

Measurement and evaluation of competence

Gerald A. Straka

In:

Descy, P.; Tessaring, M. (eds)

The foundations of evaluation and impact research

Third report on vocational training research in Europe: background report.
Luxembourg: Office for Official Publications of the European Communities, 2004
(Cedefop Reference series, 58)

Reproduction is authorised provided the source is acknowledged

Additional information on Cedefop's research reports can be found on:

http://www.trainingvillage.gr/etv/Projects_Networks/ResearchLab/

For your information:

- the **background report** to the third report on vocational training research in Europe contains original contributions from researchers. They are regrouped in three volumes published separately in English only. A list of contents is on the next page.
- A **synthesis report** based on these contributions and with additional research findings is being published in English, French and German.

Bibliographical reference of the English version:

Descy, P.; Tessaring, M. *Evaluation and impact of education and training: the value of learning*. Third report on vocational training research in Europe: synthesis report. Luxembourg: Office for Official Publications of the European Communities (Cedefop Reference series)

- In addition, an **executive summary** in all EU languages will be available.

The background and synthesis reports will be available from national EU sales offices or from Cedefop.

For further information contact:

Cedefop, PO Box 22427, GR-55102 Thessaloniki

Tel.: (30)2310 490 111

Fax: (30)2310 490 102

E-mail: info@cedefop.eu.int

Homepage: www.cedefop.eu.int

Interactive website: www.trainingvillage.gr

Impact of education and training

Preface

The impact of human capital on economic growth: a review

Rob A. Wilson, Geoff Briscoe

Empirical analysis of human capital development and economic growth in European regions

Hiro Izushi, Robert Huggins

Non-material benefits of education, training and skills at a macro level

Andy Green, John Preston, Lars-Erik Malmberg

Macroeconometric evaluation of active labour-market policy – a case study for Germany

Reinhard Hujer, Marco Caliendo, Christopher Zeiss

Active policies and measures: impact on integration and reintegration in the labour market and social life

Kenneth Walsh and David J. Parsons

The impact of human capital and human capital investments on company performance Evidence from literature and European survey results

Bo Hansson, Ulf Johanson, Karl-Heinz Leitner

The benefits of education, training and skills from an individual life-course perspective with a particular focus on life-course and biographical research

Maren Heise, Wolfgang Meyer

The foundations of evaluation and impact research

Preface

Philosophies and types of evaluation research

Elliot Stern

Developing standards to evaluate vocational education and training programmes

Wolfgang Beywl; Sandra Speer

Methods and limitations of evaluation and impact research

Reinhard Hujer, Marco Caliendo, Dubravko Radic

From project to policy evaluation in vocational education and training – possible concepts and tools. Evidence from countries in transition.

Evelyn Viertel, Søren P. Nielsen, David L. Parkes, Søren Poulsen

Look, listen and learn: an international evaluation of adult learning

Beatriz Pont and Patrick Werquin

[Measurement and evaluation of competence](#)

Gerald A. Straka

An overarching conceptual framework for assessing key competences. Lessons from an interdisciplinary and policy-oriented approach

Dominique Simone Rychen

Evaluation of systems and programmes

Preface

Evaluating the impact of reforms of vocational education and training: examples of practice

Mike Coles

Evaluating systems' reform in vocational education and training. Learning from Danish and Dutch cases

Loek Nieuwenhuis, Hanne Shapiro

Evaluation of EU and international programmes and initiatives promoting mobility – selected case studies

Wolfgang Hellwig, Uwe Lauterbach, Hermann-Günter Hesse, Sabine Fabriz

Consultancy for free? Evaluation practice in the European Union and central and eastern Europe Findings from selected EU programmes

Bernd Baumgartl, Olga Strietska-Ilina, Gerhard Schaumberger

Quasi-market reforms in employment and training services: first experiences and evaluation results

Ludo Struyven, Geert Steurs

Evaluation activities in the European Commission

Josep Molsosa

Measurement and evaluation of competence

Gerald A. Straka

Abstract

Development, measurement and evaluation of competence has become an important issue in education and training. Various factors contribute to that:

- (a) the shift of focus in education from input to output, stimulated by standard-based assessment; accountability systems;*
- (b) international comparisons of school system achievement;*
- (c) the transition from subject to literacy orientation in education;*
- (d) recognition of learning in non-formal and informal settings, the vision of lifelong learning as a prominent goal of the European Union (EU).*

Nevertheless, electronic literature research into vocational education and training (VET) using, in combination, the keywords 'measurement', 'evaluation' and 'competence' generated a negligible number of references. Better results were obtained using 'competence' and 'evaluation'. This may indicate that the link between measurement, evaluation and competence is still absent in European VET research.

Measurement requires deciding what is to be measured. In order to provide specification, a general diagnostic framework is introduced. It consists of three properly differentiated levels: external conditions (e.g. situations, products), actual episodes realised by an individual (e.g. behaviour, cognitive operations, individually created information, motivation), and personal internal conditions (e.g. knowledge, skills, motives). Within this conceptualisation measurement faces the problem that knowledge, skills, and cognitive operations are not visible to outsiders. The only way to achieve insight is by observation of specified external conditions and/or behaviour (e.g. realised behaviour, changed situations, or created products). From this observation, specified/defined elements of the internal conditions (e.g. knowledge, skills) are inferred. The relations between the observable and non-observable are established through interpretation rules and hypotheses.

Evaluation, in this context, is judging the observed competences against defined benchmarks. Such benchmarks may be measured knowledge, skills, actions, or performance of other persons (norm-referenced); or they may be theoretically specified types and levels of knowledge, skills, actions, or performances (criterion-referenced). Evaluation and measurement themselves are subsumed under the term assessment.

Against the background of this general diagnostic framework, selected competence definitions in the EU are analysed. Most of them bridge all the three levels of the model. Such a broad and multilevel concept of competence might generate more misunderstanding than understanding in public and scientific discussions. Therefore, two recommendations are made on how to define and assess competence:

- (a) an accurate description of the tasks and requirements (= external conditions);*
- (b) a specification/characterisation of the psychic attributes a person should possess or that has built up in a specific occupational domain.*

Selected procedures of assessing competence in the EU are introduced:

- (a) the bilan de compétences (France);*
- (b) the national vocational qualifications (NVQ) (UK);*
- (c) dimensions of action competence in the German dual system;*
- (d) assessing competences at work (the Netherlands);*
- (e) realkompetanse (Norway);*
- (f) recreational activities (Finland);*
- (g) competence evaluation in continuing IT-training (Germany).*

The analysis of these conceptions using interrelated assessment quality criteria (validity, reliability, objectivity, fairness and usability) revealed that, to date, empirically grounded findings about fulfilling these criteria are scarce. Furthermore, methodological considerations indicate that self and other observation, even guided by criteria and realised with more than one external assessor, produce many errors.

The validity of the approaches varies considerably when it comes to diagnosing occupational competence and occupational success in general. The discussions often focus on the format of the tasks (closed versus open-ended). In this respect, the advantages of performance-based assessment through open-ended tasks requiring complex skills (i.e. involving a significant number of decisions) should not be overestimated. The problem is that the increased cost of evaluation is not justified by the increase in external validity. In addition, there is evidence that the key is not the format but rather the content requirements of the task.

Concerning the timing of assessments (e.g. continuously, sequential or punctual) valid evidence can only be obtained through systematic empirical investigations of assessment procedures in relation to concepts of competence development.

In addition, there are methodological reasons to supplement self- and peer assessment with standard oriented assessment and accountability considerations. Assessment should be done with regard to criteria, such as those proposed by the American Educational Research Association (AERA).

On the basis of these findings, the following recommendations are given to correspond with the EU goals of transparency and mobility:

- (a) initiating a detailed and summarising review of the diverse practices of competence assessment in the EU, focusing on the methodological dimension;*
- (b) promoting empirical investigation about the measurement quality of selected and prototypical assessment procedures practised in the EU;*
- (c) activating conceptual and empirical research on how to define and validate competence and its development in VET;*
- (d) advocating a VET-PISA in selected occupations or sectors under the patronage of the EU.*

Table of contents

1. Introduction	267
2. Measurement and evaluation	270
2.1. A general model of measurement	270
2.2. Evaluation	272
3. Competence	274
3.1. General competence categorisations	274
3.2. VET competence definitions	275
4. Procedures of assessing competence	277
4.1. The bilan de compétences (France)	278
4.2. The national vocational qualifications (NVQ) (UK)	279
4.3. Assessing action competence in Germany	281
4.4. Social and methodical dimensions of competence (DaimlerChrysler)	283
4.5. Competence at work: a case in the Netherlands	284
4.6. The realkompetanse project (Norway)	285
4.7. Recreational activities (Finland)	287
4.8. New ways of assessing competences in information technology (IT) in initial and continuing VET in Germany	288
4.9. Analysis and evaluation	290
5. When should competence be assessed and by whom?	294
6. Summary and conclusion	298
List of abbreviations	301
Annex 1: Sample overall training plan (Bundesanzeiger, 1998, pp. 8-9)	302
Annex 2: Sample of skeleton school syllabus (Bundesanzeiger, 1998, p. 18)	303
Annex 3: Unit 3 – sell financial products and services (NVQ Level 3, n.d.)	304
References	307

List of tables and figures

Tables

Table 1:	Comparisons of concepts	275
Table 2:	Validity of customary approaches	292
Table 3:	Performance-based assessment versus 'classical' approaches	293

Figures

Figure 1:	General diagnostic framework	271
Figure 2:	The structure of a NVQ	279
Figure 3:	Overview of the structural features of the dual system	281
Figure 4:	Excerpt from an observation and evaluation sheet for client-counselling of bank clerks	283
Figure 5:	Rating scale for the key competence 'ability to cooperate'	284
Figure 6:	Competence card from the workplace	286
Figure 7:	Example page of the recreational activity study book	288
Figure 8:	Structure of the IT-further education system	289
Figure 9:	From a practice to a reference project – an example for a network administrator	289
Figure 10:	General model of economic processes, a basis of business administration	295

1. Introduction

Competence and its evaluation have become an important issue in both the theory and practice of general education around the globe. This trend is not restricted to vocational education and training (VET). Several triggers can be identified of which two deserve emphasis:

- (a) a shift from input to output orientation in education through implementation in the US of a standard-based assessment and accountability system from the 1990s onward (Linn, 2000);
- (b) large-scale international achievement comparisons of school systems in different domains such as the *Third international mathematics and science study* (Baumert et al., 1997) ⁽¹⁾, or the *Programme for international student assessment* (Deutsches PISA-Konsortium, 2001) ⁽²⁾. Compared with former studies of this type, such as the first (1964) and second (1980-82) international mathematics studies, much more attention was given to the results of PISA 2000. Because PISA was replicated in 2003 with another focus – results will be published in 2004 – and will again be repeated in 2006 it is expected that the public awareness and discussion of measurement and evaluation of competence will continue.

TIMSS and PISA attest to a substantial shift in the generic aims of education and training. The focus of TIMSS was the contents of the mathematics and science education in schools of the participating nations. In contrast to this ‘world curriculum approach’ for the two subjects, the theoretical foundation of PISA is a ‘literacy concept’ not only in reading but in science and mathematics too. This concept additionally takes into account that the acquisition of literacy occurs not only during formal instruction but also in informal school settings, out of school, and

lifelong. ‘PISA assesses the ability to complete tasks relating to real life, depending on a broad understanding of key concepts, rather than assessing the possession of specific knowledge’ (OECD, 2001, p. 19). Such a perspective has an impact on the assigned situations and tasks as well as on the measurement models chosen.

In the US there was from the 1990s substantial pressure on the development and use of ‘new’ approaches to assessment under different labels such as alternative assessment, authentic assessment, direct assessment, or performance-based assessment (Linn and Gronlund, 2000). Against this background Lorrie Shephard ⁽³⁾ (2000) analyses and compares the curriculum, the learning and the measurement theories dominant in the 20th and at the beginning 21st century. At first, curriculum was guided by social efficiency. The concept of an innate, unitary, fixed IQ, combined with associationist behavioural learning theories led to sorting of students by ability; this ability was measured through objective achievement tests. Now, the curriculum is starting to be student-, problem- and authenticity-centred. It is in line with cognitive-constructivist learning theories. It uses formative assessment, addressing both outcomes and process of learning, and such assessments are used to evaluate both student learning and teaching. These findings suggest that assessment is interconnected with the curriculum concept and its learning theoretical basis.

Almost at the same time, two committees of the Commission on Behavioural and Social Sciences and Education of the National Research Council (NRC) of the US summarised the state of research on the science of learning and linked it with actual practice in the classroom in *How people learn* (Bransford et al., 2000). This report was the trigger for the Committee on the Founda-

⁽¹⁾ Under the auspices of the International Association for the Educational of Educational Achievement (IEA). The sample of TIMSS III focussing on 12th or 13th year of schooling included VET students in Germany. Their performances were analysed in relation to occupations (Watermann and Baumert, 2000).

⁽²⁾ Launched by the OECD in 2001.

⁽³⁾ Presidential address *The role of assessment in a learning culture*, annual meeting of the American Educational Research Association (AERA).

tions of Assessment to review and synthesise advances in the cognitive sciences and measurement as well as exploring their implications for improving educational assessment. The report – *Knowing what students know* – ‘addresses assessments with the three purposes: to assist learning, to measure individual achievement, and to evaluate programs. The conclusion is that one type of assessment does not fit all purposes’ (NRC, 2001, p. 1). As a consequence, measurement and evaluation of competence have to take in account their purposes, the concepts of learning and development.

In European VET, analogous developments took place – although mostly without explicit reference to the discussion in the US – of which examples are:

- (a) in 1987, in Germany ‘action orientation’ (*Handlungsorientierung*) was designated as the central objective of in-company training within the dual VET system. The school side adopted this objective in 1991 and, since 1996, the mission to generate ‘action competence’ (*Handlungskompetenz*) is the central pedagogical task of the schools in the dual system (KMK, 1996/2000). Linked with this development, ‘new’ approaches to assessment and examinations are requested, tested, or introduced;
- (b) independently from the German development, the national vocational qualifications (NVQ) were introduced in England and Wales in 1987 with ‘a very particular approach to defining and measuring competencies’ (Wolf, 1998, p. 207). In contrast to Germany, the NVQ concept explicitly refers to the competence-based training and assessment of the US (Wolf, 1995);
- (c) in 1991, *the bilan de compétences* was legalised and introduced in France, focusing on the competences acquired beyond formal schooling, during working life.

According to a feasibility study titled *Research scope for vocational education in the framework of COST social sciences*, transferability, flexibility and mobility should become the targets of VET in the future (Achtenhagen et al., 1995). These three targets are to be seen as heavily interrelated with the assessment of VET outcomes (Achtenhagen and Thang, 2002).

The conclusions of the European Council held in Lisbon in March 2000 mark a shift in the focus

of education policy towards lifelong learning, which ‘is no longer just one aspect of education and training; it must become the guiding principle for provision and participations across the full continuum of learning contexts’ (European Commission, 2000, p. 3). Lifelong learning is seen as the common umbrella under which all kinds of teaching and learning should be gathered and it calls for a fundamentally new approach to education and training.

In order to take action on lifelong learning, six key messages have been formulated by the European Commission with the purpose of carrying out a wide-ranging consultation on priority issues:

- (a) new basic skills for all;
- (b) more investment in human resources;
- (c) innovation in teaching and learning;
- (d) valuing learning;
- (e) rethinking guidance and counselling;
- (f) bringing learning closer to home.

The objective of valuing learning is to ‘improve significantly the ways in which learning participation and outcomes are understood and appreciated, particularly non-formal and informal learning’ (European Commission, 2000, p. 15). Regardless of the type of learner, innovative forms of certifying non-formal learning are considered important; absolutely essential is the development of high-quality systems for the accreditation of prior and experiential learning (APEL) (European Commission, 2001, p. 17).

These examples indicate interwoven trends which may be characterised as follows:

- (a) a change from input to output orientation in education;
- (b) a shift from the discipline or subject approach to competence and competence-based authentic measurement in education;
- (c) an increasing attention to the measurement and evaluation of competences acquired outside of educational settings;
- (d) methods of measurement and evaluation (= assessment) are regarded as interrelated with concepts of learning, development, education and training.

