

DISCUSSION PAPERS

**No. 119**

Philipp Grollmann,  
Melanie Hoppe  
(Ed.)

**Methods and instruments  
for the evaluation and monitoring  
of VET- systems**

**Conference proceedings**

Königswinter, Germany  
10<sup>th</sup> and 11<sup>th</sup> December 2009  
PRELIMINARY DRAFT

Federal Institute for  
Vocational Education and Training  
Publication  
Bonn

Federal Institute for  
Vocational Education  
and Training

**BiBB**

► Researching  
► Advising  
► Shaping the future

The DISCUSSION PAPERS of the Federal Institute for Vocational Education and Training, („BIBB“) are issued by the Institute's President and are published under the respective author's name. The opinions and views expressed protected by copyright. They are published to foster discussion among players in the vocational education and training field.



The content of this work is under a Creative Commons license (Licence typ: Attribution – NonCommercial – NoDerivs 3.0 Germany).

The work is protected by copyright and/or relevant laws. Any use that is not expressly permitted by this license or copyright is prohibited.

Further information can be found on the Internet on our Creative Commons info page <http://www.bibb.de/cc-lizenz>.

**Distributed by:**

Bundesinstitut für Berufsbildung  
A 1.2 Kommunikation - VÖ  
53142 Bonn

**Order No.: 14.119**

Copyright 2008 by Bundesinstitut für Berufsbildung, Bonn  
Published by: Bundesinstitut für Berufsbildung, Bonn  
Internet: [www.bibb.de](http://www.bibb.de)  
E-mail: [zentrale@bibb.de](mailto:zentrale@bibb.de)  
Cover design: Hoch Drei, Berlin  
Publisher: Bundesinstitut für Berufsbildung, Bonn  
Printed in Germany

ISBN 978-3-88555-888-0

Diese Netzpublikation wurde bei Der Deutschen Bibliothek angemeldet und archiviert.  
URN: [urn:nbn:de:0035-0432-9](http://nbn-resolving.org/urn:nbn:de:0035-0432-9)

## Content

*Dr. Philipp Grollmann, Melanie Hoppe*

Preface “Research meets Advisory Service” .....5

*Dr. Gisela Dybowski*

Eröffnung des Expert Workshop „Methods and Instruments for the Evaluation  
and Monitoring of VET Systems“ .....7

*Dr. Gudrun Steeger*

VET Research and Measurement of Competencies by the BMBF ..... 10

*Kathrin Hoeckel*

Learning for Jobs The OECD Policy Review of Vocational Education and Training ..... 18

*Ute Hippach-Schneider*

Challenges and Chances of a European monitoring instrument in VET .....28

*Dr. Torsten Dunkel*

Reviewing European VET Policy - challenges for comparative analyses ..... 39

*Stefan Thomas*

Results Measurement and Programme-based Approaches (PBA) .....55

*Noela Eddington and Ian Eddington*

Integrating VET into other policy domains: some thoughts about monitoring  
and evaluation .....62

*Dr. Xiaorong Nie, Mr. Ziyuan Wang, Mrs. Zhibao Pan*

The Study on the Evaluation of Secondary Vocational School in China ..... 71

*Dr. Uwe Lauterbach*

Monitoring of VET systems – The perspective of comparative education ..... 87

*Prof. Dr. Ute Clement, Dr. Philipp Grollmann and Dr. Matthias Pilz*

International Handbook on VET Presentation of a draft modified monitoring grid ..... 105

*Prof. Dr. Marc Schütte and Prof. Dr. Georg Spöttl*

Entry-Level-Competence Tests: A Means for Evaluation the Fulfilment of  
Occupational Standards ..... 112

*Dr. Lorenz Lassnigg, Dr. Stefan Vogtenhuber*

Monitoring of Qualification and Employment in Austria – an empirical approach  
based on the labour force survey (LFS) ..... 122

*Prof. Jim Stewart and Victoria Harte*

Using Survey Instruments in Longitudinal Evaluations of Vocationally Orientated  
Higher Education: Problems and Pitfalls ..... 137

*Dr. Geoff Hayward*

The use of large-scale administrative data sets to monitor progression from VET  
into Higher Education in the UK: possibilities and methodological challenges ..... 147

*Dana Stroie*

Methods and instruments for quality assurance of the initial and vocational education  
and training system in Romania ..... 161

*Theodorah. U.Ogali*

Gender Sensitivity in Instrument Designs and Evaluation of Vocational Education  
and Training (VET) Systems ..... 171

*Prof. Pilar Pineda, Carla Quesada & Victoria Moreno*

The “Learning Transfer System Inventory” (LTSI) as a diagnostic tool of training  
effectiveness in Spain ..... 178

*Sandra Speer*

Peer Evaluation: Facilitating Changes on the System-Level..... 193

*Dr. Benjamin A. Ogwo*

Qualitative Techniques for Evaluating Vocational Education and Training  
(VET) Systems ..... 202

*Simonas Gaušas and Rimantas Dumčius*

Combining quantitative and qualitative research methods in comparative study  
on VET cost-sharing mechanisms and regulatory instruments ..... 211

*Dr. Ludger Deitmer, Dr. Lars Heinemann, Prof. Dr. Felix Rauner*

*and Dr. Wolfgang Wittig*

Qualitative Evaluation of Governance Structures in VET ..... 224

## **Preface “Research meets Advisory Service”**

*Dr. Philipp Grollmann, Melanie Hoppe*

“We advise international partners on the development and modernisation of vocational education and training ...” says the mission statement of BIBB. At the same time it expresses BIBB’s interest “... to foster the efficiency and competitiveness of the German system of vocational education and training and contribute to the creation of the European Education Area.” These two sides of one coin are the foundation for BIBB’s international activities and the existence of the two sections “International Benchmarking and Monitoring/European VET policies” and “International Cooperation and Advisory Services”.

With regard to the latter, BIBB has developed a concept for cooperating with countries that request advisory services in their on-going reforms (BIBB 2009). Reform processes with the objective of shaping vocational education and training in a sustainable manner and geared towards employability have been and are being supported. In doing so, the following guidelines give direction: The delivered advisory services always tie in with existing structures and advice is viewed as providing impetus for systemic development processes.

BIBB also contributes and has contributed to significant European developments such as the introduction of the European Qualifications Framework (EQF), the implementation of national frameworks as well to the development of a European Credit System for Vocational Education (ECVET). Increasingly, BIBB is also engaged into the monitoring of VET systems. E. g. BIBB contributes regularly to the CEDEFOP ReferNet (e. g. Hippach-Schneider, Woll, Hanf 2008) over the last years, and it was asked to take over the managing editorship of the “International Handbook on VET” (IHBB, Lauterbach 1994 ff.) from 2010 together with seven German universities.

For both tasks a solid basis of information is needed on the state of vocational education and training systems as well as concepts of shaping vocational education systems. Due to this BIBB also conducts its own international comparative research projects to generate solid information and sustainable instruments for VET systems’ analysis. These research activities deal with questions such as: How do we know what is needed for the further development of a VET system? Why is one approach chosen instead of the other? Which results are to be expected and how can we monitor this?

This set of questions and the needs for further development as regards to such instruments was the background for the organisation of the workshop “Methods and Instruments for the Evaluation and Monitoring of VET systems”. The workshop was preceded by an international call for papers that was disseminated through the pertinent channels of CEDEFOP, UNESCO-UNEVOC and others. Papers have been selected through a blind review proce-

dure according to their scientific quality and the contribution to the abovementioned questions.

Taking a look at the same issue from different angles, combining the joint interests of the two international sections of BIBB, the workshop provided a platform for international researchers and practitioners to discuss and address their needs. Therefore the following central topics were addressed under the headline "Research meets Advisory Service" during the workshop:

System analysis :A general overview of the available analyses of the genesis and development of VET systems;

Indicators & benchmarks: Supervising the 'status quo' or heading towards certain targets - both instruments are used for international comparison and for advising national policy makers;

Peer review & peer learning: In order to capture opinions of experts and facts and experiences that cannot be depicted through indicators solely, different models for the instruments are in use.

This documentation provides the results of the expert workshop which was held in Bonn/Königswinter in December 2009. We want to thank the participants for sharing their knowledge. A special acknowledgement needs to be made of the contributions by the reviewers who have helped in the selection of papers, namely Prof. Dr. Ute Clement, University of Kassel and Sören Nielsen, European Training Foundation, as well as the local organization team, Ms. Barbara Schulte and Ms. Heidemarie Ranfft.

## **References**

BIBB 2009: Modernising Vocational Education and Training. BIBB International Advisory Services. Bonn.

Hippach-Schneider, Ute; Woll, Christian; Hanf, Georg 2008: VET Policy Report 2008. Bonn.

Lauterbach, Uwe et. al (Eds.)1994 ff.: Internationales Handbuch der Berufsbildung. Bielefeld: W. Bertelsmann Verlag.

# **Eröffnung des Experten Workshops „Methods and Instruments for the Evaluation and Monitoring of VET Systems“**

*Dr. Gisela Dybowski*

**Ladies and Gentlemen, dear colleagues and friends!**

On behalf of BIBB and as head of the International Department in BIBB I'd like to give you a very warm welcome here in Königswinter. We are very pleased having you here as our guests and experts taking part in the Expert Workshop "Methods and Instruments for the Evaluation and Monitoring of VET Systems". Let me assure of being proud that so many experts followed our "call for papers" and our invitation to this workshop. We assume that the great response which we have received to the announcement of the workshop underlines the importance of the topics we have chosen for the discussion for today and tomorrow. Following the agenda of the workshops program, it is not my part but the part of my colleagues, Philip Grollmann und Melanie Hoppe, to introduce into aims and scope of the workshop later on. For that reason my opening remarks will focus on BIBB's role and engagement in international research and development activities in the field of VET via a short historical review and recent challenges, which will link to the topics of this workshop. Due to the limited time which requires to be short and prompt, I'd like to continue in German language.

Das BIBB feiert 2010 sein 40-jähriges Bestehen, denn es wurde 1970 gegründet. Vor 40 Jahren startete das BIBB zunächst als nationale Einrichtung für die Berufsbildung, doch schon bald nach seiner Gründung erreichten das Institut erste Anfragen und es bahnten sich internationale Kontakte an. Bevor das BIBB sich selbst international positionierte, nahmen Institute anderer Länder seine Aktivitäten wahr und suchten die Kooperation. Insbesondere die deutschsprachigen Nachbarländer (Schweiz und Österreich) zeigten Interesse an einem Informationsaustausch und an einer inhaltlichen Zusammenarbeit mit dem Bundesinstitut. Bald darauf gab es auch Anfragen nach einer Kooperation aus Frankreich (CEREQ), aus Italien (ISFOL), den Niederlanden (CINOP) und in den 1990er-Jahren darüber hinaus Anfragen aus den Mittel- und Osteuropäischen Staaten, wie Ungarn (NIVE) und Tschechien (VUOS) Institut für Berufsbildung sowie aus den Berufsbildungsforschungsinstituten in Moskau, St. Petersburg und Minsk. Zudem eröffnete die Mitarbeit in der European Training Foundation (ETF) und im CEDEFOP neue Formen der Kooperation mit zahlreichen Staaten inner- und außerhalb der EU. Die Mitarbeit an mehreren Projekten der OECD in den 1990er-Jahren bot ferner die Chance, auch die Zusammenarbeit mit außereuropäischen Partnern in den USA, Australien und in Schwellenländern auf- und auszubauen.

