative to a longitudinal study.

Research design

The main goal of a cross-sectional study is to meet the objectives of a PISA-VET (*Chapter 1*) within a shorter period of time and with fewer costs than with a panel-study. Thus, a cross-sectional study represents a compromise between the broad objectives and the feasibility of a PISA-VET. Whereas in the first workshop, all participants voted for a real panel-study, in the second workshop, one group recommended a detailed examination of cross-sectional designs.

The research design consists of one measuring point in one age cohort (e.g., 24 years). The measurement comprises proficiency levels in different competence dimensions, as well as retrospective analysis of school-, VET-, and work-biographies. Likewise, the focus of a real-panel design is on competence measurement in general and subject-specific vocational dimensions.

Measurement programme

Compared to the first option (panel design), competencies are measured at three levels: Competencies of general education, cross-occupational competencies, and subject-specific vocational competencies (*Chapter 2* and *Table 5.2*). Correspondingly, institutional and individual contexts of education are measured in the same way as in the first option (*Table 5.2*). However, institutional contexts are focused on those dimensions, which are measurable by questionnaires.

The sample

The occupational tasks performed until the age of 24 provide the basis of the sample, since information. To formally completed vocational qualifications are excluded. A stratified sample is defined on the basis of the same occupational fields as in the first option (*Table 5.3*). For the determination of the initial sample, the following approach is recommended:

- 24-year olds can be identified in the relevant occupational fields together with their employers and firms on the basis of ISCO-codes and national European labour force surveys;
- these individuals provide the basis for either an institution-based stratified sample or random sampling (*Table 5.5*).

Table 5.5: Sampling Design: Cross-Sectional Study

	Size of Calculated Basis Sample (simple random sampling, error margin 5 percent)	**Enlargement Based on the Proportion of Individuals without any Qualification (estimated 25 percent)**	**Enlargement Based on an Adjustment Coefficient for Clustering (1.08)**	**Sample Size Cross-Sectional Study**
Car Mechatronic	400	+ 100	+ 40	= 540
Mechanic (Industry)	400	+ 100	+ 40	= 540
Electrician (Industry)	400	+ 100	+ 40	= 540
Banker	400	+ 100	+ 40	= 540
Nurse	400	+ 100	+ 40	= 540
Skilled Computer Scientist	400	+ 100	+ 40	= 540
Total of VET-programmes				**= 3240**

Strengths and weaknesses of the model

With regard to the outlined objectives of a PISA-VET, the chances of a cross-sectional design with one measuring point can be summarised as follows:

- The birth cohort model avoids the problem of age variance;
- the model allows integration of individuals without formal education working within the relevant occupational fields and it provides a basis for creating a bench-mark between formal and informal learning processes;
- retrospective measurement of education and work biographies allows the classification of VET-typologies, including the entire spectrum of formal and informal learning processes (in a panel design, this is possible only to a very limited extent);

- the model provides information on occupational competencies in specific age cohorts in less time than a panel-design model;
- the model is more complex with respect to overall costs and time requirements than a panel-design model; and
- the development of measurement tools is less costly than that required for a panel-design model since instruments are needed only for one measuring point. In contrast, expenses for the measurement of individuals are much higher given that retrospective measurement of education and work biographies is not feasible on the basis of standardised face-to-face interviews.

However, despite these strengths, there are fundamental weaknesses:

- The measurement of competence developments during VET processes or in the workplace is not possible, whereas relating competencies to education typologies and work biographies is feasible;
- competencies cannot be explicitly linked to institutional settings due to the lack of dynamic competence measurements (competence development), which heavily impacts aspects of general social or meta-cognitive competencies since they can be acquired outside educational contexts;
- the reconstruction of education and work biographies has to account for the problem that individual memories cannot be considered very reliable information sources; and
- the measurement of transitions to occupations, which are distinctive from the respective vocational qualification, are not possible, only transitions *into* the relevant occupation.

6. FURTHER STEPS

The presented feasibility study highlights opportunities and necessary developments for the implementation of an international comparative large-scale assessment:

- The research perspectives and indicators discussed in *Chapters 2, 3,* and *4* must be operationalised with adequate measurement instruments. Development efforts for competence-based measurement tools are particularly high due to the focus on vocational subject-related competencies;
- a decision with regard to the research design (panel-longitudinal vs. cross-sectional study) must be made; and
- an international research community for the development of a "Programme for an International Student Assessment of Vocational Education and Training" (PISA-VET) must be established:
- The expenses for the feasibility study, including international co-operation, were paid by the German government (Federal Ministry of Economics and Labour). All international experts have appreciated this initiative very much, having in mind that a continuation of the project requires an international financial and organisational basis. Interest in participation was manifested from various institution in different countries.

The research network established at the workshops could be used as a basis for developing an international research community. During the final session of the workshop (30th of April, 2005), all the participants manifested their interest to continue co-operating in this project.

As a conclusion of the discussion at the second workshop, we can distinguish between two possible ways for further developments, which are termed "small solution" and "large solution."

Small solution

Single countries co-operate in small groups for the purpose of developing a PISA-VET. This option allows the aggregation of countries with similar structures of VET; (for example, one group of German speaking countries with dual VET-systems consisting of Switzerland, Austria, and Germany, and another group comprising Scandinavian countries and Great Britain. A "small" solution consists of less coordinating efforts and allows a quicker start. On the other hand, it hampers benchmarking due to the limited scope of VET-institutions. However, it represents a chance for comparing different VET-programmes, since even within a small number of countries many different programmes are coexisting. Moreover, possibilities of collective development of research methods as well as the option of relating different research approaches should be taken into consideration.

Large solution
An international organisation, such as EU or OECD, coordinates a comparison of VET in a number of countries. On the one hand, this allows a broad comparison based on policy-supported research. On the other hand, efforts for political coordination as well as reaching consensus on standards and measurement methods for competencies are very time-consuming. This may paralyse scientific activities as well as policy interest after the “PISA-shock”.

7. REFERENCES

Achtenhagen, F. (1996). Entwicklung und Evaluation ökonomischer Kompetenz mit Hilfe handlungsorientierter Verfahren am Beispiel der Ausbildung zum Industriekaufmann/zur Industriekauffrau (Skizze eines Forschungsprojekts). In W. Seyd & R. Witt (Eds.), *Situation – Handlung - Persönlichkeit. Kategorien wirtschaftspädagogischen Denkens. Festschrift für Lothar Reetz* (pp. 137–147). Hamburg: Feldhaus.

Achtenhagen, F. (1998). General versus vocational education: Demarcation and integration. In W. J. Nijhof & J. N. Streumer (Eds.), *Key qualifications in work and education* (pp. 133–143). Dordrecht: Kluwer.

Achtenhagen, F. (1990). Vorwort. In Deutsche Forschungsgemeinschaft (Ed.), *Berufsbildungsforschung an den Hochschulen der Bundesrepublik Deutschland* (pp. VI–VII). Weinheim, Basel: VCH.

Achtenhagen, F. (2004). Curriculum development as modelling of complex reality. In N. M. Seel & S. Dijkstra (Eds.), *Curriculum, plans, and processes in instructional design* (pp. 193–210). Mahwah, N. J.: Erlbaum.

Achtenhagen, F., & John, E. G. (Eds). (1992). *Mehrdimensionale Lehr-Lern-Arrangements – Innovationen in der kaufmännischen Aus- und Weiterbildung.* Wiesbaden: Gabler.

Achtenhagen, F. in Cooperation with J. W. Nijhof & D. Raffe (1995). *Feasibility study: Research scope for vocational education in the framework of COST Social Sciences.* European Commission: Directorate-General Science, Research and Development, COST Technical Committee Social Sciences, Vol. 3. Brussels, Luxembourg: ECSC-EC-EAEC.

Achtenhagen, F. & Baumert, J. (1999). Alternative Evaluationsverfahren – eine Einführung [Alternative evaluation methods: An introduction]. *Unterrichtswissenschaft, 27,* 98–101.

Achtenhagen, F., & Lempert, W. (Eds.) (2000). *Lebenslanges Lernen im Beruf – Seine Grundlegung im Kindes- und Jugendalter. Vol. 1: Das Forschungs- und Reformprogramm.* Opladen: Leske & Budrich.

Achtenhagen, F., & Grubb, W. N. (2001). Vocational and occupational education: Pedagogical complexity, institutional diversity. In V. Richardson (Ed.), *Handbook of research on teaching. Fourth edition* (pp. 604–639). Washington, DC: American Educational Research Association.

Achtenhagen, F., & Weber, S. (2003). „Authentizität" in der Gestaltung beruflicher Lernumgebungen. In A. Bredow, R. Dobischat & J. Rothmann (Eds.), *Berufs- und Wirtschaftspädagogik von A-Z* (pp. 185–199). Baltmannsweiler: Schneider Verlag Hohengehren.

Achtenhagen, F., Preiß, P., & Weber, S. (2005). *Lerntheoretische Grundlagen für die Entwicklung von Lehr-/Lernarrangements.* Oldenburg: Universität Oldenburg.

Aebli, H. (1987). *Grundlagen des Lehrens: Eine allgemeine Didaktik auf psychologischer Grundlage*. Stuttgart: Klett-Cotta.

Altiner, A., Wilm, S., & Sielk, M. (2002). Allgemeinmedizinische Inhalte im Medizinstudium. Implementation einer Klausur. *Zeitschrift für Allgemeine Medizin*, 425–430.

Anderson, L. W., & Krathwohl, D. R. (with Airasian, P.W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al.) (Eds.) (2001). *A taxonomy for learning, teaching, and assessing. A revision of bloom's taxonomy of educational objectives.* New York: Longman.

Andersson, R., & Olsson, A.-K. (1999). *Fields of education and training.* Luxembourg: Eurostat. Appendix A3: Commission of the European Communities (2005), S. 18–20.

Association of German Engineers (VDI) (2004). *Bildungsstandards im Fach Technik für den mittleren Schulabschluss. Arbeitspapier zum Politikdialog „Kompetent durch Technikunter-*

richt". Tagung in Berlin. http://www.vdi-jutec.de/medienarchiv/ablage/original/tagungsunterlage.pdf.

Association for Research in Professional Development (ABWF) (2005). *Kompetenzmessung im Unternehmen: Lernkultur- und Kompetenzanalysen im betrieblichen Umfeld*. Band 18. Münster et al.: Waxmann.

Baethge, M., Buss, K. P., & Lanfer, C. (2003). *Konzeptionelle Grundlagen für einen Nationalen Bildungsbericht – Berufliche Bildung und Weiterbildung/Lebenslanges Lernen*. Bonn: BMBF.

Baethge, M., & Baethge-Kinsky, V. (Eds.) (2006). *Der ungleiche Kampf um das lebenslange Lernen.* Münster: Waxmann.

Baethge, M., & Baethge-Kinsky, V. (2005). Ökonomie, Technik, Organisation: Zur Entwicklung von Qualifikationsstruktur und qualitativem Arbeitsvermögen. In R. Arnold & A. Lipsmeier (Eds.). *Handbuch der Berufsbildung.* 2. Aufl. Opladen: Leske&Budrich.

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25,* 725–739.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider W., et. al. (1999). *Soziale Bedingungen von Schulleistungen – Zur Erfassung von Kontextmerkmalen durch Schüler-, Schul- und Elternbefragungen.* http://www.mpib-berlin.mpg.de/pisa/Kontextmerkmale.pdf.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., et al. (2001). *PISA 2000*: *Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.

Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzwettbewerb im internationalen Vergleich. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, et al. (Eds.), *PISA 2000: Die Länder der Bundesrepublik Deutschland im Vergleich* (pp. 159–202). Opladen: Leske & Budrich.

Baumert, J., Artelt, C., Carstensen, C. H., Sibberns, H., & Stanat, P. (2002). Untersuchungsgegenstand, Fragestellungen und technische Grundlagen der Studie. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, et al. (Eds.), *PISA 2000: Die Länder der Bundesrepublik Deutschland im Vergleich* (pp. 11–38). Opladen: Leske & Budrich.