Do these developments since the 1990s have visible effects on VET at the EU level? A search in the Leonardo da Vinci project database for 1995-99 and 2000-06 with the combined keywords ‘measurement’, ‘evaluation’ and ‘competence’

yielded no result. Using other combinations, for example, 'assessment' and 'competence', 'diagnosis' and 'competence', 'certificate' and 'competence', only two project titles have been found. The search in Cedefop's VET-bib database with 27000 references using the keyword 'competence' found more than 800 in the titles, the keyword 'evaluation' more than 900, and 'measurement' more than 10, but the combination 'measurement', 'evaluation' and 'competence' gave no reference. The results with the two keywords ranged from 0 to 21 hits. These findings may be interpreted in several ways. The keywords may not be properly inserted in the database, or measurement in combination with

competence and evaluation is still absent in VET research at European level. In consequence, this contribution proposes to stimulate reflections on the issue by:

- (a) introducing a general diagnostic framework as a basis for measurement and evaluation;
- (b) selecting some competence concepts and analysing them against the background of this framework;
- (c) describing some assessment concepts in practice in selected EU Member States and analysing and evaluating them from a methodological angle;
- (d) outlining some implications for policy and practice.

2. Measurement and evaluation

The question of ‘how much’ refers to measurement, whereas the question of ‘how well does the individual perform’ refers to the evaluation of something observed and quantified against reference criteria. Both concepts will be developed in the following paragraphs.

2.1. A general model of measurement

Measurement as a systematic mapping ⁽⁴⁾ of psychological attributes and actions has to cope with the problem that parts of them are invisible. They have to be made visible for measurement and diagnosis. To locate and structure the observable and non-observable parts, a general diagnostic framework will be introduced (Figure 1). It differentiates three levels: the external conditions, the actual episodes, and the internal conditions as described below (Gagné, 1977; Klauer, 1973; Straka and Macke, 2002):

- (a) from the perspective of an acting or learning individual, a task, a situation, or a request (1 in Figure 1) and the created outcome (solution or product) (2) are elements of the visible socioculturally shaped external conditions;
- (b) the actual episodes operated by the individual are performed with the individual's actions. For their measurement it is crucial to differentiate between the observable and non-observable components of actions. The observable component of an action is the behaviour (3). The non-observable part of an action is the experience, which may be distinguished between cognitive (e.g. planning, controlling), motivational (e.g. achievement orientation) and emotional (e.g. fear, joy) aspects, all of them being part of episodes. It may be asked whether an action can be

realised independently of a context. The answer is no. Actions always need points of reference, such as mentally constructed representations of goals, intended outcome, pathways to the goal and criteria for bringing the process to an end. From the perspective of the individual these internal representations are actual information created by the individual and shaped by her/his prior knowledge or cultural perspectives ⁽⁵⁾ (Shephard, 2001; Bransford et al., 2000) as part of the internal conditions. It is the acting individual who is (re)constructing the situation and analysing it. From the individual's perspective a situation or a request is in itself part of her/his environment and, therefore, a ‘pure physical matter’ in the format of characters or signals (Straka and Macke, 2002, 2003).

It is important to point out that action and information are two inseparable sides of the same coin (4); they are differentiated only for analytical reasons. An example is the revised taxonomy of educational objectives (Bloom, 1956) which consists now of the ‘knowledge’ and the ‘cognitive process’ dimensions, both further subdivided and combined in a two-dimensional matrix (Anderson and Krathwohl et al., 2001). The components factual, conceptual, procedural, and meta-cognitive of the taxonomy's knowledge dimension correspond to types of information in the introduced general diagnostic framework. The cognitive processes of the taxonomy, such as remembering, understanding, applying, analysing, evaluating, and creating, name types of cognitive actions. The combination of these two dimensions and their components with the matrix-concept indicates that neither the information nor the action dimensions can exist on their own. Action and information are transient.

⁽⁴⁾ It should be noted that there are different qualities of measurement like nominal, ordinal, interval and ratio measurement level, which enable different types of conclusion and data processing (e.g. Kerlinger, 1973).

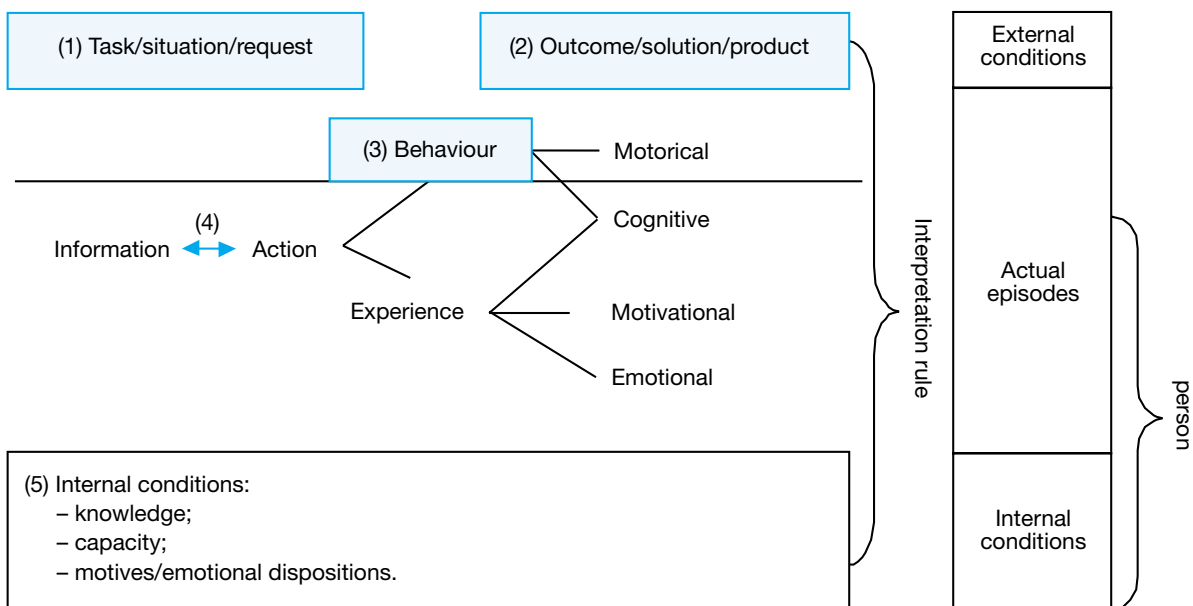
⁽⁵⁾ In using a hammer, the hammer is not in our brain but information about its composition, made of use, functions, etc.

They are present only in the moment of execution and immediately afterwards they belong to the past. Information as the result of an action stays only for fractions of a second in our working memory. It is not possible to rewind the interplay between action and information and to show it again like a movie. In order to keep on being able to act we have to rely on our knowledge and capability (located in our memory). Attributes such as knowledge, capability, motives,

emotional dispositions, and values are subsumed under the concept of internal conditions (5). In specific configurations they characterise an expert for a specific domain ⁽⁶⁾.

These considerations about the individual's interaction between the socioculturally shaped external conditions (environment), the actual episodes and the internal conditions as prerequisites and results of such a process, are put together in the following diagnostic framework:

Figure 1: **General diagnostic framework**



Source: G. A. Straka

This framework brings positive and negative implications for observation and measurement. The positive one is that outsiders can directly observe and measure the created and/or communicated outcome, the solutions and the result (2),

the task, situation, and the request (1), the difference between (2) and (1), and the behaviour (3) as well. All of them are performance elements which constitute the focus of performance assessment. The negative implication is that it is impossible for

⁽⁶⁾ 'Experts, regardless of the field, always draw on a richly structured information base; they are not just 'good thinkers' or 'smart people'. The ability to plan a task, to notice patterns, to generate reasonable arguments and explanations, and to draw analogies to other problems are all more *closely intertwined with factual knowledge than was once believed*' (Bransford et al., 2000, italics by G. A. Straka).

externals to observe, measure and evaluate directly the interplay between action and information and the internal conditions. The only way is to conclude from the observable, the non-observable. Considering that ‘observation [...] is always observation in the frame of theories’ (Popper, 1989, p. 31), theories, or at least hypotheses, about internal conditions and their development – like the pathways from novice to expert in a specific domain – are necessary. So are rules specifying the relationship between observation and theoretical constructs (indicated by the term ‘interpretation rule’ in the figure above; see also Shepard, 2000).

Summarising, assessment is always a process of reasoning from evidence. The results of such a procedure are only estimates of what a person knows and can do. ‘Every assessment, regardless of its purpose, rests on three pillars: a model of how students represent knowledge and develop competence in a subject domain (cognitions = internal conditions, actual individual operations), tasks or situations that allow one to observe students’ performance (external conditions), and an interpretation method for drawing inferences from the performance evidence thus obtained (interpretation rule). These three elements: cognition, observation, and interpretation rule constitute the angles of the assessment triangle’ (NRC, 2001, p. 2, italics by G. A. Straka).

2.2. Evaluation

In the context of individual environment interaction, evaluation judges measured competences against a defined benchmark (Straka, 1974). In this context, two approaches are distinguished: norm-referenced and criterion-referenced evaluation (Linn and Gronlund, 2000). In the case of norm-referenced evaluation, the measured competence is interpreted and judged in terms of the individual’s position relative to some known group. Criterion-referenced evaluation (7) is interpreting and judging the measured competence in terms of a clearly defined and delimited domain.

The two approaches have common and different characteristics. Both require a specification of the achievement domain to be mastered, a relevant and representative sample of tasks or test items. They use the same types of tasks and qualities of goodness (e.g. validity, reliability) for judging them. The differences – to some degree a matter of emphasis – are:

- (a) norm-referenced evaluation typically covers a large domain of requirements with a few tasks used to measure mastery, emphasises discrimination among individuals, favours tasks of average difficulty, omits very easy and very hard tasks, and requires a clearly defined group of persons for interpretations;
- (b) criterion-referenced evaluation focuses on a delimited domain of requirements with a relatively large number of tasks used to measure mastery, emphasises what requirements the individual can or cannot perform, matches task difficulty of requirements, and demands a clearly defined and delimited achievement domain (Linn and Gronlund, 2000, p. 45).

These different orientations of the two approaches have distinct impacts on the statistical measurement model. The norm-referenced assessment or the classical model is bonded with the normal curve whereas criterion referenced assessment is aligned with probabilistic models focussing on degrees of mastering a requirement and not related to the performance of other persons (Embretson and Reise, 2000). Criteria may be international, national, regional, institutional, individual standards such as personal goals, philosophies of companies (Section 4.4), state syllabuses, occupational profiles or NVQs (Sections 4.2 and 4.3). The core question remains: are these criteria concrete enough (operationalised) that decisions can be made on their basis. Since PISA there are some doubts – especially in Germany – indicated by the tendency to define national standards for subjects and age-groups. In the US, different content standards for school subjects or school types are already worked out (Linn and Gronlund, 2000, pp. 531-533). However, standards are varied. The reasons for this variety may be the specifics of

(7) Other terms for this type of evaluation are: standards-based, objective referenced, content referenced, domain referenced, and universe referenced.

the domains but also broad or diverse interpretations about what standards should look like (Bulmahn et al., 2003).

In the same way as the action is inseparably bonded to information in the course of the indi-

vidual's actual processing, the differentiation between measurement and evaluation is an analytical one. For that case the term 'assessment' is introduced and subsequently used, by subsuming measurement and evaluation.

3. Competence

The term competence has become increasingly popular since the 1980s in theory, practice and in VET political discourse. However, increasing use also led to an expanding meaning which will be illustrated by the reviews of Weinert (2001) and Eraut (1994). We will complement it by an analysis of some competence definitions used in VET against the background of the general diagnostic framework.

3.1. General competence categorisations

Weinert (2001, p. 45) introduces competence in general 'as a roughly specialised system of abilities, proficiencies, or skills that are necessary or sufficient to reach a specific goal' and he works out seven approaches by which competence has been defined, described, or interpreted:

- (a) general cognitive competences as cognitive abilities and skills (e.g. intelligence);
- (b) specialised cognitive competence in a particular domain (i.e. chess or piano playing);
- (c) the competence-performance model used by the linguist Noam Chomsky (1980) differentiating linguistic ability (= competence) enabling creation of an infinite variety of novel, grammatically correct sentences (= performance);
- (d) modifications of the competence-performance model which assume that the relationship between competence and performance is moderated by other variables such as cognitive style, familiarity with the requirements, and other personal variables (i.e. conceptual, procedural, performance competence);
- (e) cognitive competences and motivational action tendencies in order to realise an effective interaction of the individual with her/his environment; i.e. competence is a motivational concept combined with achieved capacity;
- (f) objective and subjective competence concepts distinguishing between performance and performance dispositions that can be measured with standardised scales and tests (objective) and subjective assessment of performance-relevant abilities and skills needed to master tasks;

- (g) action competence including all those cognitive, motivational, and social prerequisites necessary and/or available for successful action and learning (i.e. general problem-solving ability, critical thinking skills, domain-general and domain-specific knowledge, realistic, positive self-confidence, and social competences).

Eraut argues that competence is not a descriptive but a normative concept. 'In examining research into competence, we need to ask not only how competence is defined in general, but how it is defined in a particular situation, i.e. how these normative agreements are constructed' (Eraut, 1994, p. 169). Referring to Norris (1991) he distinguishes three main research traditions in post-war research in this field:

- (a) the behaviouristic tradition focussing more on training than on qualifications (competence-based training, e.g. in teacher education [Grant et al., 1979] widely spread in North America and heavily based on previous work in Canada [Ministry of Education, 1983]);
- (b) the generic tradition, rooted mainly in management education, widely used in Britain, for example Boyatzis (1982), listing 12 competences; e.g. concern with impact, self-confidence, or perceptual objectivity, and differentiating superior from average managers;
- (c) the cognitive competence tradition, most clearly articulated in Chomsky's theory of linguistics, which was adopted from research under the aegis of the UK Department of Employment and which is one of the rationales for the concept of national vocational qualifications (NVQ) in the UK.

Based on these two categorising reviews it can be concluded that the term 'competence' has the function of an umbrella for divergent research strands in human capacity development and its assessment. The umbrellas themselves differ in extent and differentiation with some overlap. Therefore, an additional approach will be chosen by introducing some definitions from the field of VET and analysing them against the background of the general diagnostic framework in order to draw some implications for measurement and evaluation of competence.

3.2. VET competence definitions

To give a flavour of how the competence concept in VET is used in different countries or institutions, definitions from EU level as well as from France, the UK and Germany will be introduced and discussed.

In the glossary prepared by the European Commission for the communication *Making a European area of lifelong learning a reality* (European Commission, 2001, p. 31) competence is defined as ‘the capacity to use effectively experience, knowledge and qualifications’.

Cedefop (Bjørnavold and Tissot, 2000, p. 208) define competence ‘the “proven/demonstrated” – and individual – capacity to use know-how, skills, qualifications or knowledge in order to meet usual – and changing – occupational situations and requirements’.

In France, Gilbert (1998, p. 9) defines competence as an entity of theoretical knowledge, ability, application knowledge, behaviour and motivation structured in mastering a specific situation.

Within the system of the national and Scottish vocational qualifications in the UK the focus is on

occupational competence, which is defined as ‘the ability to apply knowledge, understanding, practical and thinking skills to achieve effective performance to the standards required in employment. This includes solving problems and being sufficiently flexible to meet changing demands’ (QCA, 1998, p. 13). Based on a review of 100 NVQs and SVQs another brochure of the Qualifications and Curriculum Authority stresses that the standards are ‘expressed in outcomes [...] [which is] the required end result for the assessment of competence’ (QCA, 1999, p. 6).

The standing conference of State Ministers of Education and Culture in Germany (*Kultusministerkonferenz – KMK*) defines action competence (*Handlungskompetenz*) as the readiness and ability of a person to behave in vocational, societal and private situations in a professional, considered and responsible manner both towards him/herself and society. This competence concept is differentiated according to domain, personal, and social dimensions (KMK, 1996/2000).

In order to find out similarities and differences in these concepts they will be analysed according to the levels of the general diagnostic framework. The results are laid down in Table 1.

Table 1: Comparisons of concepts

Level	Internal conditions (knowledge, capacity, motives, etc.)	Actual episodes (action, behaviour, information)	External conditions (situation, task, requirement, outcome, product)
Definition			
European Commission (2001)	<ul style="list-style-type: none"> • capacity • experience • knowledge • qualifications 	<ul style="list-style-type: none"> • use (effectively) 	
Bjørnavold and Tissot (2000)	<ul style="list-style-type: none"> • capacity • know-how • skills • qualifications • knowledge 	<ul style="list-style-type: none"> • use 	<ul style="list-style-type: none"> • occupational situations and requirements (usual and changing)
France (Gilbert, 1998)	<ul style="list-style-type: none"> • knowledge (theoretical, applicable) • ability 	<ul style="list-style-type: none"> • behaviour and motivation, structured (mastering) 	<ul style="list-style-type: none"> • specific situation
NVQ-UK (QCA, 1998)	<ul style="list-style-type: none"> • ability • knowledge • skills (practical and thinking) 	<ul style="list-style-type: none"> • apply understanding • problem solving • being flexible 	<ul style="list-style-type: none"> • effective performance • demands (changing)
Action competence (KMK, 1996/2000)	<ul style="list-style-type: none"> • ability 	<ul style="list-style-type: none"> • readiness to behave (professional, considered, responsible) 	<ul style="list-style-type: none"> • situations (vocational, societal, private)
Source: Straka, G. A.			