16 Jahre nach seiner Gründung wurde das BIBB erstmals evaluiert. Im Rahmen dieser Evaluierung fordert die „Kommission zur Untersuchung der wissenschaftlichen und wirtschaftlichen Effizienz des BIBB“, dass der vergleichenden Berufsbildungsforschung im BIBB eine stärkere Aufmerksamkeit geschenkt werden solle. Durch eine entsprechende Satzungsänderung im Februar 1989 wurde festgeschrieben, dass die Berufsbildungsforschung des Instituts „auch internationale, vorwiegend europäische Entwicklung“ berücksichtigen soll.

Wie eine vom BIBB anlässlich seiner 20-Jahrfeier ausgerichtete internationale wissenschaftliche Tagung zeigte, stand die vergleichende Berufsbildungsforschung jedoch Anfang der 1990er-Jahre international noch weitgehend am Anfang. Zwar wurden vergleichende Betrachtungen angestellt. Diese beleuchteten jedoch eher ausschnittshafte Fragen und partielle Aspekte beruflicher Bildung. Das zentrale Forschungsinteresse der vergleichenden Forschung liegt jedoch gerade darin, Aufschlüsse darüber zu gewinnen, mit welchen Lösungen andere Staaten auf vergleichbare Herausforderungen in der beruflichen Bildung reagieren, und diese Vorgehensweisen systematisch zu analysieren. Der Vergleich kann sich auf komplexe Funktionszusammenhänge (z. B. den Übergang vom Bildungs- zum Beschäftigungssystem) oder Funktionsweisen (z. B. die Formen der politischen Steuerung von Berufsbildungssystemen) richten. Eine weiterführende Forschungsperspektive besteht in der Systemevaluation. Dabei geht es darum, die Funktionsweise und Leistungsfähigkeit der betrachteten Berufsbildungssysteme anhand von generellen Kriterien bzw. „Benchmarks“ zu bewerten.

Im Rahmen der internationalen Aktivitäten des BIBB galt es, wesentliche Forschungs- und Entwicklungsarbeiten auf diese Aspekte zu konzentrieren. Denn Erkenntnisse der vergleichenden Berufsbildungsforschung sind zunehmend auch im Hinblick auf die Weiterentwicklung der beruflichen Bildung in Deutschland, vor allem im Kontext des Europäischen Integrationsprozesses gefragt. Systematisch erarbeitete Erkenntnisse über die Stärken und Schwächen des deutschen im Vergleich zu ausländischen Berufsbildungssystemen können dazu beitragen, die Reform der beruflichen Bildung in Deutschland voranzutreiben.

Ferner wurden mit der Erklärung von Lissabon und Kopenhagen Prioritäten für die (Berufs) Bildung gesetzt. Deutschland hat sich verpflichtet, dazu seinen Beitrag zu leisten und an der kontinuierlichen Bewertung der Fortschritte im Hinblick auf die Prioritäten mitzuwirken. Ein wesentliches Instrument in der europäischen (Berufs) Bildungszusammenarbeit ist dabei die Vereinbarung von Indikatoren und Benchmarks, die als Basis der Berichterstattung zum Arbeitsprogramm 2010 der EU und zum systematischen Vergleich der Entwicklung der Systeme dienen.



Lassen Sie mich zum Schluss kommen. Festzustellen ist, Europa hat starke Impulse zur Internationalisierung von Forschung und Entwicklung in der Berufsbildung gegeben. Das BIBB wird mit seinen Forschungs- und Entwicklungsarbeiten den Prozess der europäischen Zusammenarbeit in der beruflichen Bildung weiter unterstützen. Dazu gehört neben der Beobachtung und Analyse der europäischen Berufsbildungspolitik und deren Wirkungen auf die deutsche Berufsbildungspolitik und -praxis und die Beteiligung an europäischen Erhebungen, Studien und Analysen. Ein Schwerpunkt laufender Arbeiten sind deshalb internationale Vergleiche der Leistungsfähigkeit von Berufsbildungssystemen. Ferner gibt das BIBB einen jährlichen Datenreport zur beruflichen Bildung heraus, der eine indikatorengestützte Berichterstattung vorsieht. Internationale Vergleichsdaten und Indikatoren sollen in diesen Datenreport einfließen und die nationale Berichterstattung erweitern. Schließlich hat sich das BIBB als Anbieter internationaler Bildungsdienstleistungen in den letzten Jahren stark exponiert. Daraus erwachsen wesentliche Anforderungen an ein systematisches Monitoring der Entwicklung der Berufsbildung in anderen Staaten. Gefragt sind darüber hinaus aber auch Instrumente und Methoden für die Evaluierung von Berufsbildungssystemen als Grundlagen für eine Systemberatung.

Motto der internationalen Aktivitäten des BIBB ist das „voneinander lernen“. Dazu ist auch der heutige und morgige Workshop gedacht, der auf Instrumente und Methoden des Monitoring und der Evaluierung von Berufsbildungssystemen fokussiert. Ich wünsche uns eine fruchtbare und interessante Diskussion.

# VET Research and Measurement of Competencies by the BMBF

*Dr. Gudrun Steeger*

## **Agenda:**

1. Information concerning the reform process in vocational education and training in Germany
2. Monitoring systems to support the reform process
  - a) Report on Vocational Education and Training as a monitoring system
  - b) Education Reporting
3. Measurement of competencies
  - a) National Educational Panel Study
  - b) VET-LSA initiative
  - c) Plans for BMBF research initiative "Technology-oriented Assessment of Vocational Skills"
  - d) Promoting scientific exchanges by funding scientific forums and workshops
4. Prospects

## **Ladies and Gentlemen,**

Ad 1. Germany may be poor in resources as far as classical primary factors are concerned; what it does have in abundance, however, are core competences in the field of human capital. We need an efficient education and vocational training system in order to expand these capacities and make targeted use of them in a national and international context. Last year, the Federal Government embarked on this task and introduced its **Qualification Initiative**. This was the starting signal for a process of improvement across the entire vocational education and training system.

The **Education Summit** of the Federal Chancellor and the Heads of *Länder* governments marked an important stage in this process and led to catalogue of measures to secure the next generation of skilled workers and improve the education system.

These measures include the following:

- The Federal Government and *Länder* will work together to increase flexibility and compatibility in the education system;
- The **performance levels** of adolescents, particularly adolescents from a migrant background, will be improved;
- The **transition from vocational education and training to higher education** will be made easier; and
- **Part-time study courses and continuing education programmes** will be developed for this purpose.

**European cooperation** in vocational education and training is playing an increasingly important role in this context. It is our aim to establish a European Education Area and we are currently focusing our initiatives on:

- strengthening the European dimension,
- improving transparency, information and counselling,
- providing greater recognition for skills and qualifications, and
- improving quality assurance.

The instruments that we are using for these purposes are the European Qualifications Framework together with the German Qualifications Framework, the ECVET and DECVET credit point systems, and the European Quality Framework. All of these instruments represent new ground for us in many respects.

Providing support for these processes is an essential precondition for shaping an education system which is both fit for the future and in touch with reality. Measuring outcomes and output is therefore a necessary process for shaping a future-oriented education policy.

The BMBF already has various monitoring systems at its disposal – or in the process of being developed – in order to determine the productivity of the vocational education and training system.

## **2. Monitoring systems to support the reform process**

### **a) Report on Vocational Education and Training as a monitoring system**

The oldest instrument for monitoring the vocational education and training system and market is the annual Report on Vocational Education and Training, whose legal basis was established as far back as in 1976. It is the task of the BMBF to continuously monitor developments in the field of vocational education and training and to report to the Federal Government on 1 April every year. The Report on Vocational Education and Training is an excellent instrument for presenting developments in vocational training on a political as well as scientific level. It provides a platform for monitoring developments in education policy by supplying targeted statistical evaluations of the situation on the training and continuing education market, and documents activities on the political and scientific level. Over the last 32 years, the Report has established itself as a standard reference work for a specialist public with an interest in vocational education as well as for stakeholders in the vocational training system. It is also published in English and French. Starting in 2009, the Federal Institute for Vocational Education and Training is publishing the scientific part separately as a data report to accompany the Report on Vocational Education and Training.

### **b) Education Reporting**

In complement to the specialized Report on Vocational Education and Training, the national "Education in Germany" report provides a comprehensive overview of all phases of education, ranging from early learning to continuing education. The report also indicates the transitions and interfaces between different fields of education. By compiling relevant empirical data from different sources, the report provides an indicator-based – and thus updatable – analysis of the entire education system.

Compared with the Report on Vocational Education and Training, the "Education in Germany" report is a relatively new monitoring instrument, which is published every two years. The first national education report was commissioned by the Conference of *Länder* Ministers of Education (KMK) in association with the Federal Ministry of Education and Research (BMBF) and was published in June 2006.

The report is drafted by independent researchers. The KMK and BMBF have thereby established a form of reporting which also analyses problematic developments in the education system. For example, the 2008 report studied the so-called "transition system" from general education to vocational training.

"Education in Germany" identifies opportunities for improvements in individual fields of education and pinpoints potential for development. But it also highlights cross-cutting challenges

which can only be tackled successfully through cooperation between several fields of education and policy – for example, cooperation in the areas of education, economic, labour market and family policy.

All these monitoring systems aim to provide statistical data to demonstrate developments within education systems and to supply sufficient information for an evidence-based vocational education and training policy.

Apart from the monitoring systems, we also need a reliable assessment of education output. We can only make our (vocational) education system fit for the future if we know just how efficient it actually is. We have launched various activities in recent years with this in mind. These measures include diverse initiatives in the field of skills assessment.

### **3. Measurement of competencies**

First of all I would like to mention the

#### **a) National Education Panel Study (NEPS for short)**

Large-scale international student assessment studies like TIMSS, PISA, and IGLU have delivered important findings on the cross-sectional distribution of competencies among students attending schools in the Federal Republic of Germany - both in comparison with other nations and as a function of social origins and further variables. As valuable as these cross-sectional studies are, they can only be viewed as isolated snapshots documenting a specific state at a fixed time point in the life course. To transform such snapshots into a moving picture, the NEPS is following up individuals over time so that it can reconstruct how competencies unfold over the life course, how competencies and decision-making processes relate to various critical transitions in education careers, along with how and to what extent they are influenced by the family of origin and the structure of teaching and learning processes in pre-school, school, vocational training, higher education, and later (working) life.

The National Educational Panel Study (NEPS) has been set up to find out more about the acquisition of education in Germany, to plot the consequences of education for individual biographies, and to describe central education processes across the entire lifespan. The guiding principle of the NEPS is to ask how competencies unfold over the life course, how they influence or do not influence educational careers at various critical transitions, and how and to what extent competencies are influenced in turn by learning opportunities – not only those within the family and the peer group but also those resulting from the way teaching and learning processes are shaped in kindergarten, school, higher education, vocational training, and adult education.

Methodologically, the NEPS is based on a multicohort sequence design. Six starting cohorts will be recruited between 2009 and 2012. These will contain a total of more than 60,000 participants who will be surveyed over an extended period of time. Their competencies will also be assessed at set intervals. New starting cohorts will also be recruited and integrated into the study in later years in order to document and analyze historical changes in the way people pass through the various education-related transitions. The NEPS will not only deliver innovative impulses for basic research but also provide decisive information for policymakers. In particular, it will provide an important additional source of data for national education reporting, and it will strengthen the domains of “education in the life course” and “lifelong learning” along with our knowledge about developmental processes and trajectories.