Beck, K. (1995). Die Äquivalenz von Ausbildungsabschlüssen in Europa als Voraussetzung transnationaler beruflicher Mobilität? Ein erziehungswissenschaftlicher Ansatz zur Lösung des Vergleichsproblems. In K. A. Schachtschneider (Ed.), *Wirtschaft, Gesellschaft und Staat im Umbruch. Festschrift der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg 75 Jahre nach Errichtung der Handelshochschule Nürnberg* (pp. 1–16). Berlin: Duncker & Humblodt.

Bendorf, M. (2005). Förderung von Metakognition und Lernstrategien am Fachgymnasium Wirtschaft. In P. Gonon, F. Klauser & R. Huisinga (Eds.), *Kompetenz, Kognition und neue Konzepte der beruflichen Bildung* (pp. 203–218). Wiesbaden: Verlag für Sozialwissenschaften.

Bjornavold, J. (2000). *Making learning visible: Identification, assessment and recognition of non-formal learning in Europe*. Luxembourg: Official Publication of the European Communities.

BLK (2002). *Vergleichende Bildungsstatistik – Sachstand und Vorschläge zur Verbesserung,* Heft 103. Bonn.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Ed.), *Soziale Ungleichheiten, Soziale Welt*, *Sonderband 2*, 183–198.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, D.C.: National Academy Press.

Büchel, F., & Neubäumer, R. (2001). Ausbildungsinadäquate Beschäftigung als Folge branchenspezifischer Ausbildungsstrategien. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 34,* 269–285.

CEDEFOP (1998). *Transparenz beruflicher Befähigungsnachweise in Europa – Stand und Entwicklungsperspektiven.* Thessaloniki.

Commission of the European Communities (2005). *Commission staff working document: Towards a European Qualifications Framework for lifelong learning.* http://europa.eu.int/comm/education/policies/ 2010/doc/ consultation_eqf_en.pdf.

CEDEFOP (2003). *Terminology of vocational training policy: A multilingual glossary for an enlarged Europe. http://www. europass.cedefop.eu.int/img/dynamic/c313/cv-1_en_US_ glossary_4030_6k.pdf.*

COM/DELSA/EDU (2004). *Programme for the International Assessment of Adult Competencies: Draft strategy paper.* http://www.oecd.org/dataoecd/3/3/34463133.pdf.

Corti, S. (2000). Forschungserfordernisse der Berufsbildungspolitik der Europäischen Union. In CEDEFOP/DIPF (Eds.), *Vergleichende Berufsbildungsforschung in der Europäischen Union,*(pp. 28–29). Frankfurt am Main, Thessaloniki. http://www2.trainingvillage.gr/ down-load/publication/reference/BonnReaderDE.pdf.

Deci, E. L., & Ryan, M. R. (1985). *Intrinsic motivation and self-determination in human behaviour.* New York, et al.: Plenum Press.

Descy, P., & Tessaring, M. (2002). *Kompetent für die Zukunft – Ausbildung und Lernen in Europa. Zweiter Bericht zur Berufsbildungsforschung in Europa: Synthesebericht*. Reihe Cedefop reference series, Bd. 5. Luxemburg.

Dolch, J. (1965). *Lehrplan des Abendlandes: Zweieinhalb Jahrtausende seiner Geschichte.* 2. Aufl. Ratingen: Henn.

Ellström, P.-E. (1997). The many meanings of occupational competence and qualification. In W. J. Nijhof & J. N. Streumer (Eds.), *Key qualifications in work and edu*cation (pp. 39–50). Dordrecht: Kluwer.

Erpenbeck, J., & von Rosenstiel, L. (2003a). *Handbuch Kompetenzmessung.* Stuttgart: Schäffer-Poeschel.

Erpenbeck, J., & von Rosenstiel, L. (2003b). Einführung. In J. Erpenbeck & L. von Rosenstiel (Eds.), *Handbuch Kompetenzmessung* (pp. IX-XL). Stuttgart: Schäffer-Poeschel.

European Commission (2005). *Auf dem Weg zu einem Europäischen Qualifikationsrahmen für Lebenslanges Lernen.* Brüssel.

EUR-Lex (1985). *85/368/EWG: Entscheidung des Rates vom 16. Juli 1985 über die Entsprechungen der beruflichen Befähigungsnachweise zwischen Mitgliedstaaten der Europäischen Gemeinschaften*. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do? uri=CELEX: 319 85D0368: DE:HTML.

Europäische Sozialagenda (2001). Rat der Europäischen Union. Luxemburg: Amt für amtliche Veröffentlichungen der Europäischen Gemeinschaften.

Europäische Union (2001). *Übergang vom Bildungswesen ins Erwerbsleben – Schlüsselzahlen zur Berufsbildung in Europa.* Eurostat, Cedefop.

Eurostat/CEDEFOP (2002). *Overview of National VE- systems.* Thessaloniki. http://www.trainingvillage.gr/etv/Information_resources/NationalVet/.

Federal Institute for Vocational Education and Training (BIBB) (1999). *Komplexe Prüfungsaufgaben für Bürokaufleute: KoPrA- ein Leitfaden zur Erstellung von Prüfungsaufgaben für die Kammerprüfungen*. Bielefeld: Bertelsmann.

Federal Institute for Vocational Education and Training (BIBB) (2005). *Ausbilden mit Lernaufgaben. Band 1-3*. Konstanz: Christiani.

Federal Ministry of Education and Research (2006). *Bildung in Deutschland* [Education in Germany]. Bielefeld: Bertelsmann

Fischer, G., & Siebern-Thomas (2004). Arbeit und Arbeitsmärkte im Zeitalter wissensbasierter Wirtschaft und Gesellschaft. Gibt es noch einen nationalen Arbeitsmarkt? In E. Hönekopp, R. Jungnickel & T. Straubhaar (Eds.), *Internationalisierung der Arbeitsmärkte, Beiträge zur*

Arbeitsmarkt- und Berufsforschung, *Nr. 282* (pp. 35–60). Nürnberg: Bundesagentur für Arbeit.

Fisher, C. W., & Berliner, D. C. (Eds.) (1985). *Perspectives on instructional time*. New York: Longman.

Fulst-Bleil, S., & Ebner, H. G. (2005). Duale Berufsausbildung im internationalen Vergleich. *Zeitschrift für Berufs- und Wirtschaftspädagogik, 101*, 176–185.

Franke, G., & Kleinschmitt, M. (1987). *Der Lernort Arbeitsplatz.* Berlin: Beuth.

Ganzeboom, H. B. G., & Treiman, D.J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & Ch.Wolf (Eds.), *Advances in cross-national comparison* (pp. 159–193). New York: Kluwer.

German Employers Organisation for Vocational and Further Training (2005). *Berufliche Bildung für Europa: Europäischer Qualifikationsrahmen (EQF) und Leistungspunktesystem (ECVET).* http://www.kwb-berufsbildung.de/pdf/2005_Positionspapier_EQF_ ECVET.pdf.

Gomez, P., & Probst, G. (1997). *Vernetztes Denken im Management.* Bern: Schweizerische Volksbank.

Green, F., McIntosh, S., & Vignoles, A. (2004). *Scoping study on international comparisons of non-certified learning.* http://www.dti.gov.uk/training_development/Non_certified learning_final.pdf.

Greinert, W.-D. (1997). *Das duale System der Berufsausbildung in der Bundesrepublik Deutschland.* Stuttgart: Holland & Josenhans.

Grünewald, U., Moraal, D., & Schönfeld G. (2003). Betriebliche Weiterbildung in Deutschland und Europa. Bielefeld : Bertelsmann.

Hannan, D., Raffe, D. & Smyth, E. (1997). Cross-national research on school to work transitions: An analytical framework. In P. Werquin, R. Breen & J. Planas (Eds.), *Youth transitions in Europe:Theories and evidence* (pp. 409–422). Marseille: Cereq.

Hanushek, E. A. (1989). Expenditures, efficiency, and equity in education: The Federal Government's role. *The American Economic Review, 79,* 41-51.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19,* 141–164.

Hasselhorn, M. (2000). Lebenslanges Lernen aus der Sicht der Metakognitionsforschung. In F. Achtenhagen & W. Lempert (Eds.), *Lebenslanges Lernen im Beruf – seine Grundlagen im Kindes– und Jugendalter (III) Psychologische Theorie, Empirie und Therapie* (pp. 41–53). Opladen: Leske & Budrich.

Heid, H. & Harteis, C. (2005). *Verwertbarkeit – Ein Qualitätskriterium (erziehungs-)wissenschaftlichen Wissens.* Wiesbaden: Verlag für Sozialwissenschaften.

Heikkinen, A. (2001). Ist eine stärkere Berücksichtigung der historischen und kulturellen Dimension in der Vergleichenden Berufsbildungsforschung erforderlich? In CEDEFOP/DIPF (Eds.), *Vergleichende Berufsbildungsforschung in Europa. Ansätze, Politikbezüge und Innovationstransfer* (pp. 118-134). Thessaloniki.

ILO (2005): *International Standard Classification of Occupations.* http://www.ilo.org/public/english/bureau/stat/class/isco.htm.

International Technology Education Association (2002). *Standards for technological literacy content for the study of technology, second edition.* http://www.iteawww.org.

Jeanneret, P. R. (1969). *A study of the job dimensions of "worker-oriented" job variables and of their attribute profiles.* Dissertation Abstracts International.

Jeanneret, P. R., Borman, W. C., Kubisiak, U. C., & Hanson, M. A. (2002). Generalized Work Activities. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret & E. A. Fleishman (Eds.), *An occupational information system for the 21st century: The development of O*NET* (pp. 105–125). Washington, D.C.: American Psychological Association.

Kanning, U.P. (2003). *Diagnostik sozialer Kompetenzen.* Göttingen: Hogrefe.

Kearns, P. (2001). *Generic skills for the new economy.* Australian National Training Authority. http://www.ncver.edu.au/research/proj/nr0024.pdf.

Neubrand, M., Klieme, E., Lüdtke, O., & Neubrand, J. (2002). Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung. *Unterrichtswissenschaft, 30,* 100–119.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, et al. (2003). *Zur Entwicklung nationaler Bildungsstandards – Eine Expertise.* Bonn: Bundesministerium für Bildung und Forschung.

Koch, R., & Reuling, J. (1998). Öffentliche Qualitätssteuerung der Berufsausbildung in Deutschland, Frankreich und England. In CEDEFOP (Ed.), *Berufsbildung, Bd. 15.* Thessaloniki.

Kogan, I. & Müller, W. (Eds.) (2003). *School-to-work transitions in Europe: Analyses of the EU LFS 2000 Ad hoc Module.* Mannheim. http://www.mzes.uni-mannheim.de /publications/ books /school_to_work_2ed.pdf.

Kuhl, J. (1984). *Motivation, Konflikt und Handlungskontrolle.* Heidelberg: Springer.

Kuhl, J. (1998).Wille und Persönlichkeit. *Psychologische Rundschau, 49*, 61–77.

Kunter, M. et al. (2002). *PISA 2000 – Dokumentation der Erhebungsinstrumente.* Max-Planck-Institut für Bildungsforschung, Berlin. http://www.kmk.org/schul/pisa/Datensaetze/Gesamt-Druckversion.pdf.

Lamb, S., Davies, M., Polesel, J. & Teese, R. (2003). *International indicators for vocational education and training.* Leabrook.

Lang, B. & Pointinger, M. (2002). Das selbst gesteuerte Lernen österreichischer Jugendlicher. In C. Reiter & G. Haider (Eds.), *PISA 2000 – Lernen für das Leben. Österreichische Perspektiven des internationalen Vergleichs* (pp. 101–104). Innsbruck: Studienverlag Innsbruck.

Lang-von Wins, T. (2004). *Der Unternehmer: Arbeits- und organisationspsyhologische Grundlagen.* Berlin et al.: Springer.

Lasonen, J., & Manning, S. (2001). How to improve the standing of vocational compared to general education: A collaborative investigation of strategies and qualifications across Europe. In P. Descy & M. Tessaging (Eds.), *Training in Europe. Second report of vocational training research in Europe*, *Vol. I.* (pp. 115–167), Thessaloniki.