Most of the definitions analysed span all the three levels but in varying differentiations; only the EU-definition does not include external conditions. Competence can be considered as a relational concept, bridging all the three levels involved in the individual-environment interaction. This has a strong impact for the public discourse: if the level and/or its elements are not specified, the audience may not be sure what the subject matter of a discussion is like. As a consequence, the probability of misunderstanding might be higher than of understanding. The result might be a pleasant and stimulating exchange of vague visions till others arise and dominate but no systematic empirically grounded insights and practical relevance.

In the context of measurement, behaviour is an indicator of action and the individual internal conditions. Bearing in mind that behaviour and action are transient, stimulated by the personally perceived environment and created by the internal conditions of the individual, there are two roads to specify competence:

(a) with situations, tasks or/and requirements (including the outcomes and/or criteria) for an occupation or an occupational function;

(b) with the knowledge, skills, abilities or more general the internal conditions necessary for an occupation or an occupational function.

In the first instance, much work has already been done in VET concerning occupational classifications (e.g. the international standard classification of occupations, ISCO, of the ILO, or VET curricula). However, it must be determined if they are adequate for measurement purposes. For the second approach, empirically tested notions, assumptions, or theories of development of internal conditions in occupational fields are required. Weinert (1999) argues that this context brings up a problem not yet solved in the 100-years history of scientific psychology. Therefore, he recommends the first approach which appears to be more functional (see NVQ). Eraut (2003) comes to the same solution but from a different perspective. He states that many definitions of competence can be differentiated, whether they are individually or socially situated. He prefers the second type, 'because public discourse on competence assumes that it is central to the relationship between professionals, their clients and the general public' (Eraut, 2003a, p. 2).

4. Procedures of assessing competence

According to Wolf (1995, 1998) four features are affiliated with competence-based measurement and assessment:

- (a) the emphasis on outcomes, specifically, multiple outcomes, each distinctive and separately considered;
- (b) the belief that these can and should be specified to the point where they are clear and transparent; that assessors, assessees and third parties should be able to understand what is being assessed, and what should be achieved;
- (c) the decoupling of assessment from particular institutions or learning programmes;
- (d) the idea of real-life performance essentially in non-academic fields.

These features raise the question of how to measure and evaluate competences. Direct and/or indirect observation seems to be the 'natural' and non-reactive approach. Direct observation relates to (visible) behaviour in specified work situations or generated products; counselling a client to choose a tour package or a menu in the tourism business, or realising a task on demand or a work sample are instances of this. Indirect observation refers to descriptions of activities, for example documented work samples including video tapes, project documentation and references. Such documentations may become a part of a portfolio, documenting increase in occupational competence. It is the purpose and the assortment that differentiates a portfolio from a collection or folder (Linn and Gronlund, 2000).

The observations, which are the basis for measurement, have to meet quality characteristics, some of them being accepted worldwide in both theory and practice of assessment (AERA et al., 1999 or DIN, 2002). The most important quality characteristics are validity, reliability, objectivity and, in addition, fairness, and usability (Linn and Gronlund, 2000; Lienert and Raatz, 1998; QCA, 1999; Schuler, 2000; O'Grady, 1991).

The criterion for validity is the degree to which the approach measures what is supposed to be measured (e.g. O'Grady, 1991). The recent and revised 'standards for educational and psychological testing' (AERA et al., 1999) define validity

as the degree to which accumulated evidence and theory support specific interpretations of scores. The bonding with theory is gradually reduced to evidence in specifying validity as the degree of the adequacy and appropriateness of the interpretations and conclusions drawn from observations or scores (Schuler, 2000).

Different aspects are added within the general notion of validity. Construct validity relates to the inference made about the perspective of adequate mapping of the individual internal conditions and their interrelations on the basis of the observations and scores. As an example, the assumption that information-creating actions consist of a dynamic interplay between learning and metacognitive control strategies, motivation and emotion (Figure 1) is to be tested with two steps. First, factor analyses should verify that each item of the measurement scale has an exclusive and important loading on the factor (= construct, e.g. motivation and not emotion). Second, structural analyses should at least reveal positive relations between these constructs (e.g. Straka and Schaefer, 2002). Referring to the general diagnostic framework or the assessment triangle (Section 2.1), the interpretation rule is on the test band.

The question whether the tasks represent the requirements of an occupation refers to content validity. Providing reasons why the documented practice project corresponds to the reference project in the German IT further education concept is an example of the implementation of such a quality criterion (Section 4.8). Prognostic validity refers to the accuracy of the prediction that can be made on the basis of the observations. Prognosis validity investigates, for example, the correlation between the methods used for hiring employees and their future career paths (Table 2 in Section 4.9).

Reliability expresses the degree to which scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable, and repeatable for an individual test taker (AERA et al., 1999). More generally, this quality criterion expresses the degree to which the same results are achieved irrespective of when, where or by whom a person is assessed (O'Grady, 1991). A

prerequisite of high reliability is that the subjective factors related to the assessor and unsystematic differences in administering and analysing are eliminated, minimised or held constant. These aspects are covered by objectivity, which is sometimes subsumed under reliability. A different quality criterion receiving more attention nowadays is fairness, which focuses on avoiding discrimination due to circumstances individuals cannot control. Usability addresses the practicability and, ultimately, the costs of measurement procedure.

For each of these criteria of measurement quality a sophisticated methodology for empirically based assessment has been developed since the Binet-Simon first IQ test in 1905. Furthermore, these criteria are interrelated, i.e. low objectivity influences reliability which itself influences validity (reliability is a necessary but not a sufficient condition for validity). In assessment practices in general, and also in VET, a reasonable balance between potential validity, reliability, costs and usability is essential. Fairness is becoming a major issue in scientific discourse and in some recent publications fairness is regarded as an essential part of a comprehensive view of validity (Linn and Gronlund, 2000). This should be borne in mind when evaluating a measurement procedure.

A selected sample of procedures for measuring and evaluation of competence used in the EU will be presented next with the aim of providing an impression of the range of approaches as well as their common features and differences. Afterwards, these procedures will be analysed and evaluated against the background of a general diagnostic framework (Figure 1) and the quality criteria for measurement described above.

4.1. The bilan de compétences (France)

The *bilan de compétences* is a broad diagnostic procedure taking place in accredited assessment centres. The bilan is a tool of feedback, flexibility and mobilisation of national continuing education policy. It has been a legal instrument since 1991 with several modifications since then (the latest was the decree for social modernisation in 2001). The main aims affiliated with the *bilan de compétences* are:

(a) strengthening self-responsibility for the occupational career of the individual;

(b) institutionalising the political goal of lifelong learning;

(c) enhancing the experiential knowledge, the transfer knowledge and the performance aspects in ones occupational biography (*acquis professionnels*).

This last aim targets workers with low formal education to establish a counterweight to the dominance of diplomas and formal certificates of qualifications job career paths in France (Thömmes, 2003).

The competence-based measurement of the *bilan de compétences* does not represent a uniform method or a distinct instrumentation. Different approaches are used, with common steps and phases. The starting point of the *bilan de compétences* is an occupational and/or personal desire to change, especially for people with low formal educational status. Knowledge, skills and other potentials acquired during working life are diagnosed, although formal qualifications and certificates are also considered.

The *bilans* are carried out in *Centres interinstitutionnels de bilans de compétences* (CIBC). Participation is voluntary and free of charge for the assessee. The result is given only to the participant, even though the employer contributes directly (time away from work) and indirectly (contributions to continuous training) to the procedure. The process of measurement consists of three phases, altogether taking 20 to 24 hours, with breaks between. On average about four to six weeks are necessary.

Phase 1 covers prearrangement of the contract (approximately four hours), fixing aim and areas, and compiling individually the repertoire of methods to be used.

Phase 2 is execution of the arrangement (12 to 16 hours): interviews to reconstruct the occupational biography; tests in areas such as intelligence, interests, learning ability, motivation, skilfulness and concentratedness; standardised personality inventories, simulations exercises, business games; and evaluation of the learning and developmental potentials.

Phase 3 is the final session (approximately four hours). This presents the results as a whole on the basis of a strengths and weaknesses profile and discusses the results with the intention to create a written outline for a project of an occupational change. There is also the arrangement of a follow up meeting about six month afterwards (Thömmes, 2003).

People conducting this procedure have to be

qualified (higher education degree in psychology, education, medicine and the like) and they are allowed only to use methods ensuring reliability and common scientific standards.

However, in reality, there is still a large discrepancy between educational policy aims and reality. The number of realised *bilan de compétences* decreased from 1993 (126 000) to 2000 (78 788) (Gutschow, 2003) and the proportion of unemployed persons using this instrument is above average. The future will show if the new law from 2001 initiated a change.

4.2. The national vocational qualifications (NVQ) (UK)

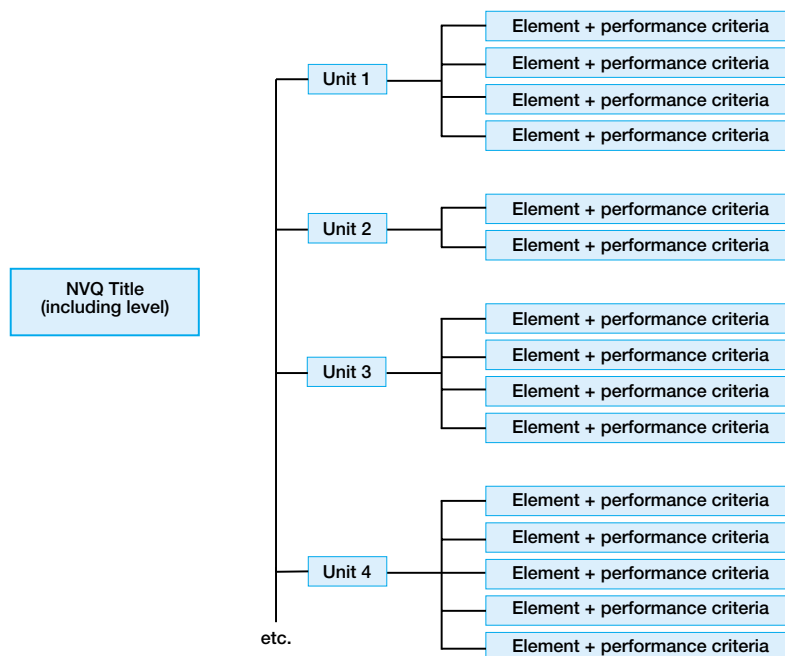
During the period 1980 to 1995 there was a strong cross-party consensus that the myriad of self-regulating awarding bodies should be reduced. The result was a competence-based system, the NVQ, in England and Wales. Employer-led organisations (lead bodies) define standards which embody and define competence in the relevant occupational context

based on a functional analysis of occupational roles (Wolf, 1995, p. 15 et seq.). Each NVQ covers a particular area of work, at a specific level of achievement related to the five levels of the NVQ framework. Levels 1 and 3 are presented as examples:

- (a) Level 1: competence which involves the application of knowledge in the performance of a range of varied work activities, most of which may be routine and predictable;
- (b) Level 3: competence which involves the application of knowledge in a broad range of varied work activities performed in a wide variety of contexts, most of which are complex and non-routine. There is considerable responsibility and autonomy, and supervisory competence in some capacity is often required.

Each level when applied to an occupation such as 'providing financial services (banks and building societies)', is divided into units, for example 'sell financial products and services' (CIB, p. 72) ⁽⁶⁾, for which standards have been developed. Each standard consists of an element and performance criteria, for example 'identify buying signals and correctly act on these' (CIB, p. 75, 3.2.5). The structure of an NVQ title is represented below.

Figure 2: The structure of a NVQ



Source: Wolf, 1995, p. 19

⁽⁶⁾ A unit itself may be allocated to different levels; for example unit 3 'providing financial products and services' of the NVQs 'Providing financial services' (banks and building societies) is allocated to level 2, 3 and 4. (CIB, n.d.). This unit is to be found in the Annex 3.

Extracts of the NVQ for ‘providing financial services (banks and building societies)’ are presented in the Annex. The standards for this sector comprise the levels 2, 3 and 4 of the NVQ system, with a total 64 units. The units are grouped in nine functional areas: sales and marketing, customer service, account services, mortgages and lending, financial advice, administration, asset management and protection, organisational effectiveness, and resource management.

The 33 units of Level 3 for the NVQ ‘providing financial services’ are subdivided into a mandatory set and two sets labelled as ‘groups’ from which units can be chosen according to regulations. Two units belong to the mandatory set (unit 45, contribute to a safe, secure and effective working environment; and unit 55, manage yourself), five units to group 1 (e.g. unit 15, maintain and improve customer service delivery) and 26 to group 2 (e.g. unit 2, sell products and service over the telephone). To gain the full NVQ at level 3, a total of eight units must be achieved: the two mandatory units, one unit selected from group 1 and the remaining five units from group 1 or 2.

As an example, unit 3, sell financial products and services (see Annex for the complete unit), relates to non-mortgage and non-regulated product and services, such as non-regulated insurance products investments and foreign currency. The assessee must show an ability to identify customer needs and promote suitable products and services by presenting them accurately and seeking customer commitment to buy (CIB, p. 72).

Unit 3 itself consists of three elements: (3.1) establish customer needs; (3.2) promote the features and benefits of the products and services; and (3.3) gain customer commitment and evaluate sales technique.

For each of these elements the assessee must demonstrate the following knowledge and understanding: ‘which products and services you are authorised to promote, key features and benefits of products and services you are responsible for promoting, the types and form of buying signals your customers might use (three specimen out of nine statements)’ (CIB, p. 72).

For element (3.2), promote the features and benefits of the products and services, nine performance criteria are specified, for example, ‘(3.2.1) explain clearly and accurately the features

and benefits of products and services that most closely match your customer’s needs; (3.2.5) identify buying signals and correctly act on these; or (3.2.9) record all relevant information accurately’ (three of nine performance criteria statements) (CIB, p. 76).

The NVQ assessment itself is based on two types of performance evidence: ‘(1) Products of the candidate’s work, e.g. items that the candidate produced or worked on, or documents produced as part of a work activity. The evidence may be in the form of the product itself, or a record or photograph of the product. (2) Evidence of the way the candidate carried out activities or evidence of the processes in demonstrating competence. This often takes the form of witness testimonies, assessor observation, authenticated candidate reports of the activity, or audio/video recordings’ (QCA, 1998, p. 19).

In unit 3 the following evidence requirements are listed: ‘You must be able to demonstrate your competence at promoting the features and benefits of your products and services through: your sales performances with customers, and your corresponding customer records’ (CIB, p. 75). The evidence criteria are further specified with appropriate examples listed: interview/discussion notes and records, referral documentation, where sales opportunities were passed to a relevant colleague, and personal customer records.

The ‘element’ also specifies the knowledge and understanding considered as necessary for performing according to the standard in this area. Examples for the descriptions are: ‘Make sure you can answer the following questions, e.g.: what is the difference between “features” and “benefits”? How does this apply to the different products and services you deal with? What sort of buying signals might customers use to show their interest, or disinterest, in the products and services being offered? How could you deal with the different types of buying signals?’ (CIB, p. 76).

These unit-based qualifications are open to everyone. There is no need to have a prior qualification to start an NVQ and because each is based on a number of units, it is possible for individuals to build up qualifications at a pace that suit them. Individuals can then achieve certificates for specific units or a full NVQ when all the units have been completed. There are no time

limits for completing NVQs. The NVQ assessment is generally carried out in the workplace, with the emphasis on outcomes rather than learning processes (CIB, p. 1).