Further information can be found at [www.bildungspanel.de](http://www.bildungspanel.de)

Beside NEPS, which is only able to measure cross-occupational competencies, we are also interested in measuring domain specific competencies. For this purpose, we already ordered a feasibility study in 2006. The initiative for a Large Scale Assessment in VET ( VET-LSA for short) developed from this study.

#### **b) VET-LSA initiative**

The VET-LSA set out to gain an insight into the strengths and weaknesses of different forms of vocational training and to establish opportunities for the various countries to learn from one another. A VET-LSA focuses on measuring skills that are acquired in the course of initial vocational training. Furthermore, it also aims to determine the institutional and individual conditions for educational processes, as these too are of decisive importance for individual skills development.

The findings of VET-LSA are intended to serve pragmatic political as well as scientific interests:

At the political level, they could be expected to produce a considerable amount of steering knowledge for shaping vocational training processes, particularly with regard to the following aspects:

Findings on the relationship between individual/biographical characteristics and forms of training to acquire skills (How do individual and institutional framework conditions influence individual skills?).

- The categorization of different vocational training qualifications in international classifications (ISCED; EQF) in order to improve recognition procedures and ensure the equality of certificates at European level.

- Establishing more transparency in vocational education and training in Europe in order to be able to better substantiate the equality of vocational qualifications.
- More detailed knowledge of the national links between the institutional regulations and skills outcomes of national vocational training processes in an international comparison, i.e. uncovering the strengths and weaknesses of different forms of training in the participating countries – not in the sense of winners and losers like in the Skills Olympics, but as an opportunity for all the countries to learn from one another.

1. On the scientific level, the following results were expected:

- The development of a reliable and valid procedure for an international comparative longitudinal study to measure the performance of vocational training organizations;
- The emergence of sound knowledge about the links between skill levels, forms of training and context variables as well as different skills dimensions (skills relating to a single occupation and inter-occupational skills);
- Findings on the links between vocational skills and outcomes on the labour market and in working life.

In contrast to the OECD studies, here a total of eight European states discussed the joint initiation and funding of such a project. According to the latest information, the planned VET-LSA study cannot be realized in the proposed form at the present time due to limited public budgets resulting from the financial crisis. In Germany, the initiative has attracted considerable interest from specialists; and outcome-oriented skills assessment is now recognized as a central national topic and field for the future. There will be further international opportunities to cooperate with other states below the VET-LSA threshold in future. The BMBF is currently planning a new research initiative at working level with this in mind.

**c) Plans for a BMBF "Technology-oriented Assessment of Vocational Skills" research initiative**

The aim of this research programme is an outcome-oriented skills assessment which can be used to measure the productivity of the education system, the quality of individual training establishments and the learning outcome of individuals. In recent years, Germany has taken part in important international comparisons in the schools sector and has put the skills of school students to the test. Similar tests in the field of vocational education are still outstanding since it is difficult to develop suitable procedures for as-

sessing our complex system of vocational education and training and no broad-based methods are available to which scientists can resort.

In order to ensure an effective and efficient evaluation of performance in various fields of vocational training, we must improve the scientific bases to enable a reliable assessment of vocational training in a national as well as international comparison. The main aim is to develop and test technology-oriented skills assessment procedures in various vocational fields. In addition, other assessment procedures are to be tested as to their reliability in measuring vocational skills. The procedure is to be developed to include different fields of application.

The **national research priority** will analyse the assessment procedure with regard to **four fields of application**:

- Are the procedures suitable for examinations at the end of training and continuing training?
- Can these procedures assess informal and non-formal skills in order to recognize these skills in the context of formal qualifications?
- Can skills be assessed reliably so that they can be credited towards other educational pathways (for example, training and continuing training measures or higher education courses)?
- In how far can the assessment of vocational skills help when checking the quality of in-company training or external training providers?

The **international funding priority** will realize projects similar to those of the VET-LSA concept in order to establish a method to assess the efficiency of training courses in international comparison. The focus will be on assessing and comparing vocational skills in different countries as well as on using the results when considering the categorization of vocational qualifications in existing classification systems. The aim is to place national vocational training on an international and European footing and to qualify trainees to meet increasing international demands on the labour market. The international comparison of vocational skills is intended to highlight national strengths and weaknesses and to thus contribute towards improving the national vocational training system. The results of the skills assessment would also enable an empirically valid classification based on international classification systems (such as EQR, ISCED).

The challenges involved in developing methods in vocational training are considerable since the tasks are relatively complex. Basically, it is a question of the technical and functional development of tests, taking into account working and business processes in



different occupational fields. Technology-based test procedures provide scope for authentic test surroundings, for example, by integrating multi-media test material such as film and video sequences or through interactive simulations. These have the advantage of being more valid and authentic with regard to the skills area under assessment than traditional test procedures. The aim is to develop and try out test items in various occupational fields. These fields can be defined by occupation-specific and cross-occupational skills. In addition, general skills should also be assessed as determinants.

Even though the introduction of innovative skills assessment processes in vocational training has been quite successful (for example, the pre-tests within the framework of the preliminary work on VET-LSA), the accompanying specialist conferences and workshops show that there is still considerable need for research in this field.

This brings me to the fourth aspect of our activities in this area.

#### **d) Promoting scientific exchanges through the funding of scientific forums and workshops**

In the context of an evidence-based and knowledge-intensive vocational training policy, it is always important to be familiar with findings from research and science and to draw the necessary consequences for education policy. Furthermore, when designing research initiatives it is essential to know what research is needed from both the scientific and the political point of view. Against this background, the BMBF has been encouraging scientific communication for many years now by providing funding for forums and workshops and by initiating feasibility studies to assess proposed research initiatives as to their functionality and practicability.

In this context, the BMBF has increased the number of conferences and small pre-tests on questions of assessing vocational skills. The proposed research initiative will intensify this dialogue still further and extend the role of research in shaping technology-based assessment procedures, in Germany as well as in interested European countries.

## **4. Prospects**

As we all know, **education systems** are not speedboats – on the contrary. Course corrections take time. And a long time often elapses between realizing that something has to change in education structures and actually adopting concrete measures. This frequently involves a difficult counselling and decision-making process. Nevertheless, we are convinced that we can generate the necessary information in the future, both in the field of monitoring and in the area of skills assessment – information which is necessary for shaping a modern, efficient vocational education and training system.

# Learning for Jobs

## The OECD Policy Review of Vocational Education and Training

*Kathrin Hoeckel*

**Abstract:** Learning for Jobs is an international comparative policy review of vocational education and training being carried out by the Organisation for Economic Co-operation and Development (OECD) between 2007 and 2010. The study is designed to help countries make their VET systems more responsive labour market needs. It aims to improve the VET evidence base, identify VET policy options and develop tools to appraise policy reforms. The review focuses on initial VET for young people. Analytical work is carried out based on data from all OECD countries. Individual policy reviews of national VET systems are being conducted in 17 countries. Drawing from a wide range of data and information gathered during field visits, the country reports analyse the strengths and weaknesses of VET systems and provide concrete policy recommendations and suggestions for implementation. An international comparative report identifies examples of good practice and provides VET policy tools.

**Keywords:** initial VET, labour market responsiveness, comparative research, OECD

### **Why is VET a priority at the OECD?**

For OECD member countries, a well-skilled workforce is one of the main supports for prosperity and growth. Some of the skills are coming from the expansion of general education at all levels. But OECD countries also need specific occupational skills – in quickly expanding fields such as in health care, in jobs driven by technology and new skills developments such as information technology, as well as in traditional trades like electricians, vehicle repairs and plumbers.

The collective importance of these skills is high in all OECD economies, although the exact mix of skills and occupations in demand vary across countries and is changing fast. Typically, initial vocational education and training (VET) systems have a big part to play in supplying these skills – and therefore play a critical role in meeting the needs of a well-functioning modern economy.

In past decades, education policy and research – including at the OECD – have focused mainly on compulsory education in schools, tertiary education at universities and increasingly also on early childhood education and care. Vocational education and training has been relatively neglected. However, some countries have realised that they have maybe invested too heavily into tertiary skills and are now facing skills gaps in specific occupational areas. Moreover, despite the heterogeneity of VET systems across OECD countries, common strains

have become apparent, including a lack of well prepared VET teachers and trainers, or insufficient engagement of social partners in the design and provision of VET programmes. Countries are keen to share experience and examples of best practice to learn from each other how to tackle such challenges.

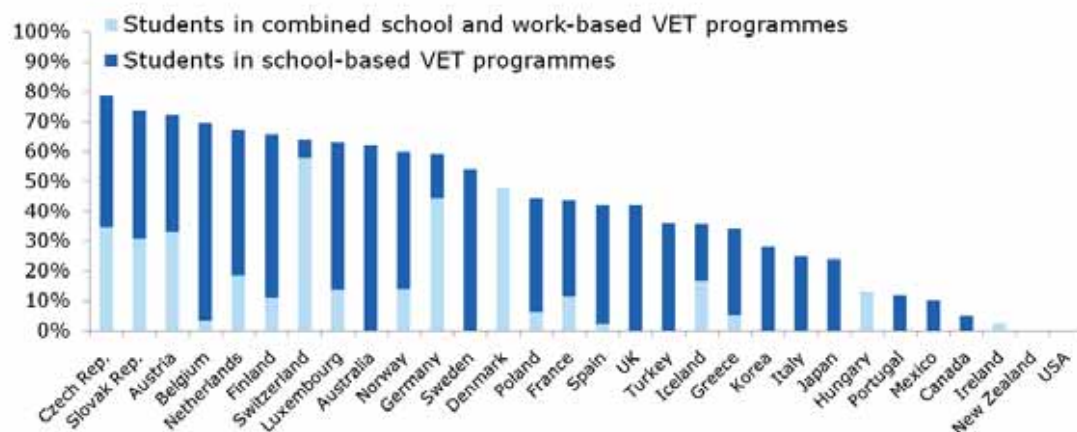
For these reasons the OECD Directorate for Education has been asked by member countries to study the challenges VET systems are facing from a comparative international perspective and to facilitate the exchange of experience and knowledge on international good practice in VET policy between countries. Launched in 2007, *Learning for Jobs*, the OECD policy review of vocational education and training is designed to help countries with the task of improving their VET systems to meet standards of international best practice.

### **Remit and methodology**

But what are the criteria of international best practice and how to assess them? The overarching question chosen for the analysis is how to make VET systems responsive the needs of the labour market. Consequently, labour market performance of VET students from different programmes, both in terms of the smoothness of the initial transition from education to work and the flexibility and success later in working life are at the core of the analysis. Elements of VET systems which have an impact on student labour market performance are determined.

The key challenge in international comparative work on VET is that vocational education and training takes highly diverse forms and is of varying importance across OECD countries (see Figure 1). For example in continental Europe, many young people enter a vocational programme in upper secondary education, sometimes linked to workplace training. In the United States, by contrast, career and technical education (as vocational education and training is described there) is concentrated largely at tertiary level, particularly in the community colleges. Some countries have extensive formal apprenticeship systems, others practically none. The aim of the study can therefore hardly be to establish a strict ranking of countries' VET systems according to a list of well defined indicators. Specific country contexts have to be taken into account.