Lassnigg, L. (2000). *Steuerung, Vernetzung und Professionalisierung in der Berufsausbildung.* Wien: Institut für Höhere Studien (IHS). http://www.lebenslangeslernen.at/downloads/ IHS_SteuerungBerufsausbildung_1100.pdf.

Lassnigg, L. (2003). *Indicators for quality in VET – European experience.* Vienna: Institute for Advanced Studies.

Lempert, W. (1979). Zur theoretischen und empirischen Analyse von Beziehungen zwischen Arbeiten und Lernen. Grundprobleme und Lösungsstrategien. In P. Groskurth (Hrsg.), *Arbeit und Persönlichkeit. Berufliche Sozialisation in der arbeitsteiligen Gesellschaft* (pp. 87–111). Reinbek bei Hamburg.

Leney, T., & Deluca, M. (2000). Einheit und Vielfalt der Systeme beruflicher Bildung der Mitgliedsstaaten der Europäischen Union – Konvergenz- und Divergenztendenzen im Bildungs- und Berufsbildungswesen der Mitgliedstaaten der Europäischen Union. In CEDEFOP/DIPF (Eds.), *Vergleichende Berufsbildungsforschung in Europa. Ansätze, Politikbezüge und Innovationstransfer* (pp. 103–118). Thessaloniki.

Leschinsky, A., & Mayer, K. U. (1999). *The comprehensive school experiment revisited.* Frankfurt: Lang.

Lisbon-to-Copenhagen-to-Maastrich Consortium Partners (2004). *Achieving the Lisbon goal: The contribution of VET – Final report to the European Commission.* http://europa.eu.int/comm/education/policies/2010/studies/maastricht_en.pdf.

Little, T. D., Oettingen, G. & Baltes, P. B. (1995). *The revised control, agency, and means-ends interview (CAMI)* (49). Berlin: Max Planck Institute for Human Development and Education.

Mertens, D. (1974): Schlüsselqualifikationen: *Thesen zur Schulung für eine moderne Gesellschaft. Mitteilungen aus Arbeitsmarkt- und Berufsforschung 7,* 36–43.

Ministry of Education in Germany KMK (1996/2000). *Handreichungen für die Erarbeitung von Rahmenlehrplänen der Kultusministerkonferenz (KMK) für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe.* http://www.kmk.org/doc/publ/handreich.pdf.

Münch, J. (1999). *Qualifikation als Standortfaktor.* Hochheim am Main: Neres.

Murray, T. S., Clermont, Y., & Binkley, M. (2004). *The International Adult Literacy and Life Skills Survey: Aspects of design, development and validation.* Unpublished paper.

National Business Education Association (2001). *National Standards for Business Education: What Americas students should know and be able to do in business*. NBEA

National Center for Education Statistics (NCES) (1998). *Adult literacy in OECD countries: Technical report on the first international Adult Literacy Survey.* Washington, D.C.: NCES.

National Council on Economic Education (2000). *Voluntary national content standards in economics.* National Council on Economic Education: New York.

National Framework of Qualifications (2003). *Determinations for the outline national framework of qualifications*. http://www.nqai.ie/determinations.pdf.

National Training information service (2005). www.ntis.gov.au.

NCVER (2003). *International indicators for vocational education and training – An Australian perspective.* Adelaide.

NCVER (2004). *International benchmarking of vocational education and training.* Adelaide.

Nieuwenhuis, L. (2004). *The learning potential of the workplace.* http://edu.fss.uu.nl/ord/ fullpapers/Nieuwenhuis-L_L%204_2.DOC.

Nijhof, W. J. (1998). Qualifying for the future. W. J. Nijhof & J. N. Streumer (Eds.), *Key qualifications in work and education* (pp. 19–38). Dordrecht: Kluwer.

Norris, N. (1991). The trouble with competence. *Cambridge Journal of Education, Vol. 21,* 1–11.

Oates, T. (2004). The role of outcome-based national qualifications in the development of an effective vocational education and training system: The case of England and Wales. *Policy Futures in Education, Volume 2.*

OECD (2001). *Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000,* Paris. http://www.pisa.oecd.org/Docs/Download/PISA 2001(deutsch).pdf.

OECD (2003): *Education at a glance,* Paris.

OECD (2004). *PIAAC draft strategy paper – Policy objectives, strategic options and cost implications*. http://www.oecd.org/dataoecd/3/3/34463133.pdf.

Office of Apprenticeship Training, Employer and Labor Services (OATELS)(2004). *The national registered apprenticeship system – Programmes and apprentices.* Fiscal Year 2003.

Onstenk, J. H. A. M. (2003). *Work based learning in continuing vocational and professional education, Paper for WACE – First Draft.* http//www.wacerotterdam2003.nl/documents/ final_papers_abstracts/140.doc.

Pintrich, P. R., Smith, J. A., Garcia, T., & McKeachie, W. (1991). A manual for the use of motivated strategies for learning questionnaire (MSLQ). University of Michigan: National Center for Research to Improve Postsecondary Teaching and Learning.

Preiß, P. (1992). Komplexität im Betriebswirtschaftslehre-Anfangsunterricht. In F. Achtenhagen & E. G. John (Eds.), *Mehrdimensionale Lehr-Lern-Arrangements – Innovationen in der kaufmännischen Aus- und Weiterbildung* (pp. 58–78). Wiesbaden: Gabler.

Preiß, P. (2004). *Strukturen der Beschreibung von Kompetenzen in England, Frankreich und Deutschland.* Vortrag auf der Herbsttagung der Sektion Berufs- und Wirtschaftspädagogik der Deutschen Gesellschaft für Erziehungswissenschaft, Siegen.

Prenzel, M., Kramer, K., & Drechsel, B. (2001). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung – Ergebnisse eines Forschungsprojekts. In K.

Beck & V. Krumm (Eds.), *Lehren und Lernen in der beruflichen Erstausbildung: Grundlagen einer modernen kaufmännischen Berufsqualifizierung* (pp. 37–63). Opladen: Leske & Budrich.

Rauner, F., & Spöttl, G. (2002). *Aufgabenanalyse für die Neuordnung der Berufe im Kfz-Sektor. Abschlussbericht: Wissenschaftliche Begleitung zur Neuordnung der fahrzeugtechnischen Berufe.* http://www.itb.uni-bremen.de/downloads/Sonstiges/Kfz-Neuordnung-Bericht4.pdf.

Rammstedt, B., Koch, K., Borg, I., & Reitz, T. (2004): *Entwicklung und Validierung einer Kurzskala für die Messung der Big-Five-Persönlichkeitsdimensionen in Umfragen.* http//www.gesis.org/Publikationen/Zeitschriften/ZUMA_Nachrichten/documents/pdfs/55/05_rammstedt.pdf.

Reetz, L. (1999). Zum Zusammenhang von Schlüsselqualifikationen – Kompetenzen – Bildung. In T. Tramm, D. Sembill, F. Klauser & E. G. John (Eds.), *Professionalisierung kaufmännischer Berufsbildung. Beiträge zur Öffnung der Wirtschaftspädagogik für die Anforderungen des 21. Jahrhunderts. Festschrift zum 60. Geburtstag von Frank Achtenhagen* (pp. 32–51). Frankfurt a. M.: Lang.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. *2. Aufl*. Bern: Huber.

Rost, J., Carstensen, C. H., & Davier, M. von (1999). Sind die Big Five Rasch-skalierbar? Eine Reanalyse der NFO-FFI-Normungsdaten. *Diagnostica, 45,* 119–127.

Roth, H. (1971). *Pädagogische Anthropologie, Bd. 2. Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik.* Hannover: Schroedel.

Rychen, D. S., & Salganik, L. H. (2003). A holistic model of competence. In D. S. Rychen & L. H. Salganik (Eds.), *Key competencies for a successful life and a well-functioning society* (pp. 41–62). Göttingen: Hogrefe & Huber.

Sandberg, J. (2000). Competence – the basis for a smart workforce. In R. Gerber & C. Lankshear (Eds.), *Training for a Smart Workforce,* (pp. 47–72). London: Routledge.

Scottish Credit and Qualifications Framework (2003). *An introduction to the Scottish Credit and Qualification Framework,* 2nd edition. Edinburgh. http://www.sqa.org.uk/files_ccc/ IntroductiontoSCQF-2ndEdition.pdf.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement, 1,* pp.61–80.

Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness.* Oxford: Pergamon Press.

Scheerens, J., (2004). *Review of school and instructional effectiveness research – Contribution to chapter 3 of the 2004 EFA Global Monitoring Report.* http://portal.unesco. org/education/en/file_download.php/23880a1c51a21961999fe362fcefde1dScheerens.pdf.

Scholz, S. (2003). *PISA 2000: Die Studie im Fokus der Statistik.* http://www.sozialforschung.uni-bremen.de/dip_scholz.pdf.

Schnell, R., Hill, P., & Esser, E. (1989). *Methoden der empirischen Sozialforschung.* München, Oldenbuorg.

Secretarys Commission on Achieving Necessary Skills (1991). *What work requires of schools. A SCANS report for America 2000: Executive summary.* Washington, DC: The Secretarys Commission on Achieving Necessary Skills, U.S. Department of Labour.

Sellin, B. (2005). *Europäischer Qualifikationsrahmen (EQF) – ein gemeinsames Bezugssystem für Bildung und Lernen in Europa.* http://www.bwpat.de/ausgabe8/ sellin_bwpat8.shtml.

Shavelson, R.J., & Ruiz-Primo, M. (2005). Windows into the mind. *International Journal of Higher Education and Educational Planning*, *48,* 260–278.

Smith, E., & Comyn, P. (2003). *The development of employability skills in novice workers. Australian National Training Authority.* http://www.ncver.edu.au/research/proj/nr2005.pdf?PHPSESSID=0f303f7a365c7efd09da16d87747e0bd.

Smits, M. H. S. B., Sluijsmans, D. M. A., & Jochems, W. M. G. (2005). *Performance assessment in nursing education: A case study.* Open University of the Netherlands. Presentation at the EARLI conference in Cyprus , August, 2005.

Spencer, L. M., & Spencer, S. M. (1993). *Competence at work.* New York et al.: John Wiley.

State of Oregon, Department of administrative Services, Human Resource Service Division (2005): http://www.hr.das.state.or.us/hrsd/class/6255.HTM.

Stasz, C., Ramsey, K., Eden, R., Melamid, E., & Kaganoff, T. (1996). *Workplace skills in practice – Case studies of technical work.* Santa Monica: National Centre for Research in Vocational Education, RAND.

Straka, G. A. (2003): *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen.* Münster: Waxmann.

Stoof, A. (2005). *Tools for identification and description of competencies.* Herleen: Open Universiteit Nederland. http://www.ou.nl/otecresearch/publications/Angela%20Stoof/Proefschrift%20Angela%20Stoof.pdf.

Tramm, T. (1992). *Konzeption und theoretische Grundlagen einer evaluativ-konstruktiven Curriculumstrategie – Entwurf eines Forschungsprogramms unter der Perspektive des Lernhandelns.* Göttingen: Seminar für Wirtschaftspädagogik der Georg-August-Universität Göttingen, Berichte: Band 17.

Treiber, B. (1982). Lehr- und Lernzeiten im Unterricht. In B. Treiber & F. E. Weinert (Eds.), *Lehr-Lern-Forschung* (pp. 12–36). München: Urban & Schwarzenberg.

Toolsema, B. (2003). *Werken met competenties. Naar een instrument voor de identificatie van competenties.* Enschede: Universiteit Twente.

Tuomi, I. (1999). *Corporate Learning.* Helsinki: Metaxis.

U.S. Department of Education & U.S. Department of Labor (1996). *Occupational skill standards project.* Washington, DC.

UNESCO (2004). *Global education digest 2004.* Montreal: UNESCO Institute for Statistics.

UNESCO (2005). *International standard classification of education.* http://www.unesco.org/-education/information/nfsunesco/doc/isced_1997.htm.

US Bureau of Labor Statistics (2005). *High Growth Industry Profile.* http://www.careervoyages.gov/.

Van den Berghe, W. (1997). *Indicators in perspective – The use of indicators in vocational education and training.* Luxembourg: CEDEFOP.