Such a procedure meets, to a degree, the requirements for competence-based assessment:

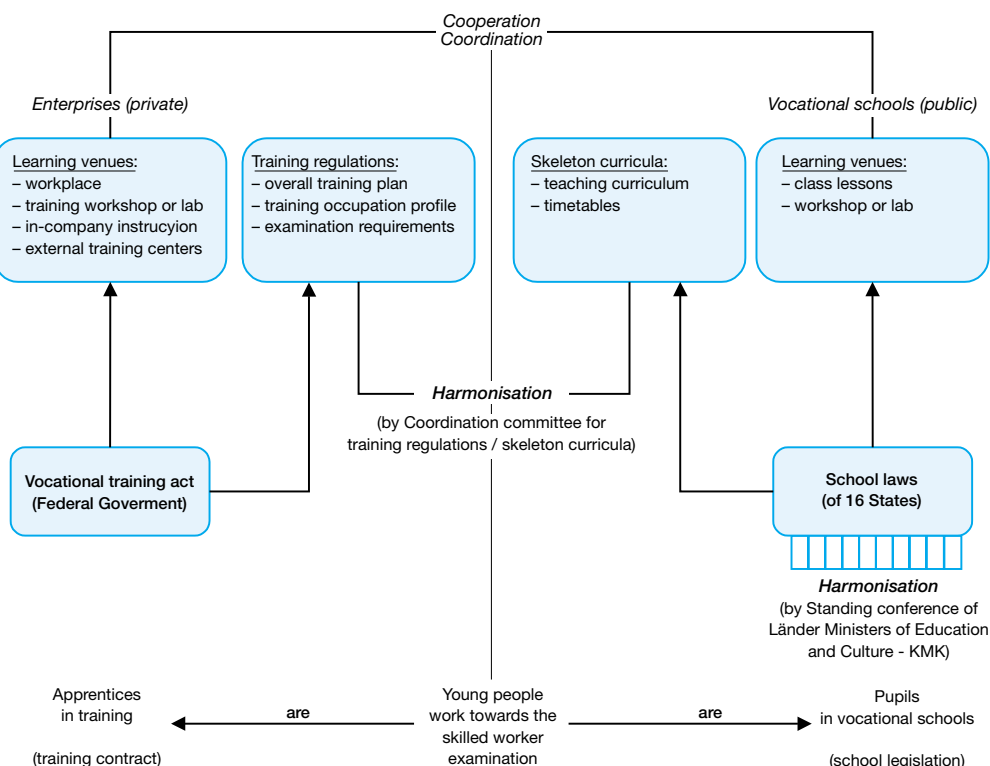
- (a) one-to-one correspondence with outcome-based standards;
- (b) individualised assessment;
- (c) competent/not yet competent judgements only;
- (d) assessment in the workplace;
- (e) no specified time for completion of assessment and no specified course of learning/study (Wolf, 1995, p. 20).

4.3. Assessing action competence in Germany

In 1996, Germany shifted to action competence, the development of which, however, is still bonded to explicit training in company and education in

school. The apprentice has to be simultaneously active in both venues for a fixed time, ranging from two to three and a half years. During this transition phase from school to work the young persons are company employees on the basis of a training contract and pupils in vocational schools (based on school legislation) at the same time. The goals for training and learning in these private enterprises are set in the training regulations consisting of an overall training plan, a training profile and examination requirements. The basis for the training regulations is determined in the 1969 Vocational Training Act of the Federal Government. Learning arrangements in the vocational schools are class lessons, workshops or labs. The teaching takes place according to state curriculum. Although the 16 German States enjoy full responsibility in education, two harmonisations are effected by the standing conference of State Ministers of Education and Culture (KMK). One is the coordination of the curricula of the 16 States; the other is tuning the training regulations to the skeleton curricula (*Rahmenlehrplan*) for vocational schools (Figure 3).

Figure 3: Overview of the structural features of the dual system



To demonstrate assessment subjects some extracts from training regulations and school syllabus, and from the profile for German bank clerks, follow.

The occupation profile is divided into the field of activity and 17 occupational skills. In the field of activity it is noted that bank clerks work in all the fields of business that lending institutions are engaged in. Their duties involve acquiring new customers, providing counselling and other services to existing customers and selling banking services, particularly standardised services and products. [...] The occupational skills of bank clerks is specified as follows: 'advise customers on selecting an appropriate type of account; [...] advise customers on investment possibilities in the form of shares, bonds, investment fund shares; [...] sell investment products; [...] evaluate the costs and revenues arising from business relationships with customers; [...] are competent in communicating, cooperating with others, solving problems and taking decisions' (Bundesanzeiger, 1998, p. 21).

The occupation profile, giving a condensed overview about the skills and knowledge of the occupation bank clerk, is made more concrete with the overall training plan (the training company is responsible for) and the skeleton school curriculum ⁽⁹⁾. According to the overall training plan for investment in shares, the following skills and knowledge should be acquired: '(a) inform the client about investments especially in shares, bonds and investment fund shares; [...] (c) assess the benefits and risks of investment in shares; [...] (f) answer client's questions concerning costs of buying and selling shares; [...] (i) describe derivatives and their risks in general' (Bundesanzeiger, 1998, p. 8).

Because of the shift from subject- to competence-orientation by the German vocational school system, the curricula should no longer consist of specifications of occupation related contents and skills but of learning fields (*Lernfelder*) ⁽¹⁰⁾ which are related occupational tasks and business processes. Each learning field is divided into specifications of objectives and contents.

For bank apprentices, the skeleton school curriculum consists of 12 learning fields for which some samples are given for learning field 4 (see Annex for a complete description): 'Offering long and short term investments in shares and bonds' comprises 100 of the 320 school hours of the first year of schooling. Examples of objectives are: The students identify client's signals for needs and motives for investment, communicate instruments of financing – client-oriented [...] use product-related calculations; explain services connected with investment decisions; describe risks associated with investment decision; heed regulations for protection of investors. Concerning contents, specifications like the following ones are specified: investment on accounts – like saving accounts [...] bonds, shares, investments funds shares, rights, risks, [...] basics and principles of investment consulting (Bundesanzeiger, 1998, p. 18).

In the German dual system of VET there is an interim and a final examination. Because the final examination comprises more assessment procedures than the interim one, only the final examination will be described. It consists of written tests in the following domains: business administration (180 minutes), control (90 minutes), and macroeconomics and civics (90 minutes). The tests consist of tasks related to practice (reality of interrelations) in selected contents to demonstrate analysis and understanding. Fixing them is the responsibility of federal and/or local examination boards in which the social partners are represented (employers, unions). For the bank clerk, the federal examination board recommends 60 % closed (multiple or other choices) and 40 % open tasks related to a national content catalogue (*Stoffkatalog*) developed and fixed by the same board, calibrated with the overall training plan and the skeleton school curriculum.

There is also an oral examination consisting of a simulation of client counselling (20 minutes and 15 minutes preparation; selected, by regulations fixed areas of content). As an example of the criteria and the way performance is measured and evaluated an extract is presented:

⁽⁹⁾ 'Skeleton' because of the sovereignty of the 16 States in education. Each State can develop its own syllabus along the skeleton school syllabus.

⁽¹⁰⁾ For a critique e.g. Straka (2001, 2002); Straka and Macke (2003).

Figure 4: Excerpt from an observation and evaluation sheet for client-counselling of bank clerks

Assessment Criteria (sample)	Remarks						
A. Action competence							
I. Interaction behaviour							
<ul style="list-style-type: none"> • produces interest in the client • establishes contact with the client • creates an agreeable atmosphere 	<hr/> <hr/> <hr/>						
II. Informational and analytical behaviour							
<ul style="list-style-type: none"> • questions and analyses customers needs • keeps the thread • 	<hr/> <hr/> <hr/>						
III. Selling behaviour							
<ul style="list-style-type: none"> • persuades the client with arguments • makes to the deal single minded 	<hr/> <hr/>						
B. Subject area competence							
<table border="1"> <tbody> <tr> <td>A. Action competence</td> <td><hr/></td> </tr> <tr> <td>B. Subject area competence</td> <td><hr/></td> </tr> <tr> <td>Total points</td> <td><hr/></td> </tr> </tbody> </table>		A. Action competence	<hr/>	B. Subject area competence	<hr/>	Total points	<hr/>
A. Action competence	<hr/>						
B. Subject area competence	<hr/>						
Total points	<hr/>						
Remarks:							
100–92=1	91–81=2	80–67=3	66–50=4	49–30=5	29–0=6		

Source: DHT, 1999

Such a simulation and evaluation of a client-clerk communication is a novel aspect of the shift of German VET to action competence in 1996 and the differentiation into domain, personal and social dimensions. A similar approach was implemented in the reform of commercial VET in Switzerland since 1999, where these three inter-related dimensions (domain, technique, and social) were integrated with the competence cube ⁽¹¹⁾ http://www.igkg.ch/deutsch/rkg_kompetenzwuerfel.htm

4.4. Social and methodical dimensions of competence (DaimlerChrysler)

Assessments are also developed and practised by companies. However, only a few have been published as these assessments are generally for internal use only. An available example was developed at DaimlerChrysler in Germany. It focuses on social and technical aspects of competence (Weisschuh, 2003).

⁽¹¹⁾ Available from Internet: http://www.igkg.ch/deutsch/rkg_kompetenzwuerfel.htm [cited 2.3.2004].

The assessment concept is named ‘training in dialogue’ (*Ausbildung im Dialog – AiD*) and, after a trial phase, was introduced company-wide around 1999. It is still in use and it is being considered whether to use this method for continuing assessments in both company education and training (Weisschuh, 2003).

‘Training in dialogue’ consists of an observation sheet to be completed by the persons in

charge of human resource development in order to record non-valuated observations of apprentices’ behaviour (see element (3) in Figure 1), a sheet for apprentices to assess their own competences and a dialogue sheet to fix target agreements between trainer and trainee.

As an example of the assessment sheet, the items and ratings for the key competence ‘ability to cooperate’ are presented.

Figure 5: Rating scale for the key competence ‘ability to cooperate’

0	10	20	<p>Ability to cooperate</p> <p><input type="checkbox"/> Remarks/reason:</p> <p>Respects the opinion of others</p> <ul style="list-style-type: none"> ➤ Takes the opinion of others seriously ➤ Checks one’s own point of view within the conversation ➤ Stands by common decisions ➤ Treats others in an open and fair way <p>Correct behaviour and support of others</p> <ul style="list-style-type: none"> ➤ Establishes and supports contact with others ➤ Takes stock concerning others ➤ Refers to her/his own knowledge to that of others ➤ Support others ➤ Integrates outsiders into the trainee group
30	40	50	
60	70		
0	10	20	
30	40	50	
60	70		
<input type="checkbox"/>			
0	10	20	
30	40	50	
60	70		
<input type="checkbox"/>			

Source: Weisschuh, 2003

The trainer evaluates the trainee with respect to the different key competences between 0 (does not fulfil the requirements at all) and 70 (exceeds the requirements). In the same way, the apprentice rates her/his own competences. After that, both sides can come together to compare and discuss the ratings. In case of disagreement, the trainer can use only the behaviours written down by her/him on the behaviour sheet. The meeting should end with fixed target agreements recorded in the dialogue sheet.

4.5. Competence at work: a case in the Netherlands

In 1999, the advisory committee for education and labour market published a proposal under the name ‘shift to core competences’. The employers’ key argument is that people should be skilled in doing a job and that these skills should preferably be acquired outside the school world of disciplines and subject matter. This kind

of reasoning has led to detailed standards of occupational competence in each industry sector, forming the basis for vocational credits. Furthermore, the decoupling of learning process and assessment was one of the principles of the qualification system, laying the ground for a new kind of development: learning competence based on experiential learning.

A cooperative pilot project at Frico Cheese, within the Leonardo da Vinci programme, aimed to test a procedure of assessing prior learning and to validate competences relative to the qualification structure in agriculture vocational education. The motive for Frico Cheese to take part was based on upgrading the workforce. Employees had been working in the company for many years with virtually no additional training or education; many had bad experiences with full time school learning processes in the past. Accreditation of achieved competences (AAC) seemed to be a better way of reducing employee resistance to learning than formal processes of schooling or training. The human resource department of Frico Cheese decided that the accreditation of achieved competences procedure should be easily accessible to employees and carried out at the workplace by an internal assessor.

'A candidate is said to be competent when he or she is able to perform in realistic work situations according to the standards defined by the lead bodies in that specific work domain' (AAC-project, 2001, p. 41). The assessment process consists of two phases: evaluating a personal portfolio and demonstrating one's competences in an authentic work situation (= on the job).

A portfolio contains descriptions of relevant experiences and diplomas of the employee obtained in a formal or non-formal way. This description is to be compared to the national standard. Conclusions are drawn regarding content and level of competences. This is a kind of matching process between qualifications demanded and competences offered. The procedure allows for the easy identification of the elements already included in the portfolios as well as of those that have to be assessed by performance-based assessment tasks. Based on these outcomes, a decision is made whether the employee is allowed to continue the procedure.

Using a checklist of authentic tasks, the assessor compares the performance with the standards. In a successive interview the candidate is asked to reflect on the performed tasks and to respond to questions on transfer of working methods and solutions to similar situations. A positive evaluation of this process leads to recognition of the competences and a certificate closes this process.

Some 77 of the lower educated employees in the assessment procedure received an average of three certificates without visiting the community college. However, it is difficult to measure this success, not knowing how many people are not selected to take part in the procedure, and not knowing what the reliability scores are for the different assessors, both internally and externally (Nijhof, 2003).

A study done at Frico Cheese after one year of the Leonardo da Vinci project records no promotion of employees; there was no link between the number of certificates and salary increase. Some of the employees felt it easier to get a job outside the company based on the certificates, but these were exceptions (Nijhof, 2003).

Most participants felt, according to the interviews, that learning on the job has helped to improve job performance; this is followed by technical training. General education was considered the least important. Employees with accredited competences are perceived as good, or even best, at their job. Nijhof (2003) however, raises the question whether a classic performance test would not have given the same result.

4.6. The *realkompetanse* project (Norway)

The *realkompetanse* project of the Norwegian Ministry of Education and Research started in August 1999 and formally came to an end in July 2002. Its aim was 'to establish a system that gives adults the right to document their non-formal and informal learning [without having to undergo traditional forms of testing]' (VOX, 2002, p. 9). *Realkompetanse* includes 'all formal, non-formal and informal learning [...] Formal learning⁽¹²⁾ is normally acquired through

⁽¹²⁾ It should be pointed that 'formal learning' is a metaphor. Learning itself is exclusively a personal matter whereas 'formal', 'non-formal' or 'informal' vaguely indicate some characteristics of the environment (Straka, 2002a; Straka, 2003b).

organised programmes delivered via schools and other providers and is recognised by means of qualifications or parts of qualifications. Non-formal learning is acquired through organised programmes, but is not typically recognised by means of qualifications, nor does it lead to certification. Informal learning is acquired outside of organised programmes and is picked up through

daily activities relating to work, family or leisure' (VOX, 2002, p. 11).

The validation process combines dialogue-supported performance assessment and portfolio assessment, possibly in conjunction with tests (VOX, 2002). It covers the following areas:

- (a) documentation methods in work life and in civil society/third sector;

Figure 6: **Competence card from the workplace**

Specification of social and personal skills		Level
Cooperation and communication	Cooperates on a daily basis with management and other members of staff. Able to contribute to finding solutions. takes part in some important decision-making. Able to make other realise the importance of following up procedures and routines.	
Effort and quality of work	Reliable, keeps deadlines and appointments. Responsible and conscientious. Able to handle longer work intensive periods. Quality conscious.	
Customer service	Open and outgoing. Achieves contact easily with other people. Good knowledge about the needs of membership businesses. Experience and skills in active recruiting of new in-service businesses.	
Initiative – flexibility – creativity	Flexible. Likes to take on new tasks. Able to come up with proposals to new solutions to e.g. the Quality Control system, or arrangements for following up apprentices.	
Work related to restructuring – acquisition/use of new knowledge	Has taken part in trial runs of a new system for wages and personnel administration, and has made suggestions to the improvement of this.	

Specification of management skills in position		Level
Staff and labour management	Responsible for the daily follow-up of an apprentice in the company.	
Training and instruction	Responsible for the information to schools and classes. Takes actively part in the establishing of training programmes in local in-service businesses. Has taught classes in clerical courses for adult apprentices and interns.	
Level A = Carries out elementary tasks under supervision Level B = Works independently within own area of responsibility	Level C = May hold professional responsibility, may council and advise Level D = Has a very good insight in subject area of profession, may be in charge of development on own workplace.	

Place:	Date:	Signature of employee:
Place:	Date:	Signature on behalf of business/company:

- (b) validation with respect to upper secondary education;
- (c) validation for admission to higher education.

The documentation of non-formal and informal learning in the workplace consists of two parts: a curriculum vitae and a skills certificate. The curriculum vitae is similar to that put forward by the European Commission and the forum on transparency of qualifications. It includes: personal data; work experience (employer, position, period, areas of responsibility); education; valid licences and public approved certificates; courses (names, period, completed year, important content); other skills and voluntary work (type of skill and activity, skills and closer description of responsibilities); and additional information. The skills certificate or the 'competence card from the workplace' describes what the employer is able to do as a part of his or her job. The categories are: personal data; main areas of work responsibility with a closer description of responsibilities; specifications of professional skills needed to carry out main responsibilities; personal capability; social and personal skills; management skills; and additional information. These specifications are graded at four levels, from level A (carries out elementary tasks under supervision) up to level D (has a very good insight in subject area of profession, may be in charge of development on own workplace). The certificate is signed by the employee and on behalf of the business/company.