**Figure 1: Heterogeneity of VET provision at upper secondary level**



Source: OECD (2008), *Education at a Glance 2008, Indicators*, Table C1.1, OECD, Paris

Moreover, VET evidence is not only difficult to compare but also often incomplete. In most countries, data allowing to follow students through their education careers and into working life, necessary to study problems related to drop out, permeability of VET systems and progress routes between different pathways and the precise labour market outcomes of different VET programmes are not available. Few countries have comprehensive surveys of students after they graduate and of employers to assess their satisfaction with and engagement in the system.

To make the scope of the study manageable it focuses on initial vocational education and training for young people before they start their first job. A twofold approach is used to study initial VET systems: individual country reviews looking at the specific national situation and identifying good practice examples; and international comparative work based on data analysis and aiming to identify general trends.

To this end, the OECD is conducting individual policy reviews of vocational education and training in 14 countries: Australia, Austria, Belgium (Flanders), the Czech Republic, Germany, Hungary, Ireland, Korea, Mexico, Norway, Sweden, Switzerland, the United Kingdom (England and Wales), and the United States (South Carolina and Texas). Short reports on Chile and the Peoples Republic of China are also being prepared. These country reviews are based on a broad set of data and information gathered through desk-based analysis and during study visits to the countries and interviews with VET stakeholders including the government, social partners, VET practitioners and experts. The focused reviews analyse the strengths and weaknesses of the respective national VET systems and provide concrete policy recommendations and suggestions for implementation.

As for the second approach, several comparative papers have been produced. An initial mapping exercise in the form of a questionnaire on the structure of VET systems sent to all countries enabled descriptive evidence on how VET systems are organised in countries, how they are funded, to what extent social partners are involved etc. to be gathered. Other working papers prepared as part of the study include reviews of previous VET literature, VET in PISA, an appraisal framework for VET policies, a study of the effect of economic downturns on the provision of apprenticeships and initial workplace training and the relationship of career guidance and VET. The initial draft of the comparative report (Field, *et al.*, 2009) drawing from the evidence was published on the OECD website in October 2009 ([www.oecd.org/edu/learningforjobs](http://www.oecd.org/edu/learningforjobs)). The final comparative report will be published as a book in autumn 2010.

## **Results**

Trends in the supply of, and demand for VET provision are very diverse. At upper secondary level some countries have been expanding VET provision, while in others it is diminishing. At post-secondary and tertiary level the global patterns are even less clear, given the weak frameworks for identifying and measuring VET at these levels. Despite the international diversity of VET systems, some common lessons emerge:

### **VET systems should deliver the right skills mix for the labour market**

VET systems seek to deliver the skills needed by the labour market as a whole. This means providing a mix of VET programmes that reflects both student preferences and employer demand, finding mechanisms to ensure that the number of people trained in different occupations matches labour market needs. Relying purely on student preferences to determine the offer of VET programmes has limitations since employer needs might not be sufficiently addressed. At the same time, it is very difficult to centrally plan provision to meet labour market needs. Forecasting - by location and by occupational sectors - of the exact number of skills needed in a given labour market is often unreliable. Ideally, VET programmes should include an element of workplace training because, apart from the learning benefit, employers' willingness to provide such workplace training reflects potential labour market demand for the skills acquired in the VET programme. Generally, employers and unions should be engaged in curriculum development to ensure that the skills taught correspond to those needed in the modern workplace.

The right balance between student preference and employer needs depends in part on funding. For VET beyond secondary level, costs should be shared between government, employers, and individual students according to the benefits obtained. If students pay full or high fees, they may reasonably expect their preferences to play a dominant role. Conversely,

where employers fund all the training, they will naturally expect to decide what is to be taught. In between these two extremes, there are many models of mixed support for training from government, students and employers. In principle, efficiency requires these models to reflect the mix of benefit obtained from the training.

On a micro level, a good balance between generic and specific skills obtained by each individual is also important. VET graduates need the occupationally specific skills that will allow them to enter skilled specialised jobs without lengthy additional training. The provision of specific skills is also attractive for employers who can make immediate use of VET graduates and encourages them to engage in the provision of training. But students also need generic transferable and basic literacy and numeracy skills to carry them through their working career, including the ability to adapt to fast-changing workplace requirements, to manage their career progression and to engage in further learning. VET systems should provide students with the right mix of all these skills.

### **VET students need adequately prepared teachers and standardised assessment**

Many countries are facing a shortage of teachers and trainers in VET institutions as the current workforce ages. Some teachers and trainers are also handicapped by lack of recent workplace experience because they fail to update their knowledge regularly. In industry, a different problem emerges. Trainers and supervisors of apprentices and trainees in companies often have no specific pedagogical training. If pedagogy is taught to VET teaching and training personnel, it is often general, rather than VET-specific taking into account the additional challenges VET teaching in various settings provides.

The task of VET systems is to deliver sufficient recruitment of teachers and trainers for VET institutions, and ensure this workforce is well-acquainted with the needs of modern industry. To this end they need to:

- Encourage part-time working, promote flexible pathways of recruitment and take advantage of the current economic slowdown to encourage those leaving industry with good practical skills to enter the workforce of VET institutions.
- Provide appropriate pedagogical and other preparation for trainers (including the supervisors) of trainees and apprentices in workplaces.
- Support interchange and partnership between VET institutions and industry.

Assessment of the skills acquired by students and apprentices during their training provides a basic test of quality in teaching and VET provision in schools and in the workplace. There

are demonstrated advantages in evaluating student performance in VET through standardised national assessments. Such national arrangements can improve student performance, enhance the signalling value of qualifications, be more cost-effective than locally organised examinations, facilitate recognition of informal and non-formal learning, and promote flexibility and innovation in learning.

### **Workplace training should be used extensively**

Workplace training for young people has major advantages. It provides a strong learning environment, it can improve transition from school to work by allowing employers and potential employees to get to know each other; it contributes to the output of the training firm, and it links training provision to a direct expression of employer needs.

But workplace training also has limitations. The basic knowledge needed for an occupation and some basic practical skills are better learnt in classroom settings and workshops; practical training involving dangerous or expensive equipment is less risky in a simulated setting; and the slower pace of a classroom or workshop setting can give students the time to develop and refine their skills in a way that would not be possible in a rapid-paced workplace environment. Workplace training also requires employer support. When the economy turns downward, it is understandably hard to convince an employer worried by the immediate survival of the enterprise to take an interest in training. For all these reasons initial VET systems will normally need a balanced mix of workplace and other forms of training.

Apprenticeship, one main model of workplace training, typically combining training in the workplace with education at schools and sometimes in workshop settings can be very effective. It faces two main challenges: the supply of training places, and their quality. Meeting both challenges at the same time is hard since quality requirements can be burdensome for employers. An ideal apprenticeship system will involve high quality training providing good transferable and occupation-specific skills, attractive to a wide range of employers, relevant and appealing to apprentices and an adequate wage. Countries use many types of financial incentives to encourage firms to offer workplace training, including direct subsidies, special tax breaks and training levies.

In addition to ensuring a sufficient quantity of apprenticeship places, countries need to control and evaluate quality of workplace training. One way of doing this is through the provision of a clear contractual framework for apprenticeships. Other measures include standards for workplace training and training plans, questionnaires administered to apprentices to assess their satisfaction with the training as well as a range of voluntary measures and checklists for quality assurance.



## **Policy tools, necessary to make reform happen**

Good tools are needed to make policy reform happen. The development and implementation of policy depends on well-informed people, working with relevant stakeholders through strong institutions. Crucially, VET policy development and implementation requires engagement with the social partners, employers and trade unions. Their involvement helps to ensure that the content of VET - what is taught in VET providers or schools and at the workplace and how exams are designed - is relevant to the labour market. Involvement of social partners ensures that VET systems react flexibly to changes in the labour market and VET qualifications are up-to-date. Employers and unions complement each other in that employers typically favour occupation specific skills and specialised trades whereas the unions ensure that training is sufficiently broad so that students, the future employees are flexible in the labour market and find jobs beyond their initial employment after leaving the VET system. Their engagement typically requires a set of interconnected institutions at national, regional and sectoral levels, with clear responsibilities for different elements in the VET system.

Information supports the link between VET and the labour market. It allows students to see their way through a VET programme into the labour market, employers to understand what competences potential recruits have acquired in a VET programme and policy makers and VET institutions to see whether VET programmes and institutions are getting their graduates into relevant work. Countries should invest in collecting good data on the labour market outcomes of VET, and in providing the capacity to analyse and disseminate that data. There are various ways to improve data on labour market outcomes. Better information might be provided either through graduate destinations surveys administered to those leaving VET programmes around one year after completion, to establish labour market outcomes of different programmes. Or, evidence can be established by tracking individuals through VET into employment thus, establishing rich longitudinal data on returns to investment into different form of education.

Drop-out and career changes due to misinformed choices have high costs. Individuals incur opportunity costs of additional time spend in education; governments typically have to cover the costs of this education and employers lose out in particular if an apprentice drops-out during the first year when an apprenticeship salary is already paid but when the apprentices spend most time with exercises to learn basic practical skills and without contributing much to the production process. While informal sources such as family and friends may provide useful information, high quality professional career guidance, well-supported by labour market intelligence, is indispensable. In many countries, the preparation of career guidance personnel focuses too strongly on psychological counselling and labour market information is neglected. Typically career guidance personnel are trained in the tertiary sector and some-



times have a bias for university education and lack knowledge about the VET sector. Standards for career guidance are essential to ensure quality. The use of external services such as information centres run by the labour market offices can complement career guidance in schools.

## **References**

The initial comparative report by Field, S. Hoeckel, K. Kis, V & Kuczera, M. 2009, Learning for Jobs, OECD, Paris, is available online: <http://www.oecd.org/dataoecd/36/24/43926141.pdf>.

Further information on the project as well as individual country reports and analytical papers on selected VET issues can be retrieved from the project website: [www.oecd.org/edu/learningforjob](http://www.oecd.org/edu/learningforjob)



## **System Analysis & Monitoring**

# Challenges and Chances of a European monitoring instrument in VET

*Ute Hippach-Schneider*

**Abstract:** As Europe continues to grow together in the field of vocational education and training, extensive and joint endeavours will be required in order to achieve transparency. The differences the VET systems exhibit in terms of structure, management and societal relevance mean that it is necessary to have instruments in place which provide information to facilitate cooperation within a spirit of trust and to enable us to learn from one another. This ties in with the activities of the European Reference Network for Vocational Education and Training – or ReferNet for short. ReferNet makes a contribution in this area via its main task focus of drawing up a range of national analytical reports. The present paper will outline the core task of ReferNet and conduct an investigation of the extent to which this task can be deemed to come under the term of “monitoring”. The strengths and weaknesses of this instrument will also be debated and options for future development will be derived.

**Keywords:** monitoring, ReferNet

The European Reference Network for Vocational Education and Training – or ReferNet – was initiated in 2002 to serve as one of the building blocks supporting endeavours to raise transparency within the colourful European VET landscape, improve knowledge of the various vocational education and training systems and thus reduce prejudice (Hanf and Tessaring 2003). The aims of the network were the establishment and ongoing updating of a knowledge and information system for vocational training. The idea behind the decentralised recording of knowledge and coordinated knowledge dissemination was to facilitate cross-border knowledge sharing with interested parties from within the relevant specialist fields (Hippach-Schneider 2007).