Weber, S. (2005). *Intercultural learning as identity negotiation.* Frankfurt am Main: Lang.

Weber, S., Achtenhagen, F., Bendorf, M., & Getsch, U. (2003). Überlegungen zur Gestaltung einer wirksamen Kooperation von Institutionen und Lernorten der beruflichen Bildung. In F. Achtenhagen (Ed.), *Neue Wirtschaftspädagogische Forschungs- und Entwicklungsarbeiten* Bd. 28 (pp. 298–331). Göttingen: Seminars für Wirtschaftspädagogik der Georg-August-Universität Göttingen.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle et al.: Hogrefe & Huber.

Weinert, F. E. (2002). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Ed.), *Leistungsmessung in Schulen,* (pp. 17–31). Weinheim & Basel: Beltz.

Weiß, M., & Preuschoff, C. (2004). *Kosten und Effizienzanalysen im Bildungsbereich. Referenzpapier für das Projekt „Bildungssteuerung".* Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.

Wigdor, A. K., & Green, B. F. (1991). *Performance assessment for the workplace, Volume I: Committee on the performance of military personnel. Commission on behavioural and social science and education.* National Research Council, Washington, D.C.: National Academy Press.

Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15,* 185–200.

Winther, E. (2005). Trait- und state-Komponenten der Motivation in Lernprozessen zwischen Interventions- und Forschungsanspruch. In P. Gonon, F. Klauser & R. Huisinga (Eds.), *Kompetenz, Kognition und neue Konzepte der beruflichen Bildung* (pp. 219–236). Wiesbaden: Verlag für Sozialwissenschaften.

Winterton, J., Delamare-Le Deist, F., & Stringfellow, E. (2005). *Typology of knowledge, skills and competences: Clarification of the concept prototype: Final draft.* Thessaloniki: CEDEFOP.

Wirth, J. (2003). *Selbstreguliertes Lernen in dynamischen Systemen. Unpublished Dissertation.* Berlin: Humbold-Universität zu Berlin.

Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education, 10,* 329–345.

Wolter, H.-J., & Petersen, H. (2003). *Ankauf und Aufbereitung von Unfallfahrzeugen für den CG-Fuhrpark: Expertenprojekt im Modellversuch GAB; Berufsgruppe Automobilmechaniker.* http://www.nibis.de/nli1/bbs/archiv/rahmenrichtlinien/gab/am_ep_fertig.pdf.

8. APPENDICES

Appendix A1: Participants of the First International Workshops on PISA-VET

Name	Institution	Department	Mail	accepted	tied up with business
Achtenhagen, Frank	University of Göttingen	Institute of Business and Human Resource Education	fachten@wipaed.wiso.uni-goettingen.de	X	
Arends, Lena	University of Göttingen	Institute of Business and Human Resource Education	larends@wipaed.wiso.uni-goettingen.de	X	
Babic, Edvin	University of Göttingen	Sociological Research Institute (SOFI)	ebabic@gwdg.de	X	
Baethge, Martin	University of Göttingen	Sociological Research Institute (SOFI)	mbaethg@gwdg.de	X	
Baethge-Kinsky, Volker	University of Göttingen	Sociological Research Institute (SOFI)	vbaethg@gwdg.de	X	
Bailey, Thomas	University of Columbia	Teachers College/Institute of Education and Economy	tb3@columbia.edu		X
Birzea, César	University of Bucarest	Institut of Educational Sciences	cesar.birzea@ise.ro		X
Brown, Alan	University of Warwick	Social Studies Building Warwick Institute for Employment Research	Alan.Brown@warwick.ac.uk	X	
Descy, Pascaline	CEDEFOP		PDE@cedefop.eu.int	X	
Grootings, Peter	European Training Foundation	Department: Eastern Europe and Central Asia/Villa Gualino	Peter.Grootings@etf.eu.int		X
Halász, Gábor	Országos Közokatási Intézet	National Institute for Public Education	halaszg@helka.iif.hu		X
Hämäläinen, Kauko	University of Helsinki	Palmenia Centre for Continuing Education	kauko.hamalainen@helsinki.fi		X

Hoffmann, Walter	Federal Ministry of Economics and Labour (BMWA)		walter.hoffmann@bmwa.bund.de	X	
Kalous, Jaroslav			jaroslav.kalous@chello.cz		X
Kauppi, Antti	University Helsinki	Palmenia Centre for Continuing Education	Antti.Kauppi@helsinki.fi	X	
Lopez, Alberto	CEREQ (Marseille)		lopez@cereq.fr	X	
Munk, Peter	Federal Ministry of Education and Research (BMBF)		Peter.Munk@bmbf.bund.de	X	
Nijhof, Wim J.	University of Twente	Department of Education	nijhof@edte.utwente.nl	X	
Olsen, Ole J.	University of Bergen	Department of Sociology	ole.j.olsen@sos.uib.no	X	
Paulsen, Bent	Federal Intitute for Vocational Education and Training (BIBB)		paulsen@bibb.de	X	
Renold, Ursula	Schweizerisches Bundesamt für Berufsbildung und Technologie (BBT)		ursula.renold@bbt.admin.ch		X
Schweri, Jürg	University of Bern	Centre for Research in Economics of Education	juerg.schweri@vwi.unibe.ch	X	
Sörensen, John Houman	Aalborg University	FIB 1-92A	houman@socsci.aau.dk	X	
Stevenson, John	Griffith University/Mt Gravatt Campus	Centre for Learning Research	J.Stevenson@griffith.edu.au	X	
Svetlik, Ivan	University of Ljubljana	Faculty of Social Sciences	ivan-svetlik@uni.lj.si		X
Wagner, Sandra	Federal Ministry of Education and Research (BMBF)		sandra.wagner@bmbf.bund.de	X	
Weber, Susanne	University of Göttingen	Institute of Business and Human Resource Education	sweber@wipaed.wiso.uni-goettingen.de	X	
Winther, Esther	University of Göttingen	Institute of Business and Human Resource Education	ewint@wipaed.wiso.uni-goettingen.de	X	

Appendix A2: Participants of the Second International Workshops on PISA-VET

Name	Institution	Department	Mail	Land
Achtenhagen, Frank	University of Göttingen	Institute of Business and Human Resource Education	fachten@wipaed.wiso.uni-goettingen.de	
Arends, Lena	University of Göttingen	Institute of Business and Human Resource Education	arends@wipaed.wiso.uni-goettingen.de	
Babic, Edvin	University of Göttingen	Sociological Research Institute (SOFI)	ebabic@gwdg.de	
Baethge, Martin	University of Göttingen	Sociological Research Institute (SOFI)	mbaethge@gwdg.de	
Baethge-Kinsky, Volker	University of Göttingen	Sociological Research Institute (SOFI)	vbaethge@gwdg.de	
Badescu, Mircea	European Training Foundation (Turin)	Department for Enlargement and South Eastern Europe	Mircea.Badescu@eft.eu.int	Romania
Bailey, Thomas	University of Columbia	Teachers College/Institute of Education and Economy	tb3@columbia.edu	USA
Bouder, Annie	CEREQ (Marseille)		bouder@cereq.fr	France
Brown, Alan	University of Warwick	Social Studies BuildingWarwick Institute for Employment Research	Alan.Brown@warwick.ac.uk	United Kingdom
Chiru, Mihaela	Institute of Educational Sciences	Department for Counselling and Continuing Learning		Romania
Coffield, Frank	London University	Institute of Education	F.Coffield@ioe.ac.uk	United Kingdom
Hamilton, Stephen F.	Cornell University, Ithaca, NY	Associate Provost for Outreach Professor Human Development	sfh3@cornell.edu	USA
Haapakorpi, Arja	University of Helsinki	Palmenia Centre for Continuing Education	arja.haapakorpi@helsini.fi	Finland

Hoffmann, Walter	Federal Ministry of Economics and Labour (BMWA)		walter.hoffmann@bmwa.bund.de	
Homs, Oriol	CIREM-Foundation (Barcelona)		oriol.homs@cirem.org	Spain
Munk, Peter	Federal Ministry of Education and Research (BMBF)		Peter.Munk@bmbf.bund.de	
Nijhof, Wim J.	University of Twente	Department of Education	nijhof@edte.utwente.nl	The Netherlands
Oates, Tim	Qualification and Curriculum Authority (QCA)	Research and Statistics	oates@qca.org.uk	United Kingdom
Olsen, Ole J.	University of Bergen	Department of Sociology	ole.j.olsen@sos.uib.no	Norway
Paulsen, Bent	Federal Intitute for Vocational Education and Training (BIBB)		paulsen@bibb.de	
Säljö Roger	University of Gothenburg	Department of Education	rsaeljo@gwdg.de	Sweden
Schweri, Jürg	Schweizerisches Bundesamt für Berufsbildung und Technologie (BBT)		juerg.schweri@vwi.unibe.ch	Switzerland
Sörensen, John Houman	Aalborg University	FIB 1-92A	houman@socsci.aau.dk	Denmark
Stevenson, John	Griffith University/Mt Gravatt Campus	Centre for Learning Research	J.Stevenson@griffith.edu.au	Australia
Svetlik, Ivan	University of Ljubljana	Faculty of Social Sciences	ivan-svetlik@uni.lj.si	Slovenia
Thang, Per-Olof	University of Gothenburg		per-olof.thang@ped.gu.de	Sweden
Verhoeven, Jef	University of Leuwen		Jef.Verhoeven@kuleuven.ac.be	Belgium
Wagner, Sandra	Federal Ministry of Education and Research (BMBF)		sandra.wagner@bmbf.bund.de	

Appendix B1: The concept of Generalized Work Activities (GWA)

The taxonomic paradigm that underlies the structure of the "Generalized Work Activities" construct is rooted in the primary foundation of modern psychology: "Behaviour in any setting is a function of stimuli (S), organism (O), and responses (R)" (Jeanneret, et al., 1999, p. 107). In the case of analysing the behaviour of individuals at work, the "O" represents the worker who is the receptor of the stimuli (S), and after processing of those stimuli, provides one or more responses (R):

- S—information that is received by the worker (i.e., the stimulus);
- O—mediation process performed by the worker; and
- R—action performed by the worker in response to the "processed stimulus.

Since the S-O-R model is limited when describing behaviour in a work setting, two factors where added:

- Interactions with others—work behaviour typically involves interactions and relationships with individuals, and
- work context—work behaviour occurs within a physical and social context that typically is described in terms of working conditions, interpersonal relationships, and structured job characteristics.

The intent of the model is to communicate that the interactive components (i.e., S-O-R), or information input, mental processes, and work output components, occur (a) while interactions with other people take place and (b) within a worker's and an organisation's work context (*Figure B1*). Factor analysis of data collected with worker-oriented job analysis questionnaires has provided considerable insights into the structure that underlies the domain of human work. Because the worker-oriented approach is not specific to technology or tasks, it permits an understanding of the general, cross-job structure that is not possible with a job-oriented job analysis methodology. Nomothetic questionnaires (PAQ: Position Analysis Questionnaire, OAI: Occupational Analysis Inventory, and JEI: Job Element Inventory) have been used to collect and measure the content of a wide spectrum of jobs across the domain of work.

As indicated previously, a number of taxonomic structures and job analysis research efforts were examined to develop both a model for the GWA construct as well as the definition and rating scale levels for each individual GWA to be included in the O*NET. The researchers began by selecting the GWA constructs, using several criteria:

- The construct should have a foundation in one or more research efforts;
- the construct should have definite underlying content that, for GWAs derived from factor analysis of job analysis data, was determined by examining the content of individual items with significant factor loadings on the factor of interest;
- the constructs as a set should be comprehensive, as much as possible reflecting work activities of all jobs in the U.S. economy; and

– the constructs should provide unique descriptive information.

By following such a strategy, matters of specificity are also addressed. GWAs can be expressed at a very broad level of generality or at successively narrower levels across the specificity-generality continuum. Clearly, the analysis of work at a more specific level will yield occupational information at a finer level of differentiation.