Another example for the documentation of non-formal and formal learning is the 3CV developed by a number of organisations of the civil or third sector. This instrument contains an introduction (in which the methodology for completion is described), an example of a completed form, a form ready for completion, and the option of creating one's own reference.

4.7. Recreational activities (Finland)

The recreational activity study book of the Finnish Youth Academy is designed as a tool to make learning visible in settings outside the formal system. The target group is all young people over 13 years. Up to March 2002 there were over 55 000 study book owners in Finland and over 250 formal educational institutions acknowledge the value of the entries in the book. This concept

of valuing learning results in non-formal settings is supported by the Finnish Ministry of Education and Culture and the Finnish Forest Industries Federation. The major youth and sport NGOs behind the development include the centre for youth work of the Evangelical Lutheran Church of Finland, the Nature league, Guides and scouts of Finland, the Finnish Red Cross and the Swedish study centre in Finland.

The concept of the recreational activity study book differentiates nine settings or types of learning activities. The study book itself is divided into nine categories, according to the nature of the learning activity: regular participation in leisure activities; holding positions of trust and responsibility within NGOs; activities as a leader, trainer or coach; participation in a project; courses; international activities; workshop activities (apprenticeship); competitions; and other activities.

Examination of the categories shows that there are non-formal and formal environments for learning. The most formalised form for learning is the category 'courses' which means organised, and often hierarchical, educational programmes offered by various youth and sport NGOs and other training providers. The eight other categories fall more or less under the umbrella of informal sets, in which the learning-by-doing approach is often the method for acquiring competences and skills.

Adults responsible for the records are instructed to focus on the learning and development of the young person, i.e. not only to document what has been done but also how it has been done. Various aspects are recorded in the book: organisation responsible for the activity; description of the activity; young person's self evaluation of learning; date(s) and duration of the activity; performance (success and development); signature of the adult person responsible; contact information of the adult; and stamp of the responsible organisation (Savisaari, 2002).

The entries in the book are always written by an adult person (over 18 years of age) who is either responsible for or well aware of the particular activity. The learners fill in the 'self-assessment of the learning'. The idea is to focus more on what and how things have been learned rather than only what has been done. The person countersigning the entry adds his/her contact information, in case someone wants to check whether the young person actually has participated in the activity or not.

Figure 7: Example page of the recreational activity study book

Type of activity: Holding positions of trust and responsibility with in NGOs	
Organisation in which the activity took place	Time/dates of the activity ____/____/____ - ____/____/____
Position of the young person in the organisation	In average _____ hours per week/month Successes and competences acquired
Description of the activity	Place Date
Young persons self – assessment of the learning	Signature of the person responsible of activity
	Contact information of the undersigned person
	Position of the undersigned person

Source: Savisaari, 2003

4.8. New ways of assessing competences in information technology (IT) in initial and continuing VET in Germany

A new type of final examination was introduced in Germany in 1997 for recognised information technology occupations (*IT-Berufe*) within the dual system. Besides subject matter, macroeconomics and civics, two ‘holistic tasks’ (*ganzheitliche Aufgaben*) are to be solved in 90 minutes each (written practical part). The former practical part of the examination – for example, simulation of client counselling for bank clerks – was replaced by a

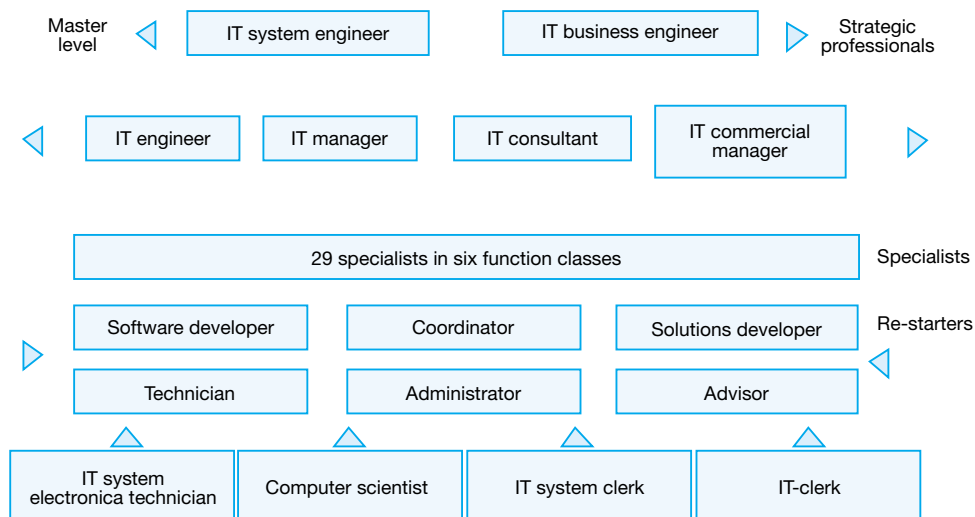
project to be realised in the company of the apprentice (written-oral part). This project work consists of an authentic assignment order or a defined partial task to be realised within 35 hours of work. The proposed project has to be accepted by the examination board and the documented result of the project work has to be presented orally in front of the local examination board, followed by a professional discussion.

A new IT continuing education system became active in May 2002 based on these recognised IT occupations of the German dual system (GDS) ⁽¹³⁾. The system consists of within-company career paths and three levels:

⁽¹³⁾ Available from Internet:
http://www.bmbf.de/pub/it-weiterbildung_mit_system.pdf,
<http://www.apo-it.de> and
<http://www.kib-net.de> [cited 2.3.2004].

- (a) the lowest level includes 28 special profiles, including software developer, network administrator, IT key account person and e-logistic developer;
 - (b) the second level, with operative functions, is differentiated IT engineer, IT manager, IT consultant, and IT commercial;
 - (c) the third level, with strategic functions, is divided into IT systems engineer and IT business engineer.
- A certificate for each of these levels will be treated as equivalent to a BA or MA certificate in the higher education system via the European credit transfer system (ECTS) (Figure 8).

Figure 8: Structure of the IT-further education system

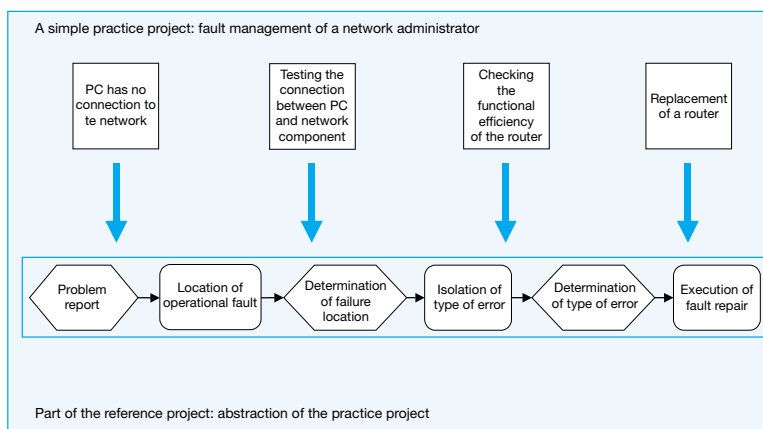


Source: Grunwald and Rohs, 2003

The basis for the certificates acquired in this system is a ‘work process oriented continuous training’ during which the employee works on actual requirements of company projects in order

to combine work and learning. The project work has to be documented, discussed with the coach in the company and checked against ‘reference projects’ (Grunwald, 2002, p. 179).

Figure 9: From a practice to a reference project – an example for a network administrator



Source: Grunwald, 2002, p. 170

Certification is run with a private certification body according to the general criteria for institutions certifying personnel according to the German Industry Norm (DIN EN 45013).

4.9. Analysis and evaluation

A review of the introduced sample of approaches to assessing occupational competences indicates a wide span of procedures. They range from broad, such as the *bilan de compétences*, to narrow targets, such as the rating scales for specific social and methodological dimension competence at DaimlerChrysler. In the general diagnostic framework, the foci of the approaches tend to be different. NVQs concentrate on the external conditions (i.e. tasks, solutions) to be solved and observable behaviour whereas, according to the German view of competence, assessment, actions and conclusion, internal conditions are the subjects of consideration.

All the procedures described tend to get away from the classical measurement, characterised by well defined internal conditions and tight tasks mostly formatted with multiple choice items. Instead they favour authentic and complex tasks from the shop floor. Combined with the constructivist approach and interconnected with the qualitative empirical research approach, introspection (self-assessment), peer assessment or observing on-the-job performance seems to be the first choice for measuring and evaluating competences, especially for softer dimensions like the social and communicative. But, no matter which approach to competence assessment is chosen, it should attain the highest possible degree of validity, reliability, objectivity, fairness and usability, i.e. the quality criteria of any assessment.

Introspection and self-evaluation are explicitly or implicitly biased, a phenomenon extensively investigated in survey research under the term 'socially desired answers' (e.g. Kerlinger, 1973). Another source of errors stems from features of general and job-related knowledge and skills. Personal explicit knowledge as a part of the internal conditions may become a matter of routine especially if practised over a long period. The opposite takes place when people learn to act successfully but are not aware of the knowl-

edge base of their actions. In both cases an identification of the knowledge incorporated in such activities makes self-evaluation impossible. Both phenomena are linked with the concepts of implicit and tacit knowledge, the latter characterised by 'we know but cannot tell' (Polanyi, 1966). These features of actions and knowledge set narrow borders to introspection as a basis for assessment of vocational aptitudes.

Work documentation and portfolios are also given high attention in the context of competence-based measurement: see *bilan de compétences*, the case from the Netherlands, or the assessment approach of 2002 German IT-continuing education. Such products are not fugacious and they can be observed, analysed and evaluated independently by different assessors. However, the selection of work pieces to be documented may be biased, and there is no guarantee whether the assessee produced the outcomes on the basis of her/his competence and without help of others.

Portfolio procedures are not easy to handle because of the heterogeneity of information being recorded, and the variation of experiences of employees. Last but not least, there is little systematic or empirical investigation of whether portfolios are acceptable predictors of future performance, especially in the case of continuous changes in the work environment (Nijhof, 2003).

At a first glance, observation by external evaluators might alleviate the weaknesses of introspection and self-compilation of documented work samples. However, this type of observation carries its own bias, because people tend to reconstruct the surrounding world under the perspectives of their prior knowledge and cultural perspectives (e.g. Shephard, 2001; Bransford et al., 2000).

Observations in natural settings by others – peers, supervisors – are more likely to be unsystematic than systematic. This fact is taken into consideration in practice, by using specifications for recording and evaluating observations, such as the check list for assessing the simulated client counselling in the German final examinations for bank clerks (Section 4.3), the rating scales of 'training in dialogue' at DaimlerChrysler (Section 4.4), the criteria of the Norwegian competence card from the workplace (Section 4.6.), or the in the Finnish recreational activity study book (Section 4.7).

Basing external observation and evaluation on defined aspects or criteria is a way of handling such sources for errors. However, the core problem is still left: the processes of recording, summarising and evaluating take place in the brain of the assessors and, therefore, they cannot be cross-checked by outsiders to converge to objectivity. Beck (1987) lists 20 sources for errors aligned with rating and he warns against 'fantastic performances of observers'.

These methodological and theoretical considerations raise the question whether any empirical evidence exists concerning the measurement quality of assessments in VET. Unfortunately such evidence is scarce.

The *bilan de compétences* in France is neither dedicated to measuring potential nor to supporting decisions concerning personal or professional development. The proportion of subjective measurement is high and the introduced quality criteria are generally seen as not suitable (Thömmes, 2003). A similar view might have guided the Dutch Frisco Cheese project funded in the Leonardo da Vinci programme when Nijhof (2003) criticises the absence of considerations concerning reliability and validity.

Eraut points out that the demonstration of competence in company A does not guarantee that this competence will work in company B. According to his investigations, Eraut (1994, 2000, 2003b) competences are context- and situation-bounded and, therefore, transfer from A to B needs an additional learning process. The NVQs prioritise on-the-job learning and the term competence was selected for 'its declared purpose of accrediting only effective performance in the workplace; but NVQs also had to be national qualifications' (Eraut, 2001, p. 94). Therefore, a generic language is used to describe activities which might have the same function but are very different in character. 'This enabled formal comparisons between "equivalent" jobs in different contexts, but did not guarantee that competence in one context could be readily transferable to another' (Eraut, 2001, p. 95). Finally, an evaluation of high level NVQs indicates

the dangers of a fragmented approach to performance, and the limited development of expertise in assessment and learning support (Eraut et al., 2001).

In Germany there is considerable critical discussion about the pros and cons of assessing competence and quite a lot of changes in methods of assessing competences in different occupations have taken place over the last decade. However, the discussion, the introduction and the practice are predominantly based on speculative reasoning instead of systematic empirical evidence. Actual assessment practice in German VET seems to have more in common with the rituals for transition from apprenticeship (novice) to craftsmanship (expert) in the Middle Ages than with the potentiality of measurement approaches in recent times (Straka, 2003a; see discussion in Chapter 5).

A meta analysis of empirical findings in occupation-related proficiency was carried out to determine empirically based devices for the potential measurement quality of assessment procedures. It researched the prognostic validity (*Berufseignungsforschung*)⁽¹⁴⁾ of different selection procedures compared to future occupational success.

The correlation (r) between data about 'occupational proficiency' and later 'occupational success'⁽¹⁵⁾ – the indicator for prognostic validity – can range between '1' and '0'. '0' represents no and '1' perfect prognostic validity. Table 2 indicates that the range of prognostic validities of the presented approaches is quite large. The lowest prognosis factor for occupational success is chronological age (.01) and the highest are standardised cognitive skill tests (.51), structured interviews (.51), and work samples (.54). However, the general validity of work samples should not be overestimated because they are often very closely related to the requirement of a specific workplace. Therefore, it is not certain that the same knowledge and skills are transferable to another context, especially if they are tacit (Eraut, 2001; Straka, 2001b).

⁽¹⁴⁾ In this context it should be pointed out that a German standard for requirements, procedures and their application in job-related proficiency assessment was published in June 2002 (DIN, 2002).

⁽¹⁵⁾ It should be noted that 'occupational success' in the table above is predominantly based on evaluations by supervisors, a method whose measurement quality is questionable.

Table 2: **Validity of customary approaches**

Validity of customary approaches to diagnosing occupational proficiency and occupational success	
Approaches to diagnosing occupational proficiency	validity (r)
conventional interview	.14
personality analysis (in general)	.15
conscientiousness	.31
integrity test	.41
school grades	.30
references	.26
work sample	.54
biographic questionnaire	.37
assessment centre	.35
structured interview	.51
time of probation	.44
cognition skills tests	.51
graphology	.02
age	.01
work experience	.18
Source: Moser, 2003	

At first glance, a correlation of 0.5 within the boundaries of '0' and '1' seems to be quite reasonable. But a more powerful indicator is the common factor variance expressing the overlap or the communality of the variances of two variables (in our case 'occupational proficiency' and 'occupational success') (Kerlinger, 1973). It is calculated by squaring the correlation coefficient ($= r$) and in the case of $r = .50$ the result is 0.25, which expresses that 25 % of the common variance, for instance, 'occupational proficiency' and 'occupational success' are due to their commonness, and 75 % of the 'residual' variance might be ascribed to other factors such as motivation, workplace design, organisation structure, work climate. Even so, hiring people on the basis of a structured interview or a cognitive skill test is, on average, significantly better for occupational success than selecting them by age or graphology (Table 2).

The EU approach to assessment procedures indicates a certain preference for performance-based

assessment. The main differences between performance-based assessment and classical approaches are presented in Table 3.

Expectations for the performance-based assessment are extremely high but a reasonable decision for the one or the other is ultimately a matter of empirical proof. However, the problem underlined by Baker et al. (1993, p. 1210) remains: 'Although interest in performance-based-assessment is high, our knowledge about its quality is low'. Some considerations indicate that this approach does not only have advantages. Complex problems linked with significant degrees of choice may increase the number of steps required to solve a problem as well as the number of solutions. For example, the cognition technology group at Vanderbilt (CTGV, 1997) constructed 'authentic' problems with at least 14 steps required to solve them. From the angle of pure combinatorics there are $14!$ ($14 \cdot 13 \cdot \dots \cdot 1$) sequences or solution paths (permutations) possible. For content reasons most of them will give no sense but there may be a number of meaningful ways and solutions. Considering that there is not a single right or wrong solution, several parties rating independently are necessary to achieving objectivity and fairness. This approach – used in VET too – has consequence on usability and some biases and errors of measurement remain.