## Construction and core task

The network is coordinated by Cedefop, the European Centre for the Development of Vocational Training, which has concluded or will conclude agreements with consortia in the 29 states. These individual consortia have been selected by means of a tender process. The central selection criterion is the composition of the consortia. The aim is for the consortia to serve as a vehicle for ensuring that the far-reaching insight into policy, practice and research in the field of vocational training necessary for the work of the network is in place. This is the only way of providing an expectation that the information will exhibit the required reliability and quality.

The German consortium currently comprises 20 members, including the German Institute for Business Research, the Dortmund Social Science Research Centre, the German Institute for Adult Education, the DEKRA Academy and, since 2008, the National Guidance Forum for Education, Career and Employment (nfb) (see [www.refernet.de](http://www.refernet.de)). BIBB has taken on responsibility for national coordination.

Although this network also enables bilateral contacts to be established and the exchange of information to take place, the network partners cooperate in projects in a number of different combinations. The core tasks of the ReferNet are, however, to prepare content for Cedefop's European Knowledge Management System and to publish and publicise such content at a national and international level in the form of reports via the national coordinators. The latter task has become significantly more prevalent over the course of recent years, a development which has been to the benefit both of the network and to interested parties within the general public in terms of creating greater visibility. All available reports, for example, are included on the national ReferNet websites, which have been in existence for three years, and are also accessible via the Cedefop website. Print versions of the reports are made available at home and abroad at workshops, seminars and at conferences in which BIBB participates.

#### *Collection and analysis of information on national vocational education and training systems*

The core element of ReferNet reporting is the preparation of coherent analytical reports on the respective national VET systems.

These reports may be divided into three groups:

1. the so-called "Country Reports – VET in Europe";
2. the national ReferNet Policy Reports;
3. the national ReferNet Research Reports.

1. Descriptions of national vocational education and training systems have been produced on an annual basis since 2002. These reports feature a standardised content structure agreed upon by Cedefop and the coordinators. This content structure is carried on from year to year, although current circumstances such as the effects of the present economic crisis on vocational training may dictate that chapters or sub-chapters are supplemented, cause the exclusion of other chapters and sub-chapters or lead to a shifting of the main areas of focus within the chapters. The new features of the 2009 Report include Chapter 10, which is of particular use to the extent that it provides a space for the alignment of the various national vocational qualifications and certificates to the ISCED classification system. This assists the reader in evaluating the comparabilities of the qualifications described in the preceding chapters.

#### Table of contents “Country report Germany – VET in Europe” 2009

- General context – framework for the knowledge society
- Policy Development – objectives, frameworks, mechanisms, priorities
- Legislative and institutional framework – provision of learning opportunities
- Initial vocational education and training
- Continuing vocational education and training
- Training VET teachers and trainers
- Matching VET provision with labour market needs
- Guidance and counseling for learning, career and employment
- financing – investment in human resources
- National VET statistics –allocation of programmes

Such a design concept for the reports is not without its problems. Specific national circumstances are accorded consideration, although there is, of course, the inherent danger that the common structure adopted may mean that these circumstances may not be discerned in a manner as prominent as the way they are perceived nationally. This represents a challenge for the authors. On the other hand, the attraction of these uniform stipulations, which certainly also provide scope for flexibility in terms of the manner in which content is presented, is that they facilitate a comparison between individual vocational training aspects in different countries. The basic idea behind the English-language reports, which require updating on an annual basis, is to offer an initial basic insight into the structure, organisational form and general prevailing conditions within the national vocational education and training systems.

**2.** The national ReferNet consortia produced a so-called Policy Report for the first time in 2008. The aim and purpose of this report is to present national developments with regard to the political priorities formulated in the Copenhagen Declaration and in the resolutions adopted at the follow-up conferences in Maastricht (2004) and Helsinki (2006).

The objective of the Copenhagen Declaration of 30 November 2002, to which the Ministers of Education from 31 countries, the European social partners and the European Commission were all signatories, was to initiate a new phase of closer European cooperation in vocational training (the so-called “Copenhagen Process”). A logical extension of this aim is periodically to prepare a European Progress Report, and indeed such a report has been drawn up on three occasions since 2004. This requires information from the member states. In 2008, the expertise and know-how of the ReferNet was used for the first time for this purpose.

The main section of the 2008 Policy Report is divided into the following main thematic areas:

- Improving access to and equity in VET;
- Lifelong learning through VET;
- Improving quality of VET;
- Strengthening the links between VET and the labour market;
- Governance and funding;
- EU tools and geographic mobility within VET;
- Development of VET statistics.
- VET beyond 2010

The introductory chapter to the report is devoted to social economic matters and aspects of relevance to the labour market, the aim being for this to serve as a background for a greater understanding of national policy structure. The concluding chapter provides an outlook of possible educational policy priorities for the time after 2010, the year in which the Lisbon Process expires. Work is currently ongoing on the national Policy Reports for 2010.

#### Structure of the ReferNet Policy Reports 2010

- Socioeconomic challenges for future VET policy development
- Economic crisis – VET policies as recovery measures
- Impact and implications of the joint work on European principles and tools
- Strengthening the links between VET and the labour market
- Addressing equity, social inclusion and active citizenship
- Quality, excellence and attractiveness of VET
- Enhancing creativity and innovation
- Financing VET
- Progress in modernising European VET systems in the Copenhagen process and priorities for future cooperation

**3.** What is occupying the “vocational education and training researchers” of a country? Is there any research of note? Which topics are being investigated and processed? Which relevant issues and material results are emerging? The answers to the above questions enable a distinct insight to be obtained into a state’s VET system by revealing focal points where “construction works” are ongoing.

In 2009, an attempt was undertaken in all countries represented in ReferNet to respond to these questions via four main areas of content focus. “Benefits of VET” and “Mobility and migration” are common topics for all. Due to the fact that there are a large number of current issues of completely differing degrees of relevance within the member states, the national coordinators were given the freedom to decide on two further topics, making their selection from a long list. In overall terms, we may state that only an extremely small number of common guidelines were in place for the ReferNet Research report compared to the other two reports described above. The targeted aim was to ensure the greatest possible leeway in terms of the main areas of focus for the description of national VET research. Germany, by the way, opted for the topics “Effectiveness and quality assurance” and “Transitions”. Interestingly, seven further states selected the topic “Effectiveness and quality assurance”, making it the most frequently selected area of focus. “Transitions” was chosen by four other countries and came 3rd in a list containing 15 proposals in total.

### **Is this monitoring?**

Where should ReferNet and its central “products”, its reports, actually be aligned? Is it a “monitoring” instrument, a “peer review” instrument, or simply a network which makes a contribution towards the transparency of vocational education and training systems in Europe?

The term “monitoring” is used in a variety of different contexts and is, therefore, subject to a range of connotations. What all definitions have in common, however, is that monitoring is a recurring act of observation which mostly takes place at regular intervals rather than being a one-off action. The varying contexts and in particular the different purposes and objectives of this “observation” then lead correspondingly to a range of different approaches.

### **Examples**

- The national educational and VET reports
- The monitoring and assessing of the performance of institutions such as vocational schools or institutes of higher education, often as a consequence of growing autonomy. Alexander speaks of “Accountability versus Autonomy” (Alexander 2000), and increasing cost pressure.
- The so-called “peer reviews”, which may be described as a systematic investigation and examination of the “performance” of a country within a certain policy area, such as the educational situation (Pagani 2002)
- Comparative analyses of several countries on the basis of uniform indicators, such as those conducted by the OECD (Education at a Glance, or the Education Policy Analysis carried out from 1997 to 2004). Pagani und Rürup draw a further distinction



here between data oriented reports or reports based on educational statistics on the one hand and inspection based reports on the other (Pagani 2002; Rürup 2003). In methodological terms, the comparative and monitoring studies in this group form their own subunit for which dedicated data surveys such as PISA are conducted.

Within this context, Radtke adopts a critical approach to the performance culture and to the role of the OECD (Radtke 2003). He particularly bemoans the fact that the picking out of individual indicators leads to the “decontextualisation” of academic research and complains that the “cost-benefit ratio” is “defined in one-dimensional terms in the way it is related to the economic system”.

Educational monitoring is defined in the following way within the context of the development of a design concept for a Swiss vocational training report: *“the systematic and long-term procurement and preparation of information on an educational system and its surrounding environment”* (Swiss Conference of the Directors of Education of the Cantons, EDK; <http://www.edk.ch/dyn/11663.php>; German language resource accessed on 30 October 2009, Swiss Coordination Office for Educational Research (<http://www.skbf-csre.ch/6.0.html>; German language resource accessed on 30 October 2009, (Maritzen 2008).

The following formulation is also used. *“Educational monitoring serves as the basis for educational planning, educational policy decisions, reporting and public discussion”* (ibid).

The inspiration provided by the debate surrounding the results of the TIMSS (1997) and the PISA Study (2000) was not the least of the reasons why the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany adopted an overall educational monitoring strategy in June 2006

([http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2006/2006\\_06\\_02-Bildungsmonitoring.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_06_02-Bildungsmonitoring.pdf); German language resource accessed on 30 October 2009). The purposes of this strategy include a *“...systematic procurement of information on the educational system..”*. The aim is for the educational monitoring system to contain a number of central procedures and instruments such as international school performance investigations and joint reporting processes for the Federal Government and the federal states. The core of these procedures and instruments is a set of parameters or key indicators, and the intention is that they will constitute the basis for political decision-making as well as contributing towards greater transparency. Endeavours are being undertaken to achieve connectivity and harmony with international reporting systems in order to facilitate a comparison with developments in the EU member states and within the OECD. Two reports have been produced thus far, the editions from 2006 and 2008 (<http://www.bildungsbericht.de/>; German language resource accessed on 30 October 2009, see also (Klieme, Avenarius et al. 2006)).

The redesign of the annual German Report on Vocational Education and Training for the 2009 edition makes clear a fundamental differentiation between the gathering of research data and the political evaluation of such data. The report now features two distinct sections. These are a political part, for which the Federal Ministry of Education and Research (BMBF) is responsible, and a non-political part, which contains academic research and data-based information and which falls into the remit of the Federal Institute for Vocational Education and Training. This latter section is defined as “a material supplement to the government report, particularly in statistical terms” (BMBF 2009), whereas the political section provides a “Abstract of central VET policy decisions, developments and measures” (ibid). The statutory task of the report is observation and: “in the event that the securing of a regionally and sectorally balanced supply of training places appears to be in jeopardy, the report should contain proposals for the rectification of such a situation” (§ 86 Vocational Training Act, BBIG). Aside from this, no express formulation exists which states that the purpose and aim of these two reports is to provide information to serve as the basis for future initiatives.

The first two chapters of the data report of the Report on Vocational Education and Training include central indicators for initial and continuing vocational education and training. The plan is for a third chapter to address different main topic focuses each year. The theme for 2009 was “training modules”.

The OECD publication “Education at a Glance” has selected indicators which appear appropriate to the measurement and comparison of the performance of member states in the field of education (for information on the definition of the “indicators” see (Döbert 2008)). According to the OECD, the aim and purpose of the report is for member states “to see themselves in the light of other countries’ performance”

([http://www.oecd.org/document/24/0,3343,en\\_2649\\_39263238\\_43586328\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/24/0,3343,en_2649_39263238_43586328_1_1_1_1,00.html), accessed 23 November 2009).