Finally, altogether 42 GWAs were identified. The structure of GWAs is presented below:

Information Input

Information input consists of two factors: "Looking for and receiving job-related information", and "identifying and evaluating job-relevant information". Research that used date obtained from the PQA, JEI, and OAI was the dominant source for defining the GWA constructs that are necessary to describe the scope of the *information input* domain: i) "getting information needed to do the job", and ii) "monitoring processes materials, and surroundings" were the two GWAs that describe activities of a worker when looking for and receiving job-related information.

Once information is received, it must then be identified or evaluated. Three GWAs consistently emerged in the research literature that describe "identifying and evaluating job-relevant information": i) "identifying objects, actions, and events", ii) inspecting equipment, structures, or materials", and iii) "estimating the characteristics of materials, products, events, or information".

Mental Processes

Once job-relevant input has been received, the worker's mental capabilities are involved and can be categorised as two factors: "Processing information or data", and "reasoning/decision making". Four GWAs were identified from the research literature that were descriptive of information/data processing activities: i) "judging the qualities of objects, services, or persons", ii) "evaluating information for compliance to standards", iii) "processing information", and iv) "analysing data or information".

These GWAs emerged from the analysis of all levels of work. On the other hand, when examining the reasoning/decision making factor, the research literature indicated that many of the relevant constructs were derived from studies of supervisory, managerial or professional work. More specifically, two GWAs: i) "making decisions and solving problems", and ii) "updating and using job-relevant knowledge", emerged from studies of all types and levels of jobs.

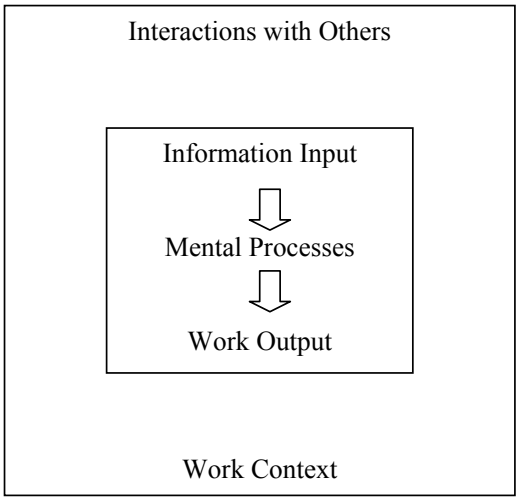
Alternatively, three other GWAs are more specific to managerial/professional work. These GWAs include: iii) "thinking creatively", iv) "developing objectives and strategies", and v) "scheduling work activities".

The complete GWA level structure is provided below (*Annex B3*). For each GWA, an operational definition and examples of high and low points on the level scale, are provided. The GWA "getting information needed to do the job" is defined as follows: "Observing, receiving, and otherwise obtaining information

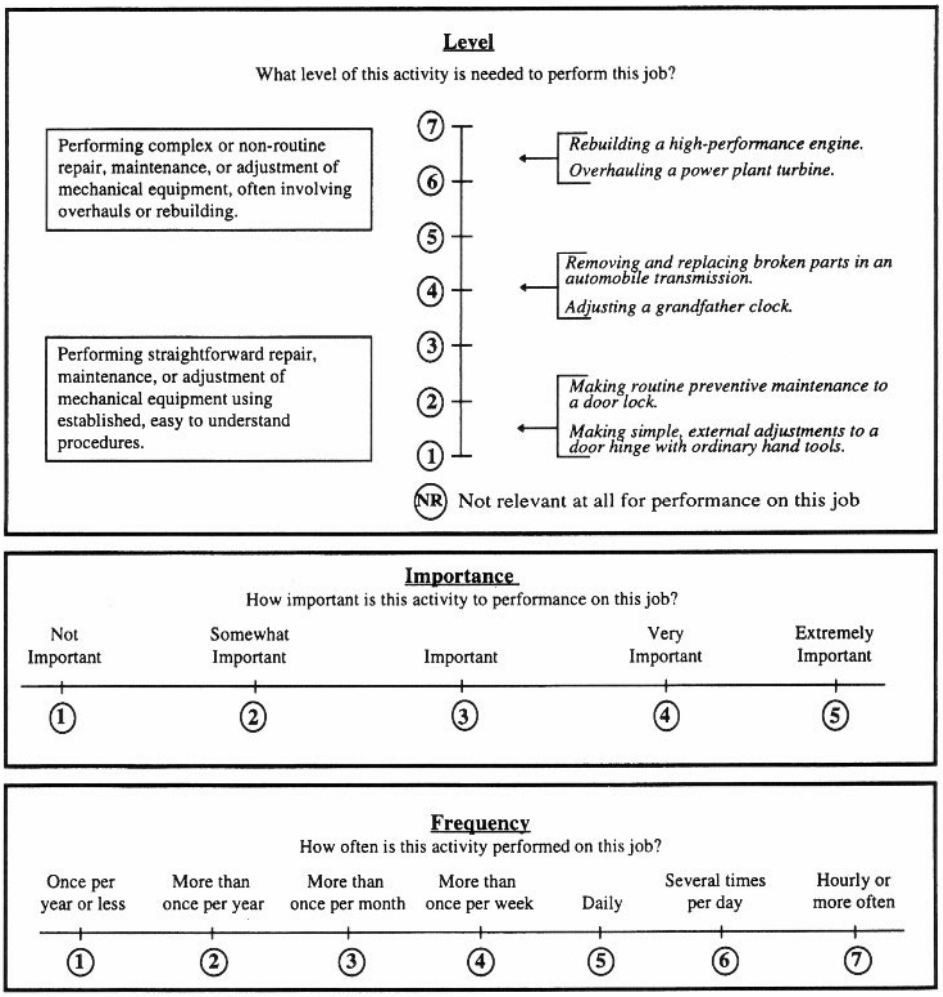
from all relevant sources". The scales consists of seven levels, the highest level is defined as "getting new information from many sources by actively interacting with the sources", whereas the lowest level refers to "making regular use of the same type of information from a single source".

For the GWA rating scales, three types of scales were developed: Level, Importance, and Frequency scales. Each GWA has a unique level scale definition, for measuring complexity. However, the scales are rather suitable for ratings of complexity across occupations, than for measurement within occupational fields. The level scales are 7-point scales (1-7) with an additional *not relevant* option. Three behavioural statements were also developed for each GWA level scale to anchor the high, mid-rage, and low levels. The *importance scale* has 5 points (1-5), with the verbal anchors: "Not important", "somewhat unimportant", "important", "very important", and "extremely important". The *frequency scale* is a 7-point scale with the verbal anchors: "Once per year or less", "more than once per year", "more than once per month", "more than once per week", "daily", "several times per day", and "hourly and more often". Both the importance and frequency scales are "within-job" ratings. *Annex B2* contains an example GWA descriptor, "repairing and maintaining mechanical equipment", and the corresponding rating scales.

Figure B1: The S-O-R Model



Annex B2: GWA "Repairing and Maintaining Mechanical Equipment"



Annex B3: Descriptions and definitions of Generalized Work Activities

Descriptions and Definitions of Generalized Work Activities

Construct label	Operational definition	Level scale
	Looking For and Receiving Job-Related Information	
1. Getting information needed to do the job.	Observing, receiving, and otherwise obtaining information from all relevant sources.	High—Getting new information from many sources, often by actively interacting with the sources. Low—Making regular use of the same types of information from a single source.
2. Monitoring processes, materials, or surroundings.	Monitoring and reviewing information from materials, events, or the environment, often to detect problems or to find out when things are finished.	High—Monitoring very complex processes, events, or circumstances. Low—Monitoring processes, events, or circumstances that are not complex.
	Identifying and Evaluating Job-Related Information	
3. Identifying objects, actions, and events.	Identifying information received by making estimates or categorizations, recognizing differences or similarities, or sensing changes in circumstances or events.	High—Making extremely difficult identifications based on very complex information. Low—Making easy identifications based on information that is not complex.
4. Inspecting equipment, structures, or materials.	Inspecting or diagnosing equipment, structures, or materials to identify the causes of errors or other problems or defects.	High—Making inspections or diagnoses of a complex system that may have many interrelated parts and determining whether conditions exist within a range of acceptable limits. Low—Making easy judgments about the quality or importance of things or people when there are many guidelines.
5. Estimating the characteristics of materials, products, events, or information.	Estimating sizes, distances, and quantities, or determining time, costs, resources, or materials needed to perform a work activity.	High—Making very difficult estimates of characteristics, time, or resources where there is limited guidance or supporting information. Low—Making straightforward estimates of characteristics, time, or resources where there is considerable guidance and supporting information.
	Information/Data Processing	
6. Judging the qualities of objects, services, or persons.	Making judgments about or assessing the value, importance, or quality of things or people.	High—Making very difficult judgments about the quality or importance of things or people for which there is limited guidance or supporting information. Low—Evaluating information against a simple criterion.
7. Evaluating information for compliance to standards.	Evaluating information against a set of standards and verifying that it is correct.	High—Evaluating complex information for compliance with regulations, laws, or technical criteria, where compliance decisions require significant interpretation or judgment. Low—Evaluating information against a simple criterion.
8. Processing information.	Compiling, coding, categorizing, calculating, tabulating, auditing, verifying, or processing information or data.	High—Processing very different and complicated data or information, where there are several ways in which the information can be processed. Low—Processing data or information that is standardized and easy to understand, where there is only one way to process the information.
9. Analyzing data or information.	Identifying underlying principles, reasons, or facts by breaking down information or data into separate parts.	High—Analyzing very different and complicated data or information that can be used for making critical decisions. Low—Analyzing data or information that is easy to understand.
	Reasoning/Decisions Making	
10. Making decisions and solving problems.	Combining, evaluating, and reasoning with information and data to make decisions and solve problems. These processes involve making decisions about the relative importance of information and choosing the best solution.	High—Reaching conclusions after considering a large number of choices that are often ambiguous or abstract, where there are competing viewpoints and alternatives that must be considered before reaching final decisions and the solutions decided upon will have very significant impact. Low—Reaching conclusions after considering a few choices that are usually well defined, where there are a limited number of possible actions, and the decisions or solutions will have minor impact.

Table 8-1 continues

Construct label	Operational definition	Level scale
11. Thinking creatively.	Originating, inventing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions.	High—Creating or inventing new and yet-to-be-proven practices, technologies, materials, products, or strategies, where the creative effort will have widespread impact and will result in substantial improvements for both an organization and its customers. Low—Offering suggestions for some change or improvement to immediate work functions or products.
12. Updating and using job-relevant knowledge.	Keeping up-to-date technically and knowing one's own job's and related jobs' functions.	High—Learning, retaining, and staying current with complex, often highly technical information. Low—Learning, retaining, and staying current with relatively easy-to-master information.
13. Developing objectives and strategies.	Establishing long-range objectives and specifying the strategies and actions to achieve these objectives.	High—Doing complex, future-oriented strategic planning. Low—Doing strategic or long-term planning that is not complex.
14. Scheduling work and activities.	Scheduling events, programs, activities, as well as the work of others.	High—Engaging in complex and difficult scheduling activities. Low—Engaging in simple or straightforward scheduling activities.
15. Organizing, planning, and prioritizing work.	Developing plans to accomplish work, and prioritizing and organizing one's own work.	High—Doing a high degree of complex planning, organizing, and prioritizing of one's own work. Low—Doing uncomplicated planning, organizing, or prioritizing of one's own work.
	Performing Physical and Manual Work Activities	
16. Performing general physical activities.	Performing physical activities that require moving one's whole body, such as in climbing, lifting, balancing, walking, stooping, where the activities often also require considerable use of the arms and legs, such as in the physical handling of materials.	High—Making repetitive and often fatiguing extensive use of the whole body in completing work activities that are done with or without the use of tools. Low—Making nonfatiguing use of the whole body in completing work activities that are done with or without the use of tools.
17. Handling and moving objects.	Using one's own hands and arms in handling, installing, forming, positioning, and moving materials, or in manipulating things, including the use of keyboards.	High—Using one's hands and arms to do the same functions almost continually. Low—Making little use of one's hands and arms.
18. Controlling machines and processes.	Using either control mechanisms or direct physical activity to operate machines or processes (not including computers or vehicles).	High—Controlling machines or processes that are very difficult to operate. Low—Controlling machines or processes that are easy to operate.
19. Operating vehicles, mechanized devices, or equipment.	Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.	High—Operating equipment or vehicles that are very difficult to run. Low—Operating equipment or vehicles that are easy to run.
	Performing Complex/Technical Activities	
20. Interacting with computers.	Controlling computer functions by using programs, setting up functions, writing software, or otherwise communicating with computer systems.	High—Using computers to develop very complex, high speed data linkages and operating systems. Low—Using computers to produce standard correspondence, graphic materials, and business-related information.
21. Drafting, laying-out, and specifying technical devices, parts, and equipment.	Providing documentation, detailed instructions, drawings, or specifications to inform others about how devices, parts, equipment, or structures are to be fabricated, constructed, assembled, modified, maintained, or used.	High—Drafting and specifying the components or technical relationships for complicated devices, parts, or equipment. Low—Drafting or specifying the components or technical relationships for devices, parts, or equipment that are easily understood.
22. Implementing ideas, programs, systems, or products.	Conducting or carrying out work procedures and activities in accord with one's own ideas or information provided through directions/instructions for purposes of installing, modifying, preparing, delivering, constructing, integrating, finishing, or completing programs, systems, structures, or products.	High—Performing highly complex and very difficult work activities with very limited guidelines to follow. Low—Performing activities that have clear-cut directions and are easy to carry out.