In parallel, multiple-choice items get much attention in public debates on assessment. Their validity has often been questioned compared with open ended items, especially in the context of performance and competence assessment. A methodological analysis in TIMSS-III-Germany of items measuring mathematic-science literacy found out that using multiple choice or extensive answer formats impacts on measured performance, but that it is not the case when multiple choice is compared to short answer format. However, independent of the format, the TIMSS-III items were a much better indicator of general ability in mathematics and science literacy. Therefore, the TIMSS analyses used an overall score ignoring the item format (Klieme et al., 2000). A recent study (Straka and Lenz, 2003) of commercial VET shows that students used more learning strategies for solving multiple choice items in the test for economic literacy (TEL), (Beck and Krumm, 1998) than for open ended teacher-made

items. Considering such evidence, it may be concluded that open-ended tasks are per se not more valid than multiple choice items.

To summarise, much supports the assumption that measurement quality is related to its format but possibly more to the content of the item. Consequently, there is a demand from Baker et al. (1993) for better research to evaluate the

degree to which newly developed or practised assessments fulfil the expectations. Similarly, Brown (1999) notes that multiple choice items 'are much more sophisticated than many of us once believed (see the UK Open University for examples of best practice) and don't always have to be self-generated by tutors'.

Table 3: **Performance-based assessment versus 'classical' approaches**

Features of tasks/situations	Classical approach	Performance approach
1. Task format	Closed (multiple choice)	Open ended
2. Required skills	Narrow, specific	High order, complex
3. Environment relation	Context free	Context sensitive
4. Task/requirement	Limited scope, single and isolated skill, short time processing	Complex problems, requiring several types of performances and significant time
5. Social relations	Individual	Individual or group performance
6. Choices	Restricted	Significant degrees
Source: Baker et al., 1993		

5. When should competence be assessed and by whom?

The approaches to measuring competence presented in Chapter 4 are embedded in different external conditions in which competence is developed and assessed. These range from the development of work-related capacity as a result of acting in a workplace during working life (e.g. NVQs, *bilan de compétences*, or the competence card from the workplace) to personal development initiated and supported by a combination of educational and authentic work settings in dual systems (e.g. the dual VET in Germany), or exclusive educational arrangements with real and/or simulated practice (e.g. technical or vocational schools).

Linked with such approaches are regulations for assessment and its timing. German dual VET, for example, has two fixed dates for assessment – in the middle and at the end of apprenticeship – whereas a *bilan de compétences* or an assessment in the framework of NVQs can be requested and run at any time.

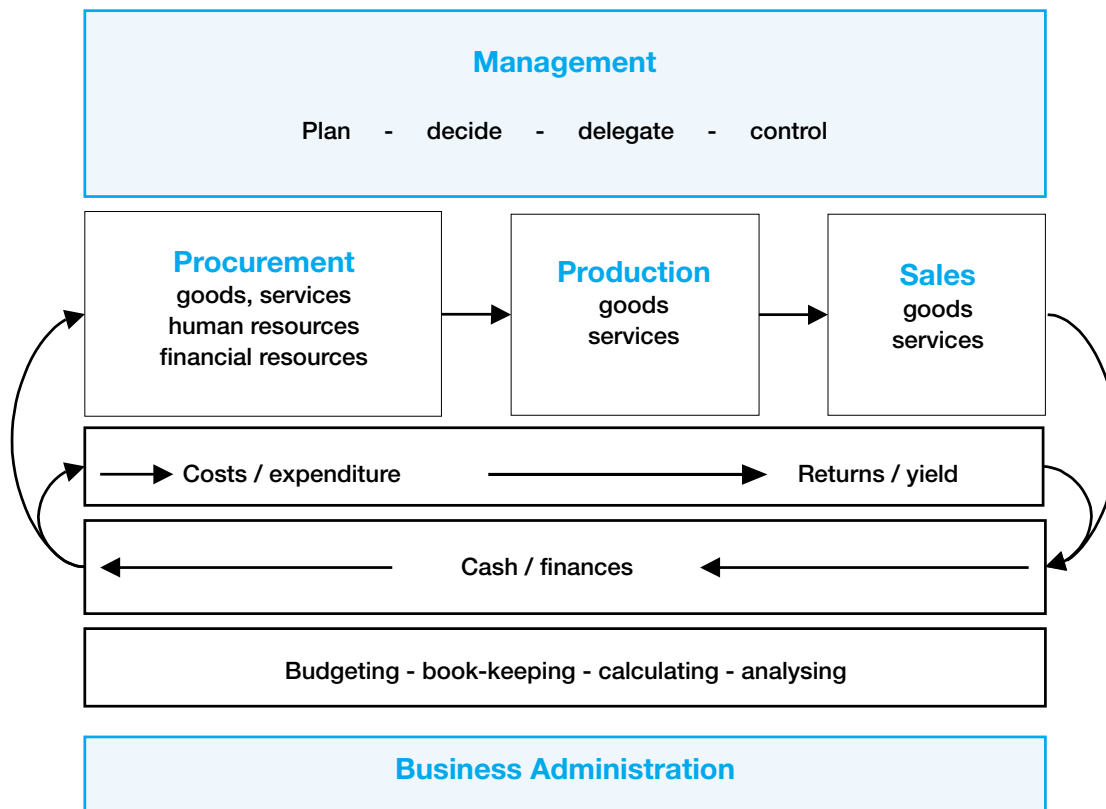
The question of the scope of assessment has to be tackled in this context. In VET courses, the contents and the objectives of the whole course might be the subject of assessment, as in the dual system in Germany. In a training system organised with modules and credits, certificates or grades confirming the mastery of a set of defined modules can be combined according to regulations for certification. For example, to get a level three certificate for providing financial services, the mastery of a defined number of units ⁽¹⁶⁾ out of different sets (Section 4.2) has to be certified by an external assessor.

There are disputes about modularisation and assessment in the EU (Ertl, 2002). Discussion between supporters of modularisation and proponents of courses concentrates, above all, on fragmentation versus wholeness, which need to be balanced for business administration (Figure 1).

Enterprises are open, dynamic, complex, autonomous, market-oriented productive social systems (Ulrich, 1970) embedded in socioculturally shaped environments. This dynamic and complexity contrasts with the general model of the economic process (Porter, 1985; Wöhe, 2002) where production and selling of goods and services results from combining material and immaterial factors according to economic principle. In order to realise such goods and services, goods (raw materials, equipment), services and human and financial resources have to be procured. Every input and sale of goods has direct or indirect financial impacts with associated costs/expenditures, returns/yield and cash/finances. The whole process is initiated, planned, decided, delegated and controlled by management. The information needed for this function is provided by budgeting, calculating, accounting and analysing the flow of goods, services, costs, expenses, and returns (Thommen, 1991).

Regarding a module as a ‘subset of a learning program’ and a unit as ‘a coherent set of learning outcomes’ (Ertl, 2002, p. 143), and noting that there are no specified criteria for coherence and subsets, there is significant freedom in structuring teaching programmes and specifying and assessing learning outcomes. Elements such as management, procurement, or sub-elements of them such as planning, deciding or further subsets, might be the subject of units or modules which themselves are considered interrelated. Their objectives might be assessment and grading. Teaching and realising the objectives, however, may take various paths through these sets: planning, to controlling, to delegating and deciding, and vice versa; sales to production to procurement or vice versa; or planning of production, sales, etc. The structural organisation of

⁽¹⁶⁾ An interpretation of a unit as a module is possible because ‘it is unclear in many cases whether the words “unit” and “module” are used interchangeably or whether they mark a difference’. The situation is similar when organising the teaching process in modules is the issue (Ertl, 2002, p. 143).

Figure 10: **General model of economic processes, a basis of business administration**

these modules is different and the structure of the business process is in part hierarchical and inter-related. However, the processing itself is always sequential and linear, i.e. one item after another (as indicated above) which has some consequences for assessment too.

Figure 1 showed that the basis of measurement and evaluation are observations of products of actions and/or the behavioural part of actions. On this level the observations are similar, independent of whether the education and/or training is structured in modules or courses. For example, if bookkeeping or sales is on the agenda, the action episodes are the same. They even have a considerable overlap across different economic cultures, as was seen in a pilot analysis of the operationalised educational objectives for bank clerks in the German VET regulations and of the performance and evidence criteria in the frame of the NVQs for 'providing financial services (banks and building societies) (Straka, 2002b; see Annex). Therefore, in both systems – German VET

and NVQs – similar observations might be the basis for assessment.

However, there are differences in the inferences draw from observations. If the focus is purely on performance, the observation is sufficient if the task is solved or the process mastered by a person. But if the knowledge base for these solutions or of the processes is targeted, assumptions about the knowledge and its structure are necessary and have to be validated (= construct validity). For example, a purely performance-oriented assessor would mark her/his observation 'explain clearly and accurately the features and benefits of products and services that most closely match your customer's needs' (see performance criteria 3.2.1 in Annex) as a satisfactory, whereas an assessor focusing on internal conditions (knowledge, skills, etc.) might interpret the same observation as an indicator of declarative and procedural knowledge. As a consequence, an investigation of the effects of holistic or fragmented, modularised or

programmed, dynamic or static assessment is not a question of assessment but of the aims – or more generally the theory (Popper, 1989) – and the organisation of vocational training and/or education.

In this context, the question may be asked whether the best practice and exchange of experience models (*Modellversuche* in Germany) recommended between Member States (often in the case VET) on assessment procedures is a good alternative to systematic empirical research in this field? Generally, reports of best practice describe what was processed, the results and the conditions. In such documentation, relations between input, output and processes are considered in most cases on the basis of qualitative data, and according to specific interpretations of the constructivist paradigm; generalisations are not intended. Therefore, it is not surprising that transfer of best practice normally does not work. The situation and context tie of the practice make the transfer its own experiment (Eraut, 2000). On the other hand, from experimental and quasi-experimental research methodology factors interfering with the validity making the transfer impossible are known (Campbell and Stanley, 1963; Straka, 1974). Consequently, systematic empirical approaches with accurate research designs are to be preferred.

The question of when assessment should be organised is connected with the question of by whom the assessment should be done. As discussed in Section 4.9., self-assessment is the weakest approach for assessment. From the perspective of validity, preference should be given to assessment by others. Where tasks have a wide range of solutions, e.g. in the case of dialectical problems – neither the target nor the ways to reach it are defined (Dörner, 1976) – or approaches, several evaluators are usually engaged.

Involving social partners in assessment is not only a way to balance it more but also contributes to fairness. In the German dual system, for example, the participation of the chambers of commerce, the trade unions and the school as examiners is regulated by law. However, from a measurement angle the question is whether the social partners and their representatives guarantee more valid assessment than trained assessors.

Because there is no empirical data available to answer this question, a fictional reflection may

reveal some problems. Certificates and/or grades should guarantee that the assessed person possesses the skills, knowledge, and attitudes which make her/him eligible for an occupation or able to perform it successfully. If social partners are involved in defining assessment standards, they might do so representing divergent class positions, i.e. the traditional antagonistic ones of capital and labour. Their judgement might be influenced primarily by power relations and only after by questions of the validity, reliability and objectivity of such a standard. As a consequence, the certificate or grade defined might rather reflect the following metaphor: a dog, having its snout in the fire and its tail frozen, feels, on average, agreeably warm (Straka, 2003a).

In VET, peer or colleague assessment – especially in the case of group work – might be an alternative form of assessment by others. Such persons know each other and are familiar with work requirements and conditions. In addition, evaluations might be done on the basis of much more observations over a far longer time than punctual or even continuous external examinations. However, this ‘natural’ procedure bears some analogue risks as a retrospective field study found out. In the context of socialist production in the former German Democratic Republic the labour units (*Produktionseinheiten*) at the Rostock Neptun shipyard gained increasing autonomy for fixing work norms. An effect of this was that, during the 1950s, the over-fulfilment of the norms exploded dramatically; exceeding norms by 200 % or 300 % was not unusual. But, in reality, shipyard productivity decreased; this was a trend which may have contributed to the implosion of the former German Democratic Republic (Alheit, 1998).

In this context, one might wonder if external and punctual final assessments might not offer a better guarantee of validity? An answer might be derived from the German situation. Very often – especially when a new group of apprentices is hired – complaints about decreasing literacy in reading, writing, and arithmetic, etc., are heard from companies or their associations. The complaint may have some truth, as TIMSS and PISA indicate for certain segments. However, success rates in the final examinations are surprising. Regularly, they reach around 85 % of those who sit for final examination in the dual

system in Germany ⁽¹⁷⁾ (*Berufsbildungsbericht*, 2003, p. 99). Different interpretations of this are possible; starting a totally new career path motivates the young persons to close the gaps from former schooling, or the efforts in instruction and training in vocational schools and companies have remedial effects on the capabilities of apprentices. But another interpretation might be possible as well. The assessment procedures carried out by the social partners have more in common with the rituals of guilds of long ago than with modern procedures of measurement and evaluation (Straka, 2003a). This justifies advocacies for a VET-PISA (Pütz, 2002), and the related feasibility studies which are planned (BIBB-Workshop PISA-B 2003).

A VET-PISA might be in accord with the recent educational reforms, emphasising accountability with the following significant new features:

- (a) ambitious, world-class standards;
- (b) forms of assessment that require students to perform more substantial tasks (e.g. construct extended essay responses rather than select answers from multiple-choice items);
- (c) imposing high-stake accountability mechanism for schools, teachers, and sometimes students;
- (d) including all students (Linn and Gronlund, 2000, p. 5 et seq.).

However, such a kind of permanent VET-PISA would have dangers as well as advantages. According to the American education research association the following conditions have to be met in high-stakes testing:

- (a) protection against high-stakes decisions based on a single test;
- (b) adequate resources and opportunity to learn;
- (c) validation for each separate intended use;
- (d) full disclosure of likely negative consequences of high-stakes testing programmes;
- (e) alignment between the test and the curriculum;
- (f) validating passing scores and achievement levels;
- (g) opportunities for meaningful remediation for examinees who fail high-stakes tests;
- (h) appropriate attention to language differences among examinees;
- (i) appropriate attention to students with disabilities;
- (j) careful adherence to explicit rules for determining which students are to be tested, sufficient reliability for each intended use;
- (k) continuous evaluation of intended and unintended effects of high-stakes testing (AERA, 2000).

⁽¹⁷⁾ The range of success rates in 2001 included crafts at 80.6 %, industry and trade at 88.6 %, public service at 91.1 %, and overall at 86.1 % (*Berufsbildungsbericht*, 2002, p. 99).

6. Summary and conclusion

The emphasis on measurement and evaluation of competence is rooted in:

- (a) a shift from the predominant input to an output view of education, linked to accountability of educational systems and institutions;
- (b) a shift from subject to literacy orientation in general and vocational education;
- (c) a rise in the cognitive-constructivist paradigm with its orientation toward qualitative assessment methods as opposed to the behavioural-cognitive paradigm of traditional quantitative procedures;
- (d) greater recognition of skills acquired in non-formal and informal settings during lifetime, linked with the request for innovative forms of certification.

Considering that any human performance is based largely on the knowledge, skills and motives of the individual, a general diagnostic framework was introduced. It differentiates three levels: internal (e.g. knowledge, skills, motives) and external conditions (e.g. situation, task, product) both bridged with the actual individual operations (e.g. behaviour, action). In this model, observing, measuring, and evaluating changes in the external conditions and the behaviour is a minor problem compared with that of making inferences on the non-visible part of the actions and the internal conditions. In order to solve this problem, explicit theories with well-defined constructs about domain specific features and the development of internal conditions, cognitive actions and interpretation rules linking observations with these constructs are necessary (Table 1).

Evaluation may have as a benchmark the measured characteristics of other persons (norm-referenced) or standards defined independently from measured characteristics of other persons (criterion-referenced). The latter is an appropriated, recommended, and used criterion in the context of VET.

Analysis of a sample of competence definitions used in the EU with the diagnostic framework revealed that the concept of competence is a

relational, one bridging all the three levels of Table 1 in most cases. Such a broad notion requires not only an accurate definition of the elements of the levels and their interrelations but also their systematic and empirical validation. Otherwise there is a significant risk of mutual misunderstanding in public and scientific discussions about VET competence, because it might often be unclear which level or element is under consideration. The result might be an interesting, even stimulating exchange of broad visions but of negligible scientific and practical relevance. Such a holistic view of competence might also become counterproductive to transparency and mobility, the officially stated objectives of the European Commission.

Practised assessment procedures in the EU were analysed against the background of the general diagnostic framework. The examples chosen included the *bilan de compétences* (France), the NVQ (England and Wales), different dimensions of action competence in the German dual system, assessing competences at work (the Netherlands), realkompetanse (Norway), the assessment of recreational activities (Finland), and valuing competences in the continuing IT-training (Germany).

The results of the analysis include an overwhelming orientation towards measuring and evaluation of performance in 'authentic' or 'natural' settings with standards of differing sophistication (e.g. NVQs) or prototypical project solutions (e.g. German IT-continuous education). There is also a preference for open-ended rather than multiple choice tasks, for complex rather than narrow and specific skills, for context-sensitive rather than context-free strategies, for complex problems requiring several types of performance and significant time rather than tasks with limited scope for a single and isolated skill to be solved in short time, and for significant degrees of choice rather than restricted choices.