The present paper does not intend to address the extent to which it is actually possible with these reports to measure performance and what performance is at all, or whether this methodological approach fails to accord consideration to correlations which would be required in order for attempts at explanation or interpretation to be undertaken or for conclusions to be reached with regard to areas where policy action is needed or in respect of quality aspects (such as the quota of academics (Müller 2009) or the unemployment rate (Hall and Schade 2005)). This mentioned “peer review” element shows clearly, the reports create a remarkable “peer pressure” and leave the sector of non-binding, purely descriptive and analytical information. This report has been published on an annual basis since 1992, although both content (chapters and indicators) and the scope of the publication change from issue to issue.

## Examples of forms and aims of monitoring

	National educational and VET reports	Evaluations of national educational institutions with regard to “performance”	Peer reviews	International educational reports
<b>Forms</b>	Systematic and long-term procurement and preparation of information on an educational system and its surrounding environment; Indicators-based	Evaluations on the basis of such indicators as institutional efficiency, consumer satisfaction, job placement and value for resources (Alexander 2000)	Systematic examination and assessment of the performance of a state by other states.	Indicators-based reports from supra-national organisations
<b>Aims</b>	Foundation for national educational planning and public debate/transparency	“Pressuring institutions to become more accountable, more efficient, and more productive in the use of publicly generated resources.”(Alexander 2000); similar (Greany and Kelaghan 1996)  “provide information on programme quality and on programme outcomes.” (Wholey and Hatry 1992)	“Helping the reviewed state improve its policy making, adopt best practices and comply with established standards and principles.” (Paganini 2002)  Using the effect of peer pressure.	“The aim is to make comparative data on resources provision, individual use and the functionality and performance of their education systems available to OECD member states” (German PISA Consortium 2001);  Using effect of peer pressure?

## **Strengths and weaknesses**

If the criteria stated above are applied to the network in overall terms, the main view which emerges is one of a collection of national educational reports which permit a comparison to be drawn on the basis of the uniform structure and the common indicators between states. The primary aims are enhancing transparency and a certain degree of “peer review”.

Differentiations need to be drawn, however, when considering the individual types of report produced by ReferNet.

The objectives of the country reports “VET in Europe” (see above) are the description and analysis of fundamental aspects of the national vocational education and training systems. The reports attempt to use the presentation of data and facts as a vehicle for highlighting correlations and explaining control mechanisms. They are primarily instruments of transparency. As far as the future is concerned, it would be desirable to develop the reports further in thematic terms and improve their quality by structuring them on a broader basis. The inclusion of an additional chapter outlining elements of historical development relevant to the current characteristics of the systems would, for example, be conceivable. Quality generally is always a topic. There are, for example, differences from report to report and from country to country. In order to maintain or even enhance the quality level, a critical review of Cedefop is needed, and there is a particular requirement for a higher degree of self-critical and self-conscious responsibility on the part of the individual national consortia.

The basic principle of establishing respective national consortia for the network is convincing in theory. The aim is to integrate a cross-section of both publicly and privately organised VET institutions into the work programme. This enables both independent academic research and policy expertise to inform the reports. Notwithstanding this, these national consortia in effect take on the character of a network. Although their features vary from state to state, the consortia predominantly consist of loosely organised unions brought together by declarations of cooperative intent which are non-binding and do not provide any specific division of labour. Funding provision from Cedefop is not sufficient for the fulfilment of the annual work programme of the consortium, meaning that remuneration for commitment within a national network can only take place in individual cases. Cedefop expects individual participation from its cooperation partners, and this extends to include a financial commitment. This aspect cannot therefore be considered as an incentive for active participation within the network. Over the course of the past three years, ReferNet Germany has undertaken a targeted attempt to make involvement in capacities such as author more attractive by seeking to make the reports better known and more visible. A number of marketing campaigns have been initiated to this end. Nevertheless, communicating the direct benefit of commitment to the national network remains very much a challenge.

The focuses of the Policy Reports are to present and describe the “progress” of individual, basic areas within the VET systems and to select examples of good practice. By the same token, this naturally means that lack of progress must be stated and that areas where action is required must be defined. The national consortia also devote varying degrees of commitment to the fulfilment of this task. Nevertheless, the Policy Reports are a valuable instrument for making the political developments of the member states visible over the past few years since the beginning of the Copenhagen Process. This is where the true benefit of the reports lies. Their aim is to make it possible to follow current developments in neighbouring countries, find inspiration in examples of good practice or provide an opportunity to identify areas of common ground. All this takes place in the interests of positive “peer pressure”.

Finally to the Research Reports. The activities or main thematic focuses of VET research in the active member states of ReferNet are almost even more heterogeneous than the vocational education and training systems themselves. Assessment of the value of relevant research in terms of its significance for evidence-based policy or as an instrument for the evaluation of initiatives and political programmes differs widely. In some areas, there are no faculties or research institutions prepared to take on responsibility for such tasks. To this extent, these reports provide information on the national situation at two levels: firstly as to which issues are being investigated by research at all and secondly as to the main results and findings which are obtained. This focus and the restriction to four main topics (see above) per state make the project operationable in the first place. The next report is planned for 2011, and the issue arises as to how this report is to be structured. It would definitely make sense to alter the main thematic focuses since the period of time of two years is too short for a continuation of the main thematic topics covered in 2009. The objective of the reports is certainly first and foremost to improve transparency. The value they provide for researchers by offering compact English language information on stakeholders and research activities which could be of relevance to their own work should not be underestimated.

## **Conclusion**

The foundations which are in place are capable of being built on. Over the course of the years, a relatively stable network unique within the field of vocational training in Europe has been established. This network fulfils monitoring tasks which are indispensable to both VET stakeholders and to interested parties within the general public in Europe in terms of improving mutual understanding and achieving joint progress.

Although some details of the design concept, structure and dissemination of the reports are required, the focus needs to remain on continuing to improve quality. All those involved will have to play their part: Cedefop, the national coordinators and the large number of national consortium members which exist on paper.

## References:

- Alexander, F. K. (2000). "The Changing Face of Accountability." *The Journal of Higher Education* 71(4): 411-431.
- BMBF (2009). *Berufsbildungsbericht 2009*.
- Döbert, H. (2008). *Indikatorengestützte Bildungsberichterstattung - eine Einführung. Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*.
- W. Böttcher, W. Bos, H. Döbert and H. G. Holtappels. Münster, New York, München, Berlin, Waxmann: 13-19.
- Greany, V. and T. Kellaghan (1996). *Monitoring the learning outcomes of education systems*. Washington D. C., The World Bank.
- Hall, A. and H.-J. Schade (2005). "Welche Ausbildung schützt besser vor Erwerbslosigkeit? Der erste Blick kann täuschen!" *Berufsbildung in Wissenschaft und Praxis* 34(2): 23-27.
- Hanf, G. and M. Tessaring (2003). "Das Refer-Netzwerk des Cedefop und seine Implementation in Deutschland." *Berufsbildung in Wissenschaft und Praxis* 32(4): 13-15.
- Hippach-Schneider, U. (2007). "Wachsendes Vertrauen durch besseres Wissen: das europäische Netzwerk ReferNet." *Berufsbildung in Wissenschaft und Praxis*(3): 30-32.
- Klieme, E., H. Avenarius, et al. (2006). "Grundkonzeption der Bildungsberichterstattung in Deutschland." *Zeitschrift für Erziehungswissenschaft Beiheft* 6: 129-145.
- Maritzen, N. (2008). *Bildungsmonitoring - Systeminnovation zur Sicherung von Qualitätsstandards. Bildungsmonitoring, Vergleichsstudien und Innovationen*. L. f. S. u. M. Berlin-Brandenburg, K. u. K. Bundesministerium für Unterricht und S. K. d. k. Erziehungsdirektoren, Berliner Wissenschaft-Verlag.
- Müller, N. (2009). "Akademikerausbildung in Deutschland: blinde Flecken bei internationalen OECD-Vergleich." *Berufsbildung in Wissenschaft und Praxis* 38(2): 42-46.
- Pagani, F. (2002). "Peer review as a tool for co-operation and change - an analysis on an OECD working method." *African Security Review* 11(4).
- Radtke, F.-O. (2003). "Die Erziehungswissenschaft der OECD - Aussichten auf die neue Performanz-Kultur." *Erziehungswissenschaft* 14(27): 109-136.
- Rürup, M. (2003) *Ausländische und internationale Bildungsberichte als Orientierung für die nationale Bildungsberichterstattung in Deutschland*. *Trends in Bildung international* 7,
- Wholey, J. S. and H. P. Hatry (1992). "The case of performance monitoring." *Public Administration Review* 52(6): 604-610.

# Reviewing European VET Policy - challenges for comparative analyses

*Dr. Torsten Dunkel*

**Abstract:** This contribution presents work in progress. It describes the process of the policy reporting which includes the analysis of different sources (for instance: country reports, including self-assessment by governments and social partners, statistical data, and findings from studies carried out by Cedefop, for example Cedefop's research report and other comparative studies). It will also reflect upon the contingencies emerging from the diversity of VET and qualifications systems. It will draw upon this specific Cedefop project to discuss methodological issues and challenges in the field of policy reviewing and comparative VET policy analysis alongside the Copenhagen process.

**Keywords:** European VET policy review, Copenhagen process, comparative analysis, evidence-based policy

## Introduction

Cedefop is currently preparing its policy report 2010. This report will take stock of progress towards commonly agreed objectives and priority areas for VET in the EU set up in the Copenhagen declaration and the subsequent Maastricht/Helsinki/Bordeaux communiqués and related Council conclusions<sup>1</sup>. To inform European VET policies, following its mandate (renewed in the Bordeaux Communiqué, 2008), Cedefop analyses progress by countries in implementing agreed European VET policy priorities in the Copenhagen Process and publishes, every two years, a comparative analysis of policy developments (see Cedefop/ Tessaring/Wannan, 2004; Cedefop/Lipinska et al., 2007, Cedefop, 2009). The outcomes of the analysis as summarised in Cedefop's fourth European VET policy report which will be presented at the informal ministerial meeting in Bruges in December 2010. Much of the information for the report comes from stakeholders as the Directors General for Vocational Training (DGVET) and the Social Partners as well as from Cedefop's European network of reference and expertise (ReferNet<sup>2</sup>), which covers all Member States, Norway and Iceland. ReferNet also provides thematic overviews of national VET systems as well as information that feed into many Cedefop projects.

---

<sup>1</sup> See the following documents: The Bordeaux communiqué on Enhanced European cooperation in vocational education and training. Communiqué of the European Ministers for Vocational Education and Training, the European social partners and the European Commission to review the priorities and strategies of the Copenhagen process. Bordeaux (26 November 2008). Council conclusions on the future priorities for enhanced European cooperation in vocational education and training (VET), Doc. 16459/08, Brussels (21 November 2008). Council conclusions on a strategic framework for European cooperation in education and training (ET 2020). Brussels (12 May 2008).

<sup>2</sup> ReferNet is Cedefop's network of reference and expertise on vocational education and training. Each national ReferNet member is backed by a consortium of VET related organisations. ReferNet Members provide Cedefop with information on VET developments in their country.



The reporting period of the 2010 report is 2002 to 2010, covering the whole period since the Copenhagen process was started, and reflecting the state of implementation of the Copenhagen priorities, changes in VET policies and their implications in this period.

With this activity, Cedefop aims to support an evidence-based European VET policy agenda and enhanced European cooperation between countries who share common policy priorities.