Table 8.1 continues

Construct label	Operational definition	Level scale
23. Repairing and maintaining mechanical equipment.	Fixing, servicing, aligning, setting up, adjusting, and testing machines, devices, moving parts, and equipment that operate primarily on the basis of mechanical (not electronic) principles.	High—Performing complex or nonroutine repair, maintenance, or adjustment of mechanical equipment, often involving overhauls or rebuilding. Low—Performing straightforward repair, maintenance, or adjustment of mechanical equipment using established, easy to understand procedures.
24. Repairing and maintaining electronic equipment.	Fixing, servicing, adjusting, regulating, calibrating, fine-tuning, or testing machines, devices, and equipment that operate primarily on the basis of electrical or electronic (not mechanical) principles.	High—Performing complex or nonroutine repair, maintenance, or adjustment of electronic equipment, where repairs are often made to complex internal components or circuitry. Low—Performing straightforward repair, maintenance, or adjustment of electronic devices or equipment using established, easy to understand procedures.
25. Documenting and recording information.	Entering, transcribing, recording, storing, or maintaining information in either written form or by electronic/magnetic recording.	High—Documenting or recording very complex information using new, unstandardized procedures. Low—Documenting or recording straightforward information using predetermined forms and procedures.
	Communicating/Interacting	
26. Interpreting the meaning of information for others.	Translating or explaining what information means and how it can be understood or used to support responses or feedback to others.	High—Making very difficult interpretations of information with limited, if any, guidance to follow. Low—Making easy interpretations of information with a high degree of guidance to follow.
27. Communicating with supervisors, peers, or subordinates.	Providing information to supervisors, fellow workers, and subordinates. This information can be exchanged face-to-face, in writing, or via telephone/electronic transfer.	High—Providing complex oral and written communications to others in the organization. Low—Providing straightforward oral or written communications to others in the organization.
28. Communicating with persons outside the organization.	Communicating with persons outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged face-to-face, in writing, or via telephone/electronic transfer.	High—Presenting complex oral and written communications to persons outside the organization. Low—Presenting routine and simple oral and written communications to persons outside the organization.
29. Establishing and maintaining interpersonal relationships.	Developing constructive and cooperative working relationships with others.	High—Developing very good interpersonal relationships with highly diverse individuals or stakeholders in difficult situations. Low—Developing very few working relationships with others.
30. Assisting and caring for others.	Providing assistance or personal care to others.	High—Providing care or assistance to others in highly stressful or difficult situations. Low—Needing to provide minimal help or assistance to others.
31. Selling or influencing others.	Convincing others to buy merchandise/goods, or otherwise changing their minds or actions.	High—Doing a lot of high-level persuading to accomplish work objectives, involving persuading a very difficult to convince audience. Low—Doing little persuading to accomplish work objectives, because there is little need to convince others in any area.
32. Resolving conflicts and negotiating with others.	Handling complaints, arbitrating disputes, and resolving grievances, or otherwise negotiating with others.	High—Handling complaints and negotiations in very challenging situations, involving complex matters and significant conflict and pressure. Low—Handling negotiations that involve very simple matters that are easily resolved or involve complaint-handling or negotiation.
33. Performing for or working directly with the public.	Performing for people or dealing directly with the public, including serving persons in restaurants and stores, and receiving clients or guests.	High—Handling interactions with the public, where the audience is hard to please or other conflict is involved. Low—Having little interaction with the public, or needing to have only brief interactions.
	Coordinating/Developing/Managing/Advising Others	
34. Coordinating the work and activities of others.	Coordinating members of a work group to accomplish tasks.	High—Coordinating the work of many employees, where a complex sequencing of others' tasks is required. Low—Needing to do little coordinating of others.

Table 8-1 continues

Construct label	Operational definition	Level scale
35. Developing and building teams.	Encouraging and building mutual trust, respect, and cooperation among team members.	High—Managing large teams and building cooperation among diverse team members toward accomplishment of highly complex or poorly defined activities/projects. Low—Doing little team building.
36. Teaching others.	Identifying educational needs, developing formal training programs or classes, and teaching or instructing others.	High—Teaching and explaining difficult tasks, concepts, or material and conducting complex training. Low—Doing little training or educating of others.
37. Guiding, directing, and motivating subordinates.	Providing guidance and direction to subordinates, including setting performance standards and monitoring subordinates.	High—Directing and motivating several organization members and building and maintaining morale in difficult or unpleasant work settings. Low—Doing little directing or motivating of subordinates.
38. Coaching and developing others.	Identifying developmental needs of others and coaching or otherwise helping others to improve their knowledge or skills.	High—Identifying effective ways of developing others to perform highly complex or difficult tasks and coaching them under these difficult conditions. Low—Doing little coaching or developing of others.
39. Providing consultation and advice to others.	Providing consultation and expert advice to management or other groups on technical, systems-related, or process-related topics.	High—Providing expert guidance on complex matters regarding the design, development, or implementation of major programs. Low—Providing little advice or consultation to others.
	Administering	
40. Performing administrative activities.	Approving requests, handling paperwork, and performing day-to-day administrative tasks.	High—Overseeing administrative activities for a large workforce, with a complex set of administrative procedures. Low—Doing very straightforward administrative activities.
41. Staffing organizational units.	Recruiting, interviewing, selecting, hiring, and promoting persons for an organization.	High—Overseeing the staff of a large and diverse workforce, with complex staffing needs. Low—Doing very straightforward staffing activities.
42. Monitoring and controlling resources.	Monitoring and controlling resources and overseeing the spending of money.	High—Monitoring and controlling a large number of resources, including managing a large budget. Low—Needing to do little monitoring or controlling of resources or money.

Annex C: VET-Programmes according to countries: Name of the programme, type of the Programme, (P = Pre-vocational, V = Vocational, G = General), ISCED-level of educational outcome, typical entrance age, and number of participants

Country	Name of VET Programme	Type of VET-Programme	ISCED Level	Typical Entrance Age	Number of Participants
Australia					
	Traditional apprenticeship	V	3b	15–18	
	Modern apprenticeship (long term)	V	3b	15–45	
	Upper secondary school-based VET (TAFE)	V	3	15–45	
Denmark					
	Basic vocational Education (EGU)	P	3c	16–20	2,000
	Social and health service assistant (SOSU)	V	3c	17–30	12,000
	Upper secondary higher technical/ commercial education	G	3a	16–17	34,000
	Individual organized Youth education	P	3	16–20	4,000
	Upper secondary VET (alternative)	V	3c	16–20	114,000

Germany					
	Basic vocational education (1 year)	V	3b	16–18	40,000
	Vocational schools	V	3b	17–18	70,000
	Vocational schools (vocational qualification)	V	3b	17–17	286,000
	Dual system	V	3b	16–18	1,337,000
	Healthcare schools		5b	19–20	
Finland					
	Upper Secondary VET-programmes (including apprenticeships)	V	3a	16–18	162,000
France					
	Vocational training for young people without qualification (CAP,BEP, BP)	V	3c	16–25	100,000
	Secondary education (2nd circle-CAP)	V	3c	15–16	80,000
	Secondary education (2nd circle-BEP)	V	3c	15–17	510,000
	Secondary education (2nd circle-CAP) combined with apprenticeship	V	3c	16–18	185,000
	Secondary education (2nd circle-BEP) combined with apprenticeship	V	3c	16–18	42,000
	Specific schools (health and social care)	V	3c	16–20	16,000
	Secondary education (2nd circle-BP) combined with apprenticeship	V	3b	18–20	20,000
	Secondary education (2nd circle-BP)-school-based programme	V	3b	18–19	150,000
	Specific schools (health and social care)	V	4a	18–20	
Netherlands					
	Different types of VET, including apprenticeship	V	3	16	449,00
	Basic vocational training, level 2	V	3c	16	
	Professional training, level 3	V	3c	16	
	Middle management training, level 4	V	3a	16	
Norway					
	Upper secondary vocational (including Apprenticeships)	V	3c	15–17	125,000
Romania					
	Vocational school	V	3c	15	
	Apprenticeship school	V	3c	15	
Switzerland					
	Basic VET (Anlehre)	V	3c	15–17	3,000
	Apprenticeship or full time schooling (2 years)	V	3c	15–17	20,000
	Apprenticeship or full time schooling (3 or 4 years)	V	3b	15–17	149,000
	Vocational education combined with apprenticeship (2 years)	G/V	3a	15–17	13,000
Sweden					
	Upper secondary school	G/V	3a	16	304,000
	Upper secondary school – individual programme	V	3c	16	20,000
	Upper secondary education for pupils with learning disabilities	V	3c	16–17	3,000
	Tertiary education (incl. Nurses)		5a	19–	
Slovenia					
	VET-programme of short duration	V	3c	15	
	VET-Programme (medium duration)	V	3c	15	

	-school type organisation				
	VET-Programme (medium duration) - Dual system/apprenticeship	V	3c	15	
Hungary					
	Secondary VET-school prematura	P	3a	14–15	188,000
	Vocational programmes preparing for NVQL examinations	V	3c	16–17	110,000
	College graduate education		5a	18–20	
	College first programmes		5a	18–20	
United Kingdom					
	GNVQ -Intermediate level	V	3c	15+	
	Activities leading to NVQ level 2	V	3c	16	
	Activities leading to NVQ level 1	V	3c	16	
	Traditional apprenticeship	V	3c	16–18	
	Work-based training	V	3c	16–17	
	Activities leading to NVQ level 3	V	3a	16+	
	GNVQ advanced level	V	3a	16+	
	Modern apprenticeship	V	3a	16–19+	
	Diploma in Higher Education (incl. Nurse training)		5b	18*	
USA					
	Vocational High Schools	P/N	?	16–17	
	Registered Apprenticeship	V	3c	16–?	
	School-based post-secondary VET (2-years college)	V	3c	18–20	

Sources: Eurostat/CEDEFOP, 2002; Office of Apprenticeship Training, Employer, and Labor Services, 2004; Lamb et al., 2003; NCVER, 2004

Annex D: Expertise by Steffen Kühnel: Sampling plan for the German pilot PISA-VET study

The outcomes of a PISA-VET obviously depend on the quality of the data as a basis for comparison. Data quality refers to relevant, valid, and reliable information, which is generalisable across countries.

In the feasibility study, different units of analysis at diverse levels are examined. The following discussion, is exclusively focused on the level of individual educational participants and VET-graduates. Two alternative sampling plans are presented for identifying generalisable information (on the basis of questionnaires for a particular number of individuals) and regarding the effectiveness of particular VET-programmes.

Focussing on the sub-group of VET-participants and -graduates, the following research questions must be addressed:

– Which competencies do participants develop during their vocational educational and trainy; and
– are they able to use these competencies in the labour market after completing VET?