Task format can affect performance but there is empirical evidence that the format is much less important than the task requirements in terms of content. Systematic empirical research on perfor-

mance measurement with 'authentic' tasks is required to examine whether they support the implicit and explicit expectations placed on them.

No published empirical evidence for the quality criteria of measurement, such as validity, reliability, objectivity, fairness and usability, can be found. However, summarised results from occupation-related proficiency research indicate a wide range of prognostic validity – from nearly 0 (chronological age) to .54 (work sample) – related to different approaches to diagnosing occupational proficiency (Table 2):

Different types of observation are used, ranging from introspection to observation, both direct (e.g. self-reports, participating observation) or indirect (portfolios). Since the act of recording, summarising and evaluating cannot be checked by outsiders, such approaches are subject to many potential errors of measurement, even when operational observation criteria are available. Recent research shows also that this type of assessment might carry sociopsychological constraints, especially if done by superiors (Krumm and Weiß, 2000). Therefore, the use of externalised work products or audio-visually recorded behaviour is recommended.

However, in the context of VET the ultimate goal is the sustainable change of personal characteristics. Hence these systematic and quantified observations become indicators of internal conditions. An approach such as the one chosen in the PISA study, that assesses the ability to complete tasks in real life, depending on a broad understanding of key concepts, rather than assessing the possession of specific knowledge in specific situations, is to be recommended. The domains are defined in terms of 'the content or structure of knowledge that students need to acquire in each domain [...] [see internal conditions]; the processes that need to be performed [...] [see actual episodes]; and the context in which knowledge and skills are applied [...] [see external conditions]' (OECD, 2001, p. 19).

This review covers only the tip of the iceberg for procedures, concepts of measurement of competence. The results underline the need for a review of VET competence assessment procedures from the methodological angle, confirming the conclusions of Björnavold (2000) in making learning visible in the field of learning in non-formal and informal settings. Starting with a

detailed description supported by concrete examples, the focus of further reviews of VET competence assessment should be on empirical evidence with respect to measurement quality criteria: objectivity, reliability, validity, fairness, and usability. If sufficient research findings are not available, empirical investigations concerning the quality of measurement of these approaches should be initiated.

This methodological undertaking has to be complemented by a conceptual domain-specific approach taking into account the non-visible parts of the actions and internal conditions and their development. Beyond VET, many such concepts from neighbouring disciplines are available (e.g. Weinert, 2001; Straka, in preparation). Their adequateness for VET should be examined instead of producing 'new' ones, in most cases not empirically validated. Such investigations could become a basis for a VET PISA in selected important forward-looking occupations or sectors under the auspices of the EU. It is time to stop initiating projects discovering the sociocultural idiosyncrasies in European VET systems and increase those mapping European VET on the basis of sound empirical data.

Valuing competences acquired in non-formal and informal settings also has elements of paradox. A rationale for this policy is that schools and their certificates have a marginal prognostic value for mastering life outside the formal settings (Resnick, 1987; Straka, 2003a). Another is that learning in out-of-school settings creates less inert knowledge than in classrooms (Cseh et al., 2000). However, competences acquired out of school are evaluated with standards of the formal system. To make the discussion even more paradoxical, because the workplace or the NGOs (see recreational activity study book, Finland, 3.7) are such wonderful learning arenas, the learning outcomes have to be valued in order to become eligible for the formal schooling.

There is another threat to this policy. Research shows that job-related continuous learning in non-formal settings is an important aspect for employed people aged 19 to 64 in Germany. In the year 2000, two thirds of the employed said they practised this type of self-education. However, people who failed to complete dual education (blue-collar workers, immigrants and working women) were under-represented (Kuwan

and Thebis, 2001). Findings confirming that some people do not reach the full operating ability during their lifetime provide another example (Oerter, 1980). As a consequence, valuing learning outcomes acquired in non-formal

settings might support the Mathew effect – i.e. giving more to those who already have – and jeopardise fairness if no support grounded in learning and domain specific development theory is added.

List of abbreviations

AERA	American educational research association
IT	Information technology
NVQ	National vocational qualification
PISA	Programme for international student assessment
SVQ	Scottish vocational qualification
TIMSS	Third international mathematics and science study
VET	Vocational education and training

Annex 1: Sample overall training plan (*Bundesanzeiger*, 1998, pp. 8-9)

Lfd. Nr.	Teil des Ausbildungsberufsbildes	Zu vermittelnde Fertigkeiten und Kenntnisse
1	2	3
4.	Geld- und Vermögensanlage (§ 3 Nr. 4)	
4.1	Anlage auf Konten (§ 3 Nr. 4.1)	<ul style="list-style-type: none"> a) Kunden über Anlagemöglichkeiten auf Konten einschließlich der Sonderformen des ausbildenden Unternehmens beraten b) Konten eröffnen, führen und abschließen c) Kunden über rechtliche Bestimmungen und vertragliche Vereinbarungen informieren d) Kunden über Verfügungsberechtigungen und Vollmachten beraten e) Kunden über Zinsgutschriften und über deren steuerliche Auswirkungen informieren
4.2	Anlage in Wertpapieren (§ 3 Nr. 4.2)	<p>Kunden über Anlagemöglichkeiten, insbesondere in Aktien, Schuldverschreibungen und Investmentzertifikaten, informieren</p> <p>Kunden über rechtliche Bestimmungen und vertragliche Vereinbarungen informieren</p> <p>Chancen und Risiken der Anlage in Wertpapieren einschätzen</p> <p>Kunden über Kursnotierungen und Preisfeststellungen Auskunft geben</p> <p>bei der Abwicklung einer Wertpapierorder mitwirken</p> <p>Kundenanfragen zu Wertpapierabrechnungen beantworten</p> <p>Kunden über Verwahrung und Verwaltung von Wertpapieren beraten</p> <p>Kunden über Ertragsgutschriften und deren steuerliche Auswirkungen informieren</p> <p>Finanzderivate und deren Risiken in Grundzügen beschreiben</p>
4.3	Anlage in anderen Finanzprodukten (§ 3 Nr. 4.3)	<ul style="list-style-type: none"> a) Vertrieb von Verbundprodukten zur Kapitalanlage und zur Risikovorsorge im Rahmen der Organisation des ausbildenden Unternehmens erklären b) beim Abschluss von Bausparverträgen mitwirken c) Kunden über Möglichkeiten der Kapitalanlage und der Risikovorsorge durch Abschluss von Lebensversicherungen informieren

Annex 2: Sample of skeleton school syllabus (*Bundesanzeiger*, 1998, p. 18)

4. Lernfeld Geld- und Vermögensanlagen anbieten

**1. Ausbildungsjahr
Zeitrichtwert 100 Stunden**

Zielformulierung:

Die Schülerinnen und Schüler ermitteln Bedarfssignale und Anlagemotive der Kunden. Sie präsentieren Finanzinstrumente kundenorientiert. Sie erläutern Preiseinflussfaktoren, Kursbildung und Kursveröffentlichungen. Sie werten Produkt- und Marketinginformationen aus. Sie nutzen produktbezogene Berechnungen. Sie erläutern aus der Anlageentscheidung resultierende Serviceleistungen. Sie beschreiben Risiken, die aus Anlageentscheidungen entstehen, und beachten die Vorschriften des Anlegerschutzes.

Inhalte:

Anlagen auf Konten am Beispiel der Spareinlage: Vertragsgestaltung aus Kunden- und Bankensicht, Bedeutung der Sparerkunde, Regelverfügungen und vorzeitige Verfügungen, Verzinsung, Besteuerung der Zinserträge

Termineinlagen, Sparbriefe

Besonderheiten des Bausparens und der Kapitallebensversicherung gegenüber anderen Anlageformen

Schuldverschreibung, Aktie und Investmentzertifikat als Grundformen der Wertpapiere: Rechtsnatur, Rechte der Inhaber, Ausstattung, Risiken, Emissionsgründe

Kursbildung und Kursnotierung am Beispiel von Aktien; Kurszusätze, Kurshinweise

Grundlagen und Grundsätze der Anlageberatung

Verwahrung und Verwaltung: Girosammelverwahrung, Wertpapierrechnung; Depotstimmrecht

Maßnahmen zum Schutz der Anleger

5. Lernfeld Besondere Finanzinstrumente anbieten und über Steuern informieren

**2. Ausbildungsjahr
Zeitrichtwert 60 Stunden**

Zielformulierung:

Die Schülerinnen und Schüler präsentieren Finanzinstrumente für besondere Anlagewünsche. Sie werten Produkt- und Marktinformationen aus und nutzen produktbezogene Berechnungen. Sie stellen Grundbegriffe des Einkommensteuerrechts und die rechtlichen Rahmenbedingungen zu Geld- und Vermögensanlage dar. Sie geben einen Überblick über die Finanzmärkte und erklären deren einzel- und gesamtwirtschaftliche Bedeutung.

Inhalte:

Wertpapiersonderformen am Beispiel von Genussschein und Optionsanleihe; Rechte der Inhaber, Ausstattung, Risiken, Emissionsgründe

Finanzderivate am Beispiel einer Aktien-Option und eines Futures: Rechte des Inhabers, Risiken, Einsatzmöglichkeiten

Grundbegriffe des Einkommenssteuerrechts

Steuerliche Gesichtspunkte bei der Anlage in Wertpapieren: Besteuerung von Erträgen und Kursgewinnen am Beispiel von Aktien und Schuldverschreibungen

Finanzmärkte: Arten, Funktionen, Bedeutung

Rahmenbedingungen des Kreditwesengesetzes und des Wertpapierhandelsgesetzes zur Geld- und Vermögensanlage

Annex 3: Unit 3 – sell financial products and services (NVQ Level 3, n.d.)

This unit relates to non-mortgage and non-regulated products and services, such as non-regulated insurance products, investments and foreign currency. You will need to show that you can identify your customer's needs and promote suitable products and services by presenting them accurately and seeking your customer's commitment to buy.

Elements

- 3.1. Establish customer needs
- 3.2. Promote the features and benefits of the products and services
- 3.3. Gain customer commitment and evaluate sales techniques

Knowledge and understanding

For each of the elements you must show that you know and understand the following:

- your organisation's requirements relating to relevant codes, laws and regulations for selling products and services
- which products and benefits of products and services you are authorised to promote
- key features and benefits of products and services you are responsible for promoting
- your organisation's sales process relevant to your area of responsibility
- the limits of your authorisation and responsibility when providing information and offering advice on your organisation's products and services
- to whom you should refer customers for information and advice outside of your authorisation and responsibility
- questioning techniques, such as when to use open or closed questions while selling
- the types and forms of buying signals your customers might use
- how to monitor and evaluate your sales technique.

Element 3.1: establish customer needs

Performance criteria

You will need to:

- 3.1.1 Give customers prompt attention and treat them politely
- 3.1.2 Give your customer, when necessary, a clear and accurate account of your role, and the levels of information and advice that you can provide
- 3.1.3 Establish the needs of your customer through discussion, using appropriate questioning
- 3.1.4 Check your customer's responses to ensure you understand them
- 3.1.5 Conduct your discussion in a manner appropriate to your customer's needs
- 3.1.6 Record all relevant information accurately
- 3.1.7 Pass promptly to the relevant person any sales opportunity outside your area of knowledge or responsibility

Evidence requirements

You must be able to demonstrate your competence at establishing customer needs. You must provide evidence from real work activities, undertaken by yourself, that:

- you have established customer needs through:
 - your discussion with the customer
 - your records of these meetings

- you have established needs of customers that are:
 - immediate
 - future.

Examples of evidence

Performance

- Interview/discussion notes and records
- Referral documentation, where sales opportunities were passed on to a relevant colleague

Knowledge and understanding

Make sure you can answer the following questions:

- What are the stages that you follow in selling your organisation's products and services? How might these be adapted to meet the particular needs of different customers?
- What different types of customer needs are there? Why is it important to prioritise these needs when speaking to individual customers?
- In what ways can you check that customers understand the information that you are giving?
- What are the limits of your authority and responsibility when advising customers about your organisation's products and services? What should you do if these limits are reached?

Element 3.2: promote the features and benefits of the products and services

Performance criteria

You will need to:

- 3.2.1 Explain clearly and accurately the features and benefits of products and services that most closely match your customer's needs
- 3.2.2 Identify, issue and explain fully to your customer the promotional material relating to products and services that meet their needs
- 3.2.3 Follow your organisation's procedures to ensure that the options relating to the products and services you offer to your customer conform to relevant codes, and legal and regulatory requirements
- 3.2.4 Address fully and accurately your customer's questions and concerns in a manner that promotes the sale
- 3.2.5 Identify buying signals and correctly act on these
- 3.2.6 Confirm with your customer their understanding of the proposed products and services
- 3.2.7 Identify opportunities for cross-selling products and services and promote these clearly to your customer in a manner that maintains goodwill
- 3.2.8 Pass promptly to the relevant person sales opportunities outside your area of knowledge or responsibility
- 3.2.9 Record all relevant information accurately

Evidence requirements

You must be able to demonstrate your competence at promoting the features and benefits of products and services through:

- your sales performance with customers
- your corresponding customer records.

Examples of evidence

Performance

- Interview/discussion notes and records
- Referral documentation, where sales opportunities were passed to a relevant colleague
- Your customer records

Knowledge and understanding

Make sure you can answer the following questions:

- What is the difference between ‘features’ and ‘benefits’? How does this apply to the different products and services you deal with?
- What are your organisation’s procedures and requirements for ensuring that you meet relevant codes, laws and regulations when selling products and services?
- What sort of buying signals might customers use to show their interest, or disinterest, in the products and services being offered? How could you deal with the different types of buying signals?
- What opportunities might you encounter for cross-selling? How would you deal with these?
- What types of sales opportunities are outside your area of responsibility? To whom do you refer them?

Element 3.3: gain customer commitment and evaluate sales technique*Performance criteria*

You will need to:

- 3.3.1 Agree the preferred products and services with your customer
- 3.3.2 Establish and agree the way forward with your customer, and record this accurately
- 3.3.3 Promptly and accurately complete documentation and check that it is signed by your customer in accordance with your organisation’s procedures under the relevant codes, and legal and regulatory requirements
- 3.3.4 Deal with documentation in accordance with your organisation’s procedures
- 3.3.5 Inform relevant parties of the outcome according to your organisation’s procedures
- 3.3.6 Conduct all business in a manner that maintains professional integrity and goodwill
- 3.3.7 Evaluate objectively the effectiveness of the sales technique and use this to influence future sales activity

Evidence requirements

You must be able to demonstrate your competence at achieving a purchase commitment and evaluating your sales technique through:

- your sales performance with customers
- your relevant records
 - You must provide evidence from real work activities, undertaken by yourself, that:
- you agreed the way forward with customers, to include both:
 - an agreed sale or successful referral with at least one customer
 - sales contact with at least one other
- you can objectively evaluate your own sales technique, through:
 - your records
 - discussions with colleagues and your assessor.

Examples of evidence

Performance

- Interview/discussion notes and records
- Product/service application documents completed by yourself and signed by your customer
- Customer records
- Notes/memos notifying the appropriate person(s) of the outcome of your sales discussions
- Notes of your evaluation of your sales technique with proposals for the future.

Knowledge and understanding

Make sure you can answer the following questions:

- What procedures should you follow for completing a sale, including gaining customer commitment, and subsequently completing and forwarding necessary documentation?
- What factors in your sales technique might influence whether a customer agrees to buy from your organisation?
- How could you check the effectiveness of differences in your sales technique?