### **Reviewing policy: exercise between research and policy analysis**

*General remarks on the relation between academic research and policy analysis*

*Generally, “policy analysis is a client-oriented advice relevant to public decisions and informed by social values” (Weimer/Vining, 2005, 24). VET policy analysis is the multidisciplinary study of VET policy seeking to answer questions about the purpose of VET policy and its objectives (societal and personal) that it is designed to attain, the measures for attaining them and the tools for measuring their effectiveness to inform the policy-maker.*

For our purpose we distinguish between academic research and policy analysis. It is important to keep in mind this difference between academic research in the social sciences and policy analysis as summarised in figure 1. Policy analysis uses methods of the social sciences to examine relationships between variables that reflect VET problems and other variables that can be manipulated by public policy. The social sciences primarily aim at constructing theories to understand society, economy, education and training etc. and testing these theories. Social scientists usually exchange with other social scientists than with VET practitioners and policy makers because the addressee of a researcher is the researcher. They apply rigorous formal methods to test theories and mostly retrospectively analyse – to provide empirical evidence. Policy analysts have a different target group as users: actors in the policy arena and researchers with an interest in the policy analysis carried out. Research problems in policy analysis stem from practical considerations of development of VET policies and their implementations rather than theoretical debates in the discipline. Outcomes are more related to the world of action than the one of ideas. Policy analysis is a mediator between research and policy makers which sometimes requires a balancing act characterised by a number of constraints such as time pressure and the necessity of making policy relevant conclusions under uncertainty based on imperfect information (soft evidence), and the difficulty of translating findings into European and government action.



**Figure 1: Differences between academic research and policy analysis**

<b>Discipline</b>	<b>Objective</b>	<b>End use(r)</b>	<b>Methodological style</b>	<b>Constraints</b>	<b>General weaknesses</b>
Social sciences	Construct theories for understanding society, economy, education	Truth as defined by the discipline	Rigorous methods	Rarely external time constraints	Often irrelevant to decision makers
Policy analysis	Systematic comparison and evaluation of alternatives identified	Actors in the policy arena, decision makers	Synthesis of existing research and theory to predict consequences of alternative policies	Policy agenda-led, often deadline pressure, perhaps mitigated by issues recurrence	Myopia resulting from client orientation and time pressure

(adapted from Weimer/Vining, 2005)

### **Specific remarks on European VET policy analysis**

The Lisbon conclusions defined the open method of coordination (OMC) as a means of spreading knowledge of best practices and achieving greater convergence towards the main EU goals while respecting the breakdown of responsibilities envisaged in various EU treaties. The OMC is a form of cooperation for the Member States, based on a fully decentralised approach relying on variable forms of partnership and designed to help them to progressively develop their own policies. It is based essentially on a set of common objectives, reporting and monitoring of progress towards objectives, including by means of indicators (16 core indicators for monitoring progress towards the Lisbon objectives covering the whole learning continuum from pre-school to adult education, teacher's professional development and investment in education and training) and benchmarks (five benchmarks to guide progress on achieving the objectives set for E/T systems of the EU), peer learning and the development of agenda-setting communications and EU reference tools supporting national reforms (European Commission, 2009a).

Indicators are used to measure progress in relation to the objectives proposed for education systems, while benchmarks are intended to act as reference points, emphasising the additional effort necessary for improving education and training systems. The newly defined "European benchmarks build on the existing ones adopted under the 'Education and Training 2010' work programme. They should be based solely on comparable data and take account of the differing situations in individual Member States. They should not be considered as concrete targets for individual countries to reach by 2020. Rather, Member States are invited to consider, on the basis of national priorities and whilst taking account of changing eco-

nomic circumstances, how and to what extent they can contribute to the collective achievement of the European benchmarks through national actions” (Official Journal, 2009, 119/7). The European Commission has stressed that indicators should not be viewed in terms of measuring progress alone. They should also be seen as a basis for establishing dialogue and exchanges between Member States and as a tool for understanding the reasons for differences in performance, so that some countries can learn from the best practice of others. The use of indicators for exchanging best practices and new policy approaches in the EU is even more relevant that in many Member States are now achieving outstanding performance, while others are facing great challenges in achieving the objectives defined.

The Copenhagen process, which is an application of the OMC, was triggered by the Copenhagen declaration on enhanced cooperation in VET which was signed and adopted by 31 European education ministers and the European Commission in November 2002. With this declaration, these countries committed themselves to increased cooperation in particular on the issues of transparency, recognition of competences and qualifications, and quality assurance (Copenhagen Declaration, 2002). The Copenhagen process provides a framework for cooperation of countries, allowing them to agree on common goals for their national VET policies, learn from one another, and regularly evaluate and assess progress achieved. In biennial follow-ups in Maastricht 2004, Helsinki 2006, Bordeaux 2008 this process was confirmed, consolidated and continued. The strategic framework for European cooperation in education and training ('ET 2020'<sup>3</sup>) follows four main strategic objectives:

1. Making lifelong learning and mobility a reality;
2. Improving the quality and efficiency of education and training;
3. Promoting equity, social cohesion and active citizenship;
4. Enhancing creativity and innovation, including entrepreneurship, at all levels of education and training.

Internationally comparable data on education, training and skills have acquired particular importance with the growing together of European countries and the implementation of common EU policies and strategies. These data are also key to comparative research on education and training that aims at establishing an overview of education and training across countries. Cedefop is considering in its policy analysis the 32 countries participating presently in the Copenhagen process of cooperation.

Political and research interest in such data and the recognition of their key role for European economic and social policies are relatively new. Most developments in internationally compa-

---

<sup>3</sup> Council Conclusions on a strategic framework for European cooperation in education and training ('ET 2020') 2941th Education, Youth and Culture Council meeting Brussels, 12 May 2009

rable education and training statistics started in the 1990s. Therefore, international sources do not always provide the information required or long time series.

Despite the substantial progress made in developing international statistics and indicators, missing or insufficient data and statistics still impede evidence-based policies, research analysis and informed decisions by individuals (e.g. for their educational or occupational choices). Indeed, comparable data on education, training and skills still suffer some drawbacks: a considerable set of key data needed by policy-makers and researchers is missing; many available data have not yet been fully exploited; many limitations exist in comparing data across countries and over time.

Although typologies for comparing national systems like the varieties of capitalism literature (Hall/Soskice, 2001) or that of welfare regimes (Esping-Andersen, 1990) are useful for comparison at system's level, there is a remaining tension field between observation and explanation, which could be used as a creative field for policy learning. Cultural embeddedness, tradition and path-dependency and pressure for modernisation are underlying factors not to be neglected when applying streamlined and standardised methods to reform VET systems.

### **Preparing the VET policy report 2010**

The Cedefop policy report follows informs about and analyses the performance and VET progress in the European Member States agreed in the Copenhagen process, and in cooperation with the European Training Foundation (ETF) monitors the candidate countries with slightly adapted questions to their specific needs. The preliminary findings of the analysis will feed into various related policy documents such as the work of the European Commission and the Bruges working group preparing the 2010 communiqué on cooperation in VET beyond 2010. The report also contributes to policy making activities at national levels by identifying good policy practice, and at European level as it is in line with the EU 2020 strategy.<sup>4</sup>

Policy analysis typically focuses on a single policy – here VET policy – and often follows the 'evaluation model' which seeks to assess the impacts actually produced by some intervention. Good policy analysis clearly documents the essential elements of the policy, which leads us to the following questions:

1. What are the essential elements – the key themes – of VET policy from Cedefop's and stakeholders' perspectives in the context of the Copenhagen process?
2. How is progress of the Copenhagen process as a whole measured?

Before answering these questions some remarks on defining the themes in which the progress should be monitored and the whole preparatory process as well as its further planning.

---

<sup>4</sup> EU 2020 is being designed as the successor to the current Lisbon Strategy, which has been the EU's reform strategy for the last decade. (European Commission, 2009b)

### *Defining the key themes for VET policy review in the Copenhagen process*

The preparation of the 2010 Policy report started in mid 2009 with identifying the essential elements of VET policy in form of key themes. 0-versions of the questionnaire for the Directors General for Vocational Training (DGVt) and ReferNet template were discussed at the meeting of the Enlarged Coordination Group Policy reporting and ReferNet Core group meeting in June 2009. A cognitive test for the questionnaire was performed.

For the first time, a draft questionnaire was prepared for social partners and their involvement in the process<sup>5</sup>. All draft questionnaires were finalised in July 2009 and sent for comments.

The themes and the roadmap (see figure 2) for the 2010 policy report were presented and discussed at the Commission's Bruges Working Group meeting and at the spring DGVt and Advisory Committee for Vocational Training (ACVT) meetings. The elaboration procedure is shown hereafter using the example of the DGVt questionnaire.

The questionnaire for the DGVts was prepared by Cedefop on the basis of:

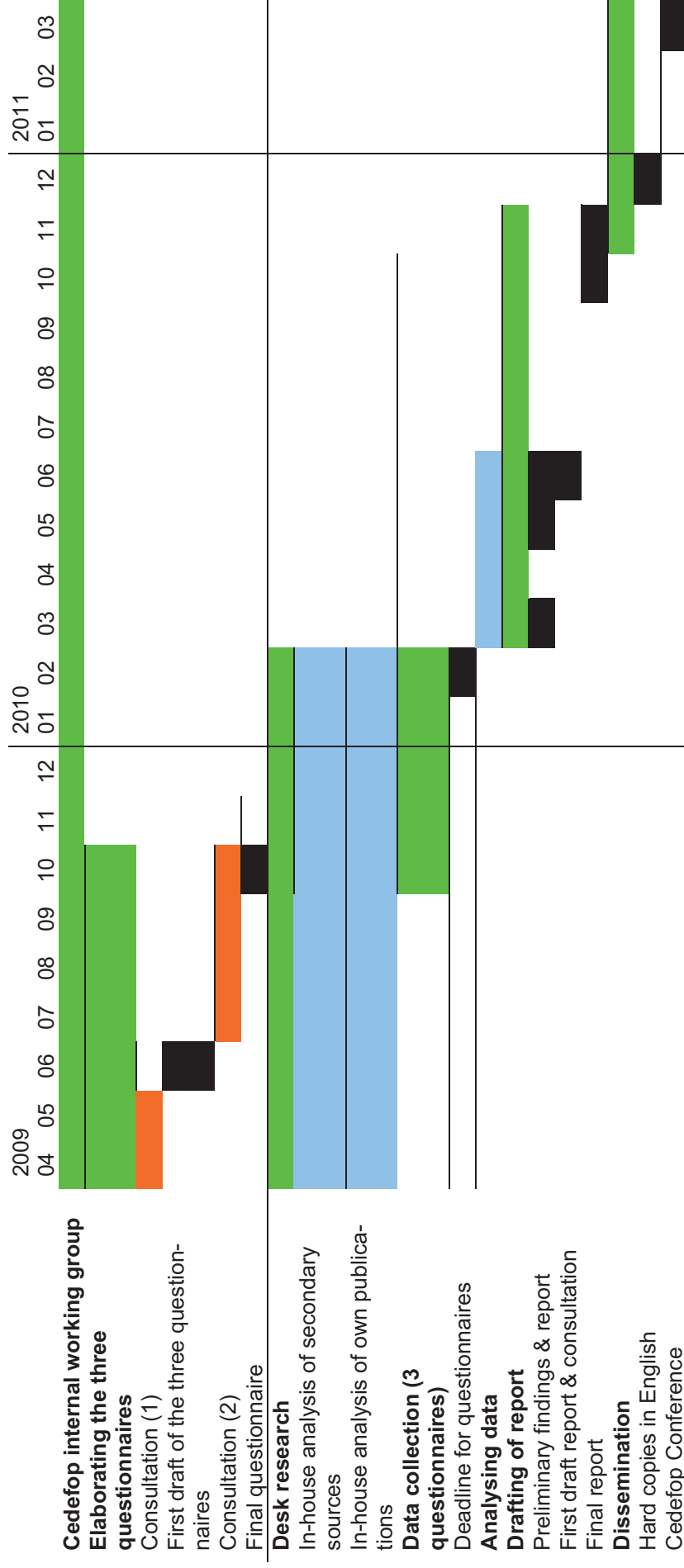
- themes discussed with the Belgian Presidency, European Commission and DGVts at their meeting on 18 and 19 May 2009;
- written comments on the first draft DGVt questionnaire sent out for comments on 22 July 2009;
- comments from the European Commission (DG Education and Culture);
- comments from the Enlarged Coordination Group Policy Reporting (meeting on 3 June 2009);
- in-house expertise;
- based on feedback on initial outline of themes by stakeholders participating in the working group of the Bruges review.