In the feasibility study it is assumed that the development of competencies not only depends on national institutional settings of VET and the quality of

educational supply, but also on individual preconditions, such as motivation or knowledge before entering VET. The utilisation of competencies in turn, depends on the respective situational context factors. In particular with regard to an international comparison of VET, the development of competencies must be considered within the respective VET-contexts. With regard to the effective utilisation of competencies in a wide range of different situations, the impact of ongoing changes of occupational contexts, due to technical, scientific, social, and political developments, must be taken into consideration.

Panel design

The sampling plan for a panel design cannot be limited to VET; the entire educational pathway from VET entry to occupational performance must be incorporated. Thus, a longitudinal panel study, including information of the same individuals at the beginning, the end, and during occupational performance, is considered an optimal solution. This allows controlling varying preconditions of VET entrants, individual determinants of competence development during VET, and hence, a more valid analysis of influencing factors and the impact of acquired competencies in the labour market.

Panel mortality, however, is a serious problem of this design; it is not possible interviewing exactly the same individuals at the beginning and end of VET. Tracking occupational careers in the labour market will be even more difficult. In principal dropouts can have two consequences: The sample number decreases or results are biased.

In the case of sample dropouts, of a proportion of q ($0 < q < 1$), between two measurement points and an intended sample number of n at the final measurement point, n/(1-q) units of analysis must be interviewed at the penultimate measurement point.

Generally, in multiple-wave panels with K measurement points and q_k dropouts between the measurement points k and k-1, the required initial sample size n_1 for the first measurement point with an intended sample size n at the final measurement point is calculated as follows:

$$n_1 = \frac{n}{\prod_{k=2}^{K} 1-q_k} = \frac{n}{(1-q_2)\cdot(1-q_3)\cdot\ldots\cdot(1-q_K)}$$

The amount of dropouts depends on whether systematic panel maintenance, such as tracking changes of participants' addresses, establishing relationships, or providing incentives for keeping up participation, is conducted.

For example, in the GFM panel (Heitemeyer, et al., 2002)[29], from a total of 2,722 German participants (without migration background) initially interviewed

29 Heitemeyer, W. (2002). Deutsche Zustände. Vol. 1. Frankfurt a. M.: Surkamp.

in 2002, 1,175 were interviewed in 2004, and only 825 in 2004. Consequently, the dropout quota amounted to 56.8 percent between the first and the second, and 29.8 percent between the second and third measurement point[30]. Phone interviews were conducted in 1-year intervals; interventions for reducing dropout rates were not conducted. In contrast, very high efforts for panel maintenance were maintained in an election survey by Falter and colleagues. From a total of 3,000 initially interviewed participants, 1,600 participants were willing to continue, which corresponds to a dropout rate of 46.7 percent[31].

In the case of a panel survey for VET-participants, with the first measurement point at the beginning, the second at the end, and the third three years after completing VET, the dropout rate between the first and the second measurement point probably corresponds with the number of VET dropouts, whereas the dropout rate between the second and the third measurement point will be much higher. Intensive panel maintenance, such as regular contact and address tracking, is required to minimise dropout rates. Nevertheless, dropout rates up to 50 percent must be expected.

More problematic than the absolute number of dropouts, which can be absorbed by increasing the initial panel size, is the risk of systematic dropouts. With regard to VET-graduates this implicates that above average dropout rates of unsuccessful participants will result in misinterpretations of VET effectiveness, and vice versa for above average dropout rates of successful participants.

To minimise these error sources, dropout analysis are required to investigate systematic dropouts on the basis of available data. For systematic control, cross-sectional surveys in the same population are required in addition to repeated panel surveys.

Repeated panel surveys are notably more expensive than singular cross-sectional studies for the following reasons:

- A larger initial sample;
- additional costs for panel maintenance;
- additional cross-sectional surveys for control; and
- additional effort for dropout analysis.

Expenditures increase according to the number of measurement points and time intervals between them; reliable conclusions regarding the effectiveness of national VET-programmes can only be achieved on the basis of repeated panel surveys and it must be taken into consideration that relevant results are not available until the trial wave. For PISA-VET it can be assumed that the consortium is not willing to wait six years until reliable results are available.

Therefore, alternative measurement designs must be taken into consideration. One possibility is a rotating panel design: At one time point t_i independent surveys are conducted with diverse target populations in different samples—the

30 Typically, dropout rates are most significant between the first and second measurement point.

31 In reality, repeated interviewing must account for additional dropouts, which results in a total dropout rate of 50 percent after three years.

first sample consists of VET-entrants, the second of VET-graduates, and the third of individuals outside VET for the time of their training duration[32]. Measurement is repeated after the period of time corresponding to the respective training duration. At the same time, a new series of samples is defined. The advantage of this design is, that a comprehensive analysis of VET-effectiveness is available in a relatively short period of time at an individual level. In addition, it allows continuous observation of developments as a basis for initiating changes in education and VET-systems.

The preconditions for rotating panel designs are comparable training periods; if this is not the case, independent samples must be defined. A comparison of VET-effectiveness at the time of completion, is only possible on the basis of the smallest common multiple of training durations. For example, in the case of VET-programmes of one-, two-, and three-year duration (A, B, and C), the first measurement point t_1 refers to entrants of the VET-programmes A, B, and C. At the same time, independent samples of VET-graduates are defined in each VET-programme. An additional sample is defined, consisting of individuals having completed the VET-programmes A and B two or six years ago, and VET-programme C three or six years ago.

Another option for saving costs refers to defining a sample consisting of graduates in different VET-programmes (A, B, and C) six years before. One year later, measurement is repeated with those individuals having entered VET-programme A one year ago. Correspondingly, after two years, measurement will be repeated with those individuals having completed VET-programme B two years ago. Finally, after three years, individuals having entered a three-year VET-programme are interviewed, and correspondingly, those individuals interviewed initially after completing VET, would be interviewed again after two, three, and six years. The following table (*Figure D1*) illustrates this measurement design:

32 The design can be extended with additional samples of target populations that are outside VET for a much longer period of time than their own training period.

Figure D1: Rotating Panel Design

Sample	12	3	4	5	6	7	8	9	10	11	12	
VET Programme	A	B	C	A	B	C	A	B	C	A	B	C
VET Entrance	t_1	t_1	t_1									
VET Completion	t_2	t_2	t_2	t_1	t_1	t_1						
1 Year after A – Stop												
2 Years after A – Stop	t_3	t_3		t_2	t_2		t_1	t_1				
3 Years after A – Stop			t_3			t_2			t_1			
4 Years after A – Stop												
5 Years after A – Stop												
6 Years after A – Stop	t_4	t_4	t_4	t_3	t_3	t_3	t_2	t_2	t_2	t_1	t_1	t_1

A = VET programme one-year duration, B = VET programme two-year duration, C = VET programme two-year duration

t_i = measurement point i; all samples are defined at the same time point, time intervals between measurement points are indicated by the information provided in the first column.

Retrospective design

Retrospective designs can be considered a reasonable alternative to panel designs. They are limited to the population of employees. Information regarding previously attended VET-programmes is identified ex post, on the basis of interviews. The measurement, however, is limited to the present occupational competencies. On the basis of group comparisons, different competencies can be identified only on an aggregated level and consequently causal interpretations of influencing factors remain theoretical to a large extent.

Trend study design

Instead of panel designs trend studies are another option. The trend study design corresponds with the panel design illustrated above, with the exception that instead of repeated measurement with the same population, independent samples with different populations are defined. The advantage is that panel mortality is impossible and expenses are much lower. On the other hand, valid measurement results can only be identified at an aggregated level, and causal interpretations are as problematic as in retrospective designs. To conclude, trend study designs are recommendable for identifying differences of VET-programmes within and between countries. However, the reasons for the differences can not be explained.

Modularised design

As long as the overall expenses for PISA-VET can not be determined, a modularised design, which allows flexible combination of single modules according to the provided financial resources, is recommended. The measurement plan for rotating panels provides the starting point for the modularised design. In the simplest case only three-year VET-programmes are taken into consideration. Training can be school-based, firm-based, or a combination of the two.

Module 1: Survey of VET graduates six years after entering VET
In this module, the initial sample is S9a (*Figure D2*), consisting of VET-graduates six years after entering a VET-programme. For Germany, the sample refers to individuals having entered a three-year VET programme six years ago. Within this group, comparisons between apprentices of different three-year VET-programmes and dropouts of different VET-programmes are possible. Additional external information on VET-programmes, in particular with regard to dropout rates, allows the identification of a comprehensive picture of the effectiveness of different VET-programmes. A control group, comparable to sample 9, but consisting of employed people without VET-qualification, could provide additional reliability.

Module 2: Survey of apprentices enrolled in VET
The samples S3a and S6a are identified in addition to module 1. With regard to sample 3 apprentices should be interviewed ideally after concluding the article of apprenticeship or before entering VET. Module 2 allows detailed measurement of competence developments in VET and allows drawing inferences regarding prior competencies before entering VET in comparison to acquired competencies after completing VET. Assuming that competence allocation before entering VET and competence development during VET has not fundamentally changed during the past three years, a valid measurement of competence development during and after completing VET is possible. Moreover, it can be evaluated to what extent the competencies acquired in VET are relevant with regard to employment.

Module 3a: Repeated measurement after three years
In this module, the apprentices of samples S3a and S6a are interviewed again, which allows measurement of learning achievements in VET at an individual level (comparison of t_2 and t_1 in sample 3), as well as utilisation and transformation of competencies in the first three years of employment (comparison of t_2 and t_1 in sample 6). Assuming unsystematic panel dropouts, it can be analysed whether competencies after completing VET have changed within the three years. Implementing module 3, however, involves elaborated panel maintenance in the samples 3 and 6. An increase of the initial panel sizes samples 3 and 6 should be taken into consideration to assure enough participants in the second measurement.

Module 3b: Control samples after completing VET and three years later
In addition to module 3a two samples should be defined: One consisting of individuals having entered a VET-programme three years ago (new sample 6b), and another consisting of individuals having entered a VET-programme six years ago (new sample 9b). The new samples allow systematic panel control of dropouts in module 3b and measurement of competence developments during

VET (comparison of samples 6a and 6b) as competence transformations in the job (development of professional expertise) (comparison of samples 9a and 9b).

Module 3c: Control sample at the beginning of VET
In addition to module 3a one sample consisting of VET-entrants should be defined (new sample 3b) for valid measurement of the changes in preconditions for VET.

Module 4: Repeated measurement after six years
The respondents of sample 3a will be measured six years later for the third time, and if necessary, samples 6b and 3b for the seconds time. For sample 3a, a complete track from the beginning of VET until six years later is available now. In addition, for two age cohorts with a difference of three years, information of individual competence development during VET can be provided.

Module 5: Control samples at the end of VET and three years after completion
In addition to repeated measurement, control samples must be defined, three (S6c) and six years (S9c) after VET-entrance. This allows control of panel mortality and a comparison of three cohorts at a time; at the beginning, at the end, and three years after completion. The following table summarises the modularised design

Figure D2: Modularised design

Start of Measurement	+ Three Years	+ Six Years
Module 1: S9a (Six Years after A-B)		
Module 2: S6a (Three Years after A-B)	M 3 a: S6a (W2, Six Years after A-B)	
Module 2: S3a (VET Entrance)	M 3a: S3a (W2, Three Years after A-B)	M 4: S3a (W3, Six Years after A-B)
	M 3b: S9b (Six Years after A-B)	
	M3b: S6b (Six Years after A-B)	M 4: S6b (W2, Six Years after A-B)
	M3c: S3b (A-B)	M 4: S3b (W2, Three Years after A-B)
		M 5: S9c (Six Years after A-B)
		M 5: S6c (Three Years after A-B)

The illustrated design can easily be expanded to a longer period of time. Moreover, it is possible to track individual occupational careers for a longer time—respondents of sample 9a are measured again three and six years later, and sample 9b six years after VET-entrance.

Recruitment

The question of sampling methods is still open in the measurement design. Since the research objective refers to a comparative evaluation of VET, the target population should comprise individuals of different VET-programmes. However, given the multitude of VET-programmes only in Germany, it is recommended to focus on typical occupational fields for the identification of VET-outcomes.