References

- AAC-project. *Accreditation of achieved competencies*. Leonardo da Vinci project. Wageningen: Stoas Research, 2001.
- Achtenhagen, F. et al. *Social sciences. COST technical committee*. Brussels: European Commission, DG Science, Research and Development, 1995, Vol. 3.
- Achtenhagen, F.; Thang, P. O. *Mehrdimensionale Lehr-Lern-Arrangements*. Wiesbaden: Gabler, 2002, p. 444-459.
- AERA - American educational research association. *AERA position statement concerning high stakes testing in PreK-12 education*. Washington, DC: AERA, 2000. Available from Internet: <http://www.aera.net/about/policy/stakes.htm> [cited 13.2.2004].
- AERA; APA; NCME. *Standards for educational and psychological testing*. Washington, DC: AERA – American Educational Research Association, 1999.
- Alheit, P. *Biographien in der späten Moderne: Verlust oder Wiedergewinn des Sozialen?* Bremen: IBL, 1998. Available from Internet: www.ibl.uni-bremen.de/publik/vortraege/98-01-pa.htm [cited 13.2.2004].
- Anderson, L. W.; Krathwohl, D. R. *A taxonomy for learning, teaching and assessing*. New York: Longman, 2001.
- Anderson, N. Decision making in the graduate selection interview. *Journal of Occupational Psychology*, 1990, No 63, p. 63-76.
- Baker E. L. et al. Policy and validity prospects for performance-based assessment. *American Psychologist*, 1993, No 48, p. 1210-1218.
- Baumert, J. et al. *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich*. Opladen: Leske und Budrich, 1997.
- Baumert, J. et al. *TIMSS/III – Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Opladen: Leske und Budrich, 2000, Vol. 1.
- Beck, K. *Die empirischen Grundlagen der Unterrichtsforschung. Eine kritische Analyse der deskriptiven Leistungsfähigkeit von Beobachtungsmethoden*. Göttingen: Hogrefe, 1987.
- Beck, K.; Krumm, V. *Wirtschaftskundlicher Bildungs-Test*. Göttingen: Hogrefe, 1998.
- Berufsbildungsbericht 2003. In: *BMBF Publik.* Bonn: Bundesministerium für Bildung und Forschung (ed.), 2003.
- Björnavold, J. *Making learning visible: identification, assessment and recognition of non-formal learning in Europe*. Luxembourg: Office for Official Publication of the European Communities, 2000 (Cedefop Reference series).
- Björnavold, J.; Tissot, P. Glossary. In: Björnavold, J. *Making learning visible: identification, assessment and recognition of non-formal learning in Europe*. Luxembourg: Office for Official Publication of the European Communities, 2000, p. 199-221 (Cedefop Reference series).
- BMBF. *IT-Weiterbildung mit System. Neue Perspektiven für Fachkräfte und Unternehmen*. Bonn: BMBF, 2002.
- Bloom, B. S. (ed) *Taxonomy of educational objectives: handbook I: cognitive domain*. New York: Adavid McKay, 1956.
- Boyatzis, R. E. *The competent manager: a model for effective performance*. New York: Wiley, 1982.
- Bransford, J. D. et al. *How people learn*. Washington, DC: National Academic Press, 2000.
- Brown, S. *Fit for purpose assessment*. Paper presented at the EARLI conference, Gothenburg, August 1999.
- Bulmahn, E.; Wolff, K.; Klieme, E. *Zur Entwicklung nationaler Bildungsstandards: eine Expertise*. Frankfurt a.M.: IPF/BMBF, 2003. Available from Internet: http://www.dipf.de/aktuelles/expertise_bildungsstandards.pdf [cited 16.2.2004].
- Bundesanzeiger. Bekanntmachung der Verordnung über die Berufsausbildung zum Bankkaufmann/zur Bankkauffrau nebst Rahmenlehrplan vom 23. Januar 1998. *Bundesanzeiger*, 1998, Vol. 50, No 82a.
- Campbell, D. T.; Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In: *Handbook of research on teaching*. Skokie, Ill: Rand McNally, 1963, p. 171-246.
- Cedefop. *Vocational education and training – the European research field: background report*

- (two volumes). Luxembourg: Office for Official Publication of the European Communities, 1998 (Cedefop Reference series).
- Chomsky, N. Rules and representations. *The behavioural and brain sciences*, 1980, Vol. 3, p. 1-61.
- Cseh, M. et al. Informal and incidental learning in the workplace. In: Straka, G. A. (ed.) *Conceptions of self-directed learning*. Münster: Waxmann, 2000, p. 59-74.
- CTGV. *The Jasper project*. The cognition and technology group at Vanderbilt. Mahwah, NJ: Lawrence Erlbaum, 1997.
- Deutsches PISA-Konsortium (eds). *PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske und Budrich, 2001.
- Deutsches PISA-Konsortium. *Die Bundesländer der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske und Budrich, 2002.
- DIHT. *Abschlussprüfung Bankkaufmann/Bankkauffrau: Ausbildungsordnung vom 30. Dezember 1997: Prüfungsfach 'Kundenberatung'*. Bonn: Deutscher Industrie- und Handelstag, 1999.
- DIN. *DIN 33430: Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen [job related proficiency assessment]*. Berlin: Deutsches Institut für Normung, 2002.
- Dörner, D. *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer, 1976.
- Embretson, S. E.; Reise, S. P. *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum, 2000.
- Eraut, M. *Developing professional knowledge and competence*. London: Taylor and Francis, 1994.
- Eraut, M. Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, 2000, No 70, p. 113-136.
- Eraut, M. The role and use of vocational qualifications. *National Institute Economic Review*, 2001, No 178, p. 88-98.
- Eraut, M. Editorial. *Learning in health and social care*, 2003a, p. 1-3.
- Eraut, M. National vocational qualifications in England: description and analysis of an alternative qualification system. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003b, p. 117-123.
- Eraut, M. et al. *Evaluation of higher level S/NVQs*. Sussex: University, 2001 (Research report).
- Ertl, H. *Modularisation of vocational education in Europe: NVQs and GNVQs as a model for the reform of initial training provisions in Germany?* Wallingford: Symposium Books, 2000.
- Ertl, H. *The role of EU programmes and approaches to modularisation in vocational education: fragmentation or integration?* München: Herbert Utz Verlag, 2002.
- European Commission. *A memorandum on lifelong learning*. Brussels: European Commission – DG for Education and Culture, 2000 (SEC (2000) 1832).
- European Commission. *Making a European area of lifelong learning a reality: communication from the Commission*. Luxembourg: Office for Official Publication of the European Communities, 2001 (Documents COM, (2001) 678).
- Gagné, R. M. *The conditions of learning*. New York: Holt, Rinehart and Winston, 1977.
- Gilbert P. *L'évaluation des compétences à l'épreuve des faits*. Luxembourg: Entreprise et Personnel, 1998.
- Grant, G. et al. *On competence: a critical analysis of competence-based reforms in higher education*. San Francisco: Jossey-Bass, 1979.
- Grunwald, S. Zertifizierung arbeitsprozessorientierter Weiterbildung. In: Rohs, M. (ed.) *Arbeitsprozessintegriertes Lernen: neue Ansätze für die berufliche Bildung*. Münster: Waxmann, 2002, p. 165-179.
- Grunwald, S.; Rohs, M.. Zertifizierung informell erworbener Kompetenzen im Rahmen des IT-Weiterbildungssystems. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003, p. 207-221.
- Gutschow, K. Erfassen, Beurteilen und Zertifizieren non-formell und informell erworbener beruflicher Kompetenzen in Frankreich. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003, p. 127-139.
- Kerlinger, F. N. *Foundations of behavioural research*. London: William Clowes, 1973.
- Klauer, K. J. *Revision des Erziehungsbegriffs*. Düsseldorf: Schwann, 1973.
- Klieme, E. et al. *Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung*

- von Kompetenzen. In: Baumert, J. et al. *TIMSS/III*, Vol. 1. Opladen: Leske und Budrich, 2000, p. 85-133.
- KMK. *Handreichungen für die Erarbeitung von Rahmenplänen der KMK für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe*. Sekretariat der ständigen Konferenz der Kultusminister der Länder in der BRD, 1996/2000.
- Krumm, V.; Weiß, S. Ungerechte Lehrer. *Psychosozial*, 2000, Vol. 23, No 79, p. 57-73.
- Kuwan, H.; Thebis, F. *Berichtssystem Weiterbildung VIII: Erste Ergebnisse der Repräsentativbefragung zur Weiterbildungssituation in Deutschland*. Bonn: Bundesministerium für Bildung und Forschung, 2001.
- Lienert, G. A.; Raatz, U. *Testaufbau und Testanalyse*. München: Beltz Psychologie Verlags Union, 1998.
- Linn, R. L. Assessments and accountability. *Educational Researcher*, 2000, Vol. 29, No 2, p. 4-16.
- Linn, R. L.; Baker, E. L. Comparing results from disparate assessments. *CRESST Line*, Winter 1993, p. 1-3.
- Linn, R. L.; Gronlund, N. E. *Measurement and assessment teaching*. New Jersey: Prentice Hall, 2000.
- Ministry of Education. *DACUM booklets*. British Columbia: Ministry of Education, 1983.
- Nijhof, W. J. Assessing competencies at work – reflections on a case in the Netherlands. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003, p. 141-151.
- Norris, N. The trouble with competence. *Cambridge Journal of Education*, 1991, Vol. 21, No 3, p. 331-341.
- NRC – National Research Council. *Knowing what students know: the science and design of educational assessments*. Committee on the Foundations of Assessment. Washington, DC: National Academies Press, 2001.
- NVQ. *Providing financial services (banks and building societies)*. NVQ Level 2. Standards and assessment guidance. The chartered institute of bankers. n.d.
- NVQ. *Providing financial services (banks and building societies)*. NVQ Level 3. Standards and assessment guidance. The chartered institute of bankers. n.d.
- CIB. *Providing financial services (banks and building societies)*. NVQ Level 2. Standards and assessment guidance. The chartered institute of bankers, n.d.
- CIB. *Providing financial services (banks and building societies)*. NVQ Level 3. Standards and assessment guidance. The chartered institute of bankers, n.d.
- OECD. *Knowledge and skills for life – first results from PISA 2000*. Paris: OECD, 2001.
- Oerter, R. *Moderne Entwicklungspsychologie*. Donauwörth: Auer, 1980.
- O'Grady, M. Assessment of prior achievement/assessment of prior learning: issues of assessment and accreditation. In: *The Vocational Aspect of Education*, 1991, No 115.
- Pettersen, I. L. (ed.) *Competence card from the workplace. A cooperation between trade and industry and the educational system in Nordland.*, n.d.
- Pettersen, I. L. (ed.) *CV. Curriculum Vitae. A cooperation between trade and industry and the educational system in Nordland*, n.d.
- Polanyi, M. *The tacit dimension*. Garden City, NY: Doubleday, 1966.
- Popper, K. R. *Logik der Forschung*. Tübingen: Mohr, 1989.
- Porter, M. E. *Competitive advantage*. New York: Macmillan, 1985.
- Pütz, H. 'Berufsbildungs-PISA' wäre nützlich. *Berufsbildung in Wissenschaft und Praxis*, 2002, Vol. 31, No 3, p. 34.
- QCA. *Assessing NVQs. March 1998, QCA Guarding Standards*. London: QCA – Qualifications and Curriculum Authority, 1998.
- QCA. *Developing an assessment strategy for NVQs and SVQs*. London: QCA – Qualifications and Curriculum Authority, 1999 (Ref: QCA/99/396).
- Resnick, L. B. Learning in school and out. *Educational Researcher*, 1987, Vol. 12, p. 13-20.
- Savisaari, L. *Youth academy & recreational activity study book system*. Helsinki: Nuorten Akatemia, 2002. Available from Internet: http://www.nuortenakatemia.fi/index.php?lk_id=64 [cited: 17.2.2004].
- Savisaari, L. The recreational activity study book system as a tool for certification of youth non-formal and informal learning in Finland – a practitioner's approach. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erwor-*

- bener beruflicher Kompetenzen. Münster: Waxmann, 2003, p. 223-229.
- Schuler, H. *Psychologische Personalauswahl*. Göttingen: Verlag für Angewandte Psychologie, 2000.
- Sembill, D. *Problemlösefähigkeit, Handlungskompetenz und emotionale Befindlichkeit*. Göttingen: Verlag für Psychologie, 1992.
- Shepard, L. A. The role of assessment in a learning culture. *Educational Researcher*, 2000, Vol. 29, No 7, p. 4-14.
- Shepard, L. A. The role of classroom assessment in teaching and learning. In: Richardson, V. (ed.) *Handbook of research on teaching* (fourth edition). Washington, DC: AERA – American Educational Research Association, 2001, p. 1066-1101.
- Sternberg, R. J.; Kolligian, J. Jr. *Competence considered*. New Haven: Yale University Press, 1990.
- Straka, G. A. *Forschungsstrategien zur Evaluation von Schulversuchen*. Weinheim: Beltz, 1974.
- Straka, G. A. (ed.) *European views of self-directed learning*. Münster: Waxmann, 1997.
- Straka, G. A. (ed.) *Conceptions of self-directed learning*. Münster: Waxmann, 2000.
- Straka, G. A. Action-taking orientation of vocational education and training in Germany – a dynamic concept or rhetoric? In: Nieuwenhuis, L. F. M.; Nijhof, W. J. (eds) *The dynamics of VET and HRD systems*. Enschede: Twente University Press, 2001a, p. 89-99.
- Straka, G. A. Lern-lehr-theoretische Grundlagen der beruflichen Bildung. In: Bonz, B. (ed.) *Didaktik der beruflichen Bildung*. Baltmannsweiler: Schneider Verlag, 2001b, p. 6-30.
- Straka, G. A. Valuing learning outcomes acquired in non-formal settings. In: Nijhof, W. J.; Heikkinen, A.; Nieuwenhuis, L. F. M. (eds) *Shaping flexibility in vocational education and training*. Dordrecht: Kluwer, 2002a, p. 149-165.
- Straka, G. A. Empirical comparisons between the English national vocational qualifications and the educational and training goals of the German dual system – an analysis for the banking sector. In: Achtenhagen, F.; Thång, P.-O. (eds) *Transferability, flexibility and mobility as targets of vocational education and training*. Proceedings of the final conference of the COST Action A11, Gothenburg, 13 and 16 June, 2002b, p. 227-240.
- Straka, G. A. Rituale zur Zertifizierung formell, non- und informell erworbener Kompetenzen. *Wirtschaft und Erziehung*, 2003a, Vol. 55 (7-8), p. 253-259.
- Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003b.
- Straka, G. A. *Verfahren und Instrumente der Kompetenzdiagnostik – der Engpass für ein Berufsbildung-PISA?* (In preparation).
- Straka, G. A.; Lenz, K. Was trägt zur Entwicklung von Fachkompetenz in der Berufsausbildung bei? *Schweizerische Zeitschrift für kaufmännisches Bildungswesen*, 2003, Vol. 97, p. 54-67.
- Straka, G. A.; Macke, G. *Lern-Lehr-Theoretische Didaktik*. Münster: Waxmann, 2002.
- Straka, G. A.; Macke, G. Handlungskompetenz und Handlungsorientierung als Bildungsauftrag der Berufsschule – Ziel und Weg des Lernens in der Berufsschule? *Berufsbildung in Wissenschaft und Praxis*, 2003, Vol. 32, No 4, p. 43-47.
- Straka, G. A.; Schaefer, C. *Validating a more-dimensional conception of self-directed learning*. Paper presented at the international research conference of the AHRD, in Honolulu, Hawaii, from 27 February to 3 March 2002.
- Thommen, J.-P. *Allgemeine Betriebswirtschaftslehre. Umfassende Einführung aus managementorientierter Sicht*. Wiesbaden: Gabler, 1991.
- Thömmes, J. Bilan de compétences. In: Erpenbeck; J.; von Rosenstiel, L. (eds) *Handbuch Kompetenzmessung*. Stuttgart: Schäffer-Poeschel, 2003, p. 545-555.
- Ulrich, H. *Die Unternehmung als produktives soziales System*. Bern: Haupt, 1970.
- VOX. *The Realkompetanse Project*. Validation of non-formal and informal learning in Norway. Oslo: VOX, 2002.
- Watermann, R.; Baumert, J. Mathematische und naturwissenschaftliche Grundbildung beim Übergang von der Schule in den Beruf. In: Baumert, J.; Bos, W.; Lehmann, R. (eds) *TIMSS/III – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Opladen: Leske und Budrich, 2000, Vol. 1, p. 199-259.
- Weinert, F. E. Concept of competence: a conceptual clarification. In: Rychen, D. S.; Salganik, L. H. (eds) *Defining and selecting key competencies*. Göttingen: Hogrefe, 2001, p. 45-66.

- Weinert, F. E. *Concepts of competence*. München: Max Planck Institut, 1999. Available from Internet: http://www.statistik.admin.ch/stat_ch/ber15/deseco/weinert_report.pdf [cited: 18.2.2004]
- Weisschuh, B. Nachhaltige Qualitätssicherung bei der Beurteilung und Förderung von Schlüsselqualifikationen im Produktionsbetrieb. In: Straka, G. A. (ed.) *Zertifizierung non-formell und informell erworbener beruflicher Kompetenzen*. Münster: Waxmann, 2003, p. 231-242.
- Wöhe, Günter. *Einführung in die Allgemeine Betriebswirtschaftslehre*. München: Vahlen, 2002.
- Wolf, A. *Competence-based assessment*. Buckingham: Open University Press, 1995.
- Wolf, A. Competence-based assessment. Does it shift the demarcation line? In: Nijhof, W. J.; Streumer, J. N. (eds) *Key qualifications in work and education*. Dordrecht: Kluwer Academic Press, 1998, p. 207-220.