In the feasibility study the following occupational fields have been selected: Industrial/technical occupations in industry and trade, commercial and comercial/administrative occupations in commerce and other services, social and healthcare occupations in the field of personal services, and information and communication technology in the field of information/technical services. With regard to a sampling plan for a German pilot study this entails including the following VET-programmes: Nursing (Krankenschwester/Krankenpfleger); car mechatronic (KFZ-Mechatroniker) formally auto mechanic (KFZ-Mechaniker) and auto electrician (KFZ-Elektriker); information technologist (Informationselektroniker); and office occupations (Bürokauffrau/ Bürokaufmann).

In principal, the target population can be recruited on the basis of household samples. Provided an appropriate selection, a representative sample, *within* differrent occupational fields as well as *between* particular occupations, at the beginning and the end of VET as well as after occupational entry can be identified. However, to get a sufficiently large sample for every occupational field, a very large initial sample must be defined which comes along with a number of individuals outside the target population irrelevant for measurement.

Joining existing samples, such as the socio-economic panel (Sozio-ökonomisches Panel, SOEP), and supplementing them with additional measurements that are in line with the respective studies, is another option. It must be examined, however, whether the samples of the respective studies are large enough. If this is not the case, representativeness at an individual level cannot be ensured. Identifying the target population on the basis of a large access panel of survey institutions is another option.

Another disadvantage of recruitment based on household panels is that individuals are related to a number of different institutions and firms. The central research questions of the feasibility study, however, refers to the identification of institutional context factors. Therefore, in addition to measuring individual characteristics of the target population, information on relevant educational institutions and firms must be identified. Typically, in household samples the relevant addresses of institutions are identified and consolidated in an institution sample. It must be anticipated, however, that individual information based on memories can be incomplete and incorrect. Therefore, it is possible that the number of institutions exceeds the number of individuals. The high efforts of in household panels suggest not implementing this method for the purpose of a comparison of VET.

To conclude, a sampling plan consisting of at least two stages is recommended for a PISA-VET; instead of identifying households in a first step, educational providers/firms are identified first, the corresponding participants in the VET-programmes/employees in a second step. In the case of random sampling of institutions this sampling plan is also representative for affiliates of institutions.

Individuals of the target population outside institutions (educational providers/firms) are not incorporated. With regard to a PISA-VET, this refers only to those individuals, which are currently unemployed but having entered or completed a VET-programme corresponding to the selected occupational fields. Thus, it is recommended to supplement the surveys with participants, dropouts, and graduates of VET identified in educational institutions/firms, with an unemployment analysis of individuals in the respective occupational fields. For recruitments of the target population a random sample of schools/firms, based on interviews with VET-participants in schools or VET-graduates working in firms, is recommended.

For participants in the dual German VET system it must be considered, whether respondents should be recruited in firms or schools. The advantage of recruitment in schools is that a large number of individuals can be interviewed within a short period of time. However, it is possible that schools are not ready to provide their instruction time for the surveys. In this case, surveys must be conducted after school in the classrooms, which requires incentives for students to participate.

Provided that competencies in VET require an infrastructure of training firms, the survey must be conducted in these firms. This also refers to the question of incentives for apprentices and firms. Respondents of firm-based surveys can be recruited either in schools or firms. The advantage of sampling in firms is that dropouts and workers without VET-qualification, performing comparable tasks can be easier identified. For a PISA-VET, recruitment of individuals in firms is recommended. For firm selection, defining a sub-sample on the basis of the firm panel provided by IAB is another option. Smaller firms and trades, however, should be identified on the basis of chambers and comparable participating firms.

Stratification

For the selection of firms, stratification according to the IAB firm-panel requirements, represents a reasonable option. Within strata, selection probability must be determined in a way that the same number of firms is contained in each stratum. If there is no adjustment at the second stratum, employees in larger firms have a higher chance to be included in the sample.

Sample size

Conclusions regarding the required sample size are problematic without information on methods of analysis, variations of target population characteristics, and aimed accuracy.

As a rough rule of thumb the required sample size for dichotomous items with maximal variance can be based on a 5 percent error rate and margin of error. In the case of simple random sampling this results in a sample of about 400 individuals, with 5 percent accuracy for about 1,100 individuals. In general, the sample size for simple random sampling can be calculated using the following model[33]:

$$n = \frac{\left(z_{1-\alpha/2}\right)^2 \cdot s^2}{\varepsilon^2}$$

$z_{1-\alpha/2}$ denotes the 95%-quantile of the normal distribution, α the error rate, s the standard deviation of the sample, and ε the maximal acceptable margin of error.

In the case of cluster sampling intra-class correlation must be taken into consideration; this refers to the fact that respondents within one cluster (in this case firms and VET-institutions) are more alike than respondents from different clusters. As a rule of thumb the following formula can be used for sample calculation:

$$1 + (M - 1) \cdot ICC$$

ICC denotes intra-class correlation, and M the cluster size. The formula indicates the required factor for increasing the sample to a size that includes all relevant individuals of the relevant institutions.

For example, in the case of an average intra-class correlation within school classes of 0.1 of the size M = 20 students and with surveys conducted in each class with each individual, a factor of $1+19 \times 0.1 = 2.9$ results. In the case of simple random sampling, 1,160/3,190 students must be interviewed, which results in a sample of 55/160 classes.

For implementing modules 3 and 4 the sampling plan requires repeated measurement. On the basis of average dropout rates for each VET-programme, panel mortality can be estimated. For example, in the case of a 20 percent dropout rate, the sample size in cross-sectional surveys must be increased by the factor 1.25 and consequently, 400/1,100 instead of 500/1,375 respondents must be interviewed.

In the case of repeated measurement three years after completing VET, a panel mortality of 50 percent can be expected, even with accurate panel maintenance (see above). Thus, the sample size of respondents, initially interviewed after completing VET, should comprise twice as much individuals as required; which means instead of 400/1,100 the sample comprises 800/2,200 individuals.

33 Sharon L. Lohr (1999). Sampling: Design and analysis. Bonn, et al.: Duxbury Press.

To provide generalisable conclusions, a separate sample should be defined in each occupational field, with a 5 percent margin of error. In the case of simple random sampling, 400 individuals are interviewed in each group, which finally amounts to a total of 1,600 individuals.

Module 1 requires a cross sectional survey of individuals six years after VET-entrance. The sample consists of employees of approximately the same age and performing the same or comparable working tasks. On the basis of secondary data the percentage of employers without VET-qualification, or transferred from a different occupation, can be analysed. For example, in the case of 25 percent, the initial sample number must be increased by a factor of 100 percent in addition to the respective percentage, resulting in a factor of 1.25, and sample number of 500 individuals for each occupational field.

In comparison to school classes, lower intra-class correlation, of about 0.02, can be assumed across different firm sizes. For calculating the adjustment factor, the (weighted) mean of employees in the respective age group included in the respective occupational field must be identified across all firms. For example, for a value of 5, an adjustment factor of $1+(5-1) \times 0.02=1.08$ results and conesquently, the sample number increases from 500 to 540 respondents in each occupational field.

The number of identified firms depends on the mean of individuals belonging to the target group and working in firms of the required size. For a value of 5 across all firm sizes, approximately 108 firms must be identified for each occupational field. Sample 9a of module 1, requires a sample of 432 firms with a total of 2,160 respondents across all occupational fields. The calculation of sample numbers for S3a and S6a depends on whether recruiting is conducted in firms or schools. In firms, sample numbers can be calculated in the same way as in S9a; identifying the average number of individuals who are employed in their firms at the beginning and at the end of their training period, however, is a suspect matter. In the case of 5 individuals in each firm and an ICC of 0.02, a sample size of 432 individuals at the beginning and additionally 432 individuals at the end of each VET-programme results.

In schools, sample numbers are calculated differently; for example, in the field of nurses, a higher number of individuals within the clusters and higher intra-class correlation can be expected. Based on the above assumptions (ICC=0.1, and an average of 20 individuals in each class) the number of respondents at the beginning and at the end of VET amounts to 1,160 individuals in 58 classes each. For each of the samples S3a and S9a, a total of 2,456 respondents from 260 firms and 58 classes result.

For repeated measurement of S3a and S6a after three years, the sample number of S3a, with an assumed dropout rate of 20 percent, would increase to 3,070 individuals at the beginning of VET, and with an expected panel mortality of 50 percent, increase to 4,912 individuals at the end of VET. A measurement

design incorporating the entire modules, results in a sample depicted in *Figure D3*:

Figure D3: Estimated sample numbers in the sub-samples

Beginning of Measurement	+ 3 Years	+ 6 Years
S9a(6 years after A-B) n=2160		
(3 years after A-B) n=4,912/2,456	(W2, 6 years after A-B) n=2,456	
(VET entrance) n=3,070/2,456	(W2, 3 years after A-B) n=2,456	(W3, 6 years after A-B.) n= 1,228
	S9b (6 years after A-B) n=2,160	
	(3 years after (A-B) n=/2,456	(W2, 6 years after A-B) n=1,228
	S3b (A-B) n=2,456	(W2, 3 years after A-B) n=1,228
		(6 years after A-B) n=2,160
		(3years after A-B) n=2,456

In the case of two sample numbers: First number with balancing for panel mortality, second number without balancing.

Interview expenses

Calculating interview expenses without additional information is very difficult. Assuming that measurement requires four hours for each respondent, two interviews of two hours each must be conducted at each measurement point. For each respondent about 120 Euro for interviews and additionally fixed costs of about 75,000 Euros in each sample for pre-tests, the development of measurement tools, data analysis, and documentation management of measurements must be calculated[34]. With regard to the modules the following estimations result:

34 Rough estimations on the basis of offers by professional survey institutions for standardised surveys of one-hour duration for 2,000 to 3,000 respondents.

Figure D4: Estimated measurement expenses in the sub-samples:

Beginning of Measurement	+ 3 Years	+ 6 Years
n=2,160; 334,200 €		
a) n=4,912; 664,440 €	n=2,456; 369,720 €	
b) n=2,456; 369,720 €		
a) n=3,070; 443,400 €	n=2,456; 369,720 €	n= 1,228; 222,360 €
b) n=2,456; 369,720 €		
	n=2,160; 334,200 €	
	n=2,456; 369,720 €	n=1,228; 222,360 €
	n=2,456; 369,720 €	n=1,228; 222,360 €
		n=2,160; 334,200 €
		n=2,456; 369,720 €
between 1;073,640 € and 1.442.040 €	1.813.080 €	1,371,000 €

In addition, expenses for the definition of the samples arise. Their final amount depends on whether there are lists of institutions that are relevant for the sample and information regarding their interest in participating. Additional costs, such as panel maintenance, repeated measurement, and possible incentives for respondents or individuals, are not included.

Limitations of results

The proposed design provides opportunities for representative analysis of VET in the depicted occupational fields. Without additional analysis of unemployment risks, the outcomes are exclusively focused on individuals working in one of the occupational fields or apprentices striving for employment in those fields.

Moreover, the design is limited to an analysis of occupational success of those individuals transferring to an occupational field outside the selected occupational fields, after completing VET. Measurement only accounts for transitions from sample S3a/b and S6a/b. However, this limitation could be solved by expanding the occupational fields.

The growing together of the European labour markets proves to be a genuine challenge for education, particularly for VET-systems. VET-Systems have to provide future employees with the necessary qualifications that they are well equipped for flexibility, mobility and challenges of international cooperation on the labour market. If today's VET-systems are able to impart these key qualifications is unknown.

Therefore, the German Government (BMWA) initiated a feasibility study on VET – modelled on the basis of PISA research studies. The concept of a PISA for vocational education and training presented in this book was developed by leading international experts in the field. It shows how a measuring of expertise and competence is actually possible. It describes the criteria to be used for the evaluation of training programs, and discusses how the efficiency of different learning places are to be evaluated.

This study explores the conditions and processes necessary for an improvement of vocational education and training in the participating countries. It thereby provides a significant contribution for the political debate, but also for the academic discussion, for instance on the establishing of a European system of VET.

www.steiner-verlag.de

Franz Steiner Verlag

ISBN-10: 3-515-08968-3
ISBN-13: 978-3-515-08968-5