After revision the final versions of all three questionnaires were sent out on 20 October 2009. The Policy Analysis Team is working on the 2010 Policy report with a preliminary analysis of statistical data, studies and other evidence already available.

---

<sup>5</sup> A separate questionnaire targeted at the social partners will additionally collect information on their evaluation of the progress made and, have them reflect their role, involvement and contribution to the Copenhagen process. This questionnaire will for reasons of limited space not further be dealt with in this short paper.

**Figure 2: Roadmap and planning of the policy report**



Elaboration and revision of questionnaires

Consultation (1) Feedback from Communiqué working group, Enlarged coordination group policy reporting, DGVT meeting,

Consultation (2) Cedefop Governing Board, ACVT, DGVT, ReferNet, Enlarged coordination group policy reporting.

## **Back to the initial questions:**

### *Ad (1) Key themes of the policy report*

The report covers the whole Copenhagen process since 2002, with a vision beyond 2010. All three questionnaires are centred on key themes as building blocks for the overall analysis. These key themes are in line with Copenhagen process priorities and reflect those of the Belgian Presidency. They are equally consistent with the new strategic framework for European cooperation in education and training (ET 2020).

The nine themes comprise:

1. Progress in modernising European VET systems in the Copenhagen process and priorities for future cooperation;
2. Emerging challenges and priorities for future VET policy development;
3. Economic crisis - VET policies as recovery measures;
4. Impact and implications of joint work on European principles and tools;
5. Strengthening links between VET and the labour market;
6. Addressing equity, social inclusion and active citizenship;
7. Quality, excellence and attractiveness of VET;
8. Enhancing creativity and innovation;
9. Financing VET.

These themes will be sub-divided into some more detailed questions as shown in the example below (see figures 3 and 4).

### *Ad (2) Mixed method and focussed comparison to analyse progress of the Copenhagen process*

A mixed method approach with a variety of sources such as replies to three questionnaires adapted to the target groups (DGVT, ReferNet, and social partners) will be applied. While the DGVT questionnaire covers more the strategic level of VET policy, the ReferNet one inquires more the operational level of policy measures. This is complemented by the social partners' assessment of progress and their involvement in the process. Although the levels of information to be gathered are different it might arise some overlaps. Redundancy is here used as a safety mechanism. Multiple studies using different data sources addressing the same question allowing policy analysts to triangulate each of the imperfect studies to support decision-making.

Replies should address VET comprehensively, covering initial and continuing vocational training, including training provisions for the unemployed in active labour market policies and vocationally-oriented programmes in higher education. If relevant, distinction between initial and continuing VET should be made.

Additional information from other networks, clusters, and working groups will also be used for the Cedefop analysis. For example, this relates to activities performed or organised within the Education and training 2010 work programme, such as the joint report on its implementation and thematic clusters (recognition of learning outcomes, teachers and trainers, key competences and curriculum development, etc.), Cedefop's Skillsnet network and other (European) working groups (on quality assurance, guidance, credit transfer, etc.). Analysis will also draw on findings of studies carried out by Cedefop such as the research report, skill needs analyses, NQF and ECVET mappings etc. as well as statistical data will be used to be fed into the analysis also taking into account progress reports by the European Commission mostly based on quantitative benchmarks, or other European and international institutions, as well as relevant statistical data from the European statistical system or other international organisations.

The key themes of the questionnaires can be understood as building blocks for the analysis. To this end a focussed comparison is applied to identify measures resp. obstacles that facilitate resp. hinder reforms as opposed, which basically ask the same questions across the nine pre-defined themes to discern similarities among them that suggest generalisations. Although this is not characterised by high level of formal verification the advantage lies in taking into account of institutional and political complexities and explore certain (causal) relationships among variables. Thus, it allows some kind of process tracing. It also includes variables that can be extremely difficult to quantify. This exercise is completed by proxies out of statistical databases.

Further this is to be more than a state-of-play; it intends to reflect interlinkages and impact. It also addresses the potential impact of the economic crisis and measures to prepare for recovery. Analysis and report drafting will be done by chapter teams horizontally by theme and vertically by country.

### **The survey instruments used for direct data collection**

The purpose of this DGVT questionnaire is to gather information on how countries perceive and assess progress in modernising their VET systems, also looking at developments and implementation of VET policies beyond 2010 and taking into account the European level. The questionnaire builds on information DGVTs provided in previous years. When assessing progress, DGVTs are asked to provide explanations, qualitative and quantitative data or examples, where possible, to substantiate the self-assessment.

DGVTs' replies will be complemented by input from ReferNet which will provide more detailed information on VET developments in their countries about existing strategies and initiatives and thus, support the governments' assessment.

Standard questions under the same heading will be asked in the questionnaires except for the crisis theme. Additional questions are asked to governments and social partners, e.g. to assess progress made, or involvement into the process respectively.



The following figures illustrate an example of the theme 'Strengthening the links between VET and the labour market' with its explanatory introduction and its related subquestions of the DGVET and the ReferNet questionnaire.

**Figure 3: Sample theme (DGVET) 'Strengthening the links between VET and the labour market'**

### **Introduction**

Europe is facing several socioeconomic challenges which accelerate the pace of change in the labour market and skills requirements. To make sure that VET responds adequately, a key priority of European cooperation in VET is to forge better links between VET and the labour market (Bordeaux communiqué). It is essential to involve all labour market actors in VET development, including promoting workplace learning, and encourage closer cooperation.

Partnerships between education and training institutions and employers have a particular role to play in improving learners' employability, developing their entrepreneurial potential and make them more familiar with the working world. Greater awareness of trends in the labour market can help to make learning more responsive to future needs and increase students' motivation by providing a clear context for learning. Cooperation with education and training institutions can also help to update skills and ensure professional development of staff <sup>6</sup>.

In this context, it is necessary to develop systems for early identification and anticipation of skill needs and mismatches to ensure that the skill needs identified are incorporated in VET on time, as advocated in the relaunched Lisbon strategy and the 'New skills for new jobs' <sup>7</sup> initiative.

### **1 Identifying and anticipating skill needs**

- a) Please assess progress in your country in developing systems for early identification and anticipation of skill needs and mismatches (forecasts, employers' surveys, regional and sectoral councils or forums, etc.).
- b) What are the main obstacles?
- c) Which initiatives are you planning?

### **2 Integrating changing skill needs into VET**

- a) Please assess progress in your country in incorporating results of identifying and anticipating skill needs in VET provision and guidance services.
- b) What are the main obstacles?
- c) Which initiatives are you planning?

<sup>6</sup> Conclusions of the Council and Representatives of the Governments of Member States, meeting within the Council, of 12 May 2009 on Enhancing partnerships between education and training institutions and social partners, in particular employers, in the context of lifelong learning. 2941. Education, youth and Culture Council meeting Brussels (12 May 2009).

<sup>7</sup> Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. 'New skills for new jobs. Anticipating and matching labour market and skills needs'. Brussels, European Commission, 2008 (COM (2008) 868 final).



### **3 Involving labour market actors in VET**

- a) Please assess progress in your country in involving labour market actors (social partners, sectors, companies) in VET in terms of system development, governance and management.
- b) What are the main obstacles?
- c) Which initiatives are you planning?

### **4 Promoting workplace learning**

- a) Please assess progress in promoting workplace learning in your country.
- b) What are the main obstacles?
- c) Which initiatives are you planning?

(Cedefop, DGVET questionnaire, 2009)

The following figure illustrates an example of the theme ‘Strengthening the links between VET and the labour market’ and its related subquestions. The same explanatory introduction as in the DGVET questionnaire precedes the questions.

**Figure 4: Sample theme (ReferNet) ‘Strengthening links between VET and the labour market’**

#### **Introduction**

Europe is facing several socioeconomic challenges which accelerate the pace of change in the labour market and skills requirements. To make sure that VET responds adequately, a key priority of European cooperation in VET is to forge better links between VET and the labour market (Bordeaux communiqué). It is essential to involve all labour market actors in VET development, including promoting workplace learning, and encourage closer cooperation.

Partnerships between education and training institutions and employers have a particular role to play in improving learners’ employability, developing their entrepreneurial potential and make them more familiar with the working world. Greater awareness of trends in the labour market can help to make learning more responsive to future needs and increase students’ motivation by providing a clear context for learning. Cooperation with education and training institutions can also help to update skills and ensure professional development of staff <sup>8</sup>.

In this context, it is necessary to develop systems for early identification and anticipation of skill needs and mismatches to ensure that the skill needs identified are incorporated in VET on time, as advocated in the relaunched Lisbon strategy and the ‘New skills for new jobs’ <sup>9</sup> initiative.

<sup>8</sup> Conclusions of the Council and Representatives of the Governments of Member States, meeting within the Council, of 12 May 2009 on Enhancing partnerships between education and training institutions and social partners, in particular employers, in the context of lifelong learning. 2941. Education, youth and Culture Council meeting Brussels (12 May 2009).

<sup>9</sup> Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. ‘New skills for new jobs. Anticipating and matching labour market and skills needs’. Brussels, European Commission, 2008 (COM (2008) 868 final).

### **1 Identifying and anticipating skill needs**

- ⇒ Describe policy progress in your country in identifying and anticipating skill needs.
- ⇒ If you apply systems of early identification and anticipation of skill needs and mismatch (e.g. forecasts, projections, employers' surveys, regional and sectoral councils, fora, etc.) please state them.
- ⇒ Discuss the strengths and weaknesses of these approaches.
- ⇒ Please provide 1-2 examples of initiatives.

### **2 Integrating skill needs of the labour market into VET provision**

- ⇒ What were the main approaches in your country to ensure that the results of identification and anticipation of skill needs are incorporated into VET provision and guidance?
- ⇒ Discuss the strengths and weaknesses of the approaches.
- ⇒ Please provide 1-2 examples of initiatives.

### **3 Involving labour market actors in VET**

- ⇒ Describe policy progress in your country to involve labour market actors (social partners, sectors, companies) in VET system development, governance and management (such as formulating VET policies, designing and assessing VET programmes, developing and validating qualifications and standards, guidance and counselling, quality assurance, financing VET)?
- ⇒ Discuss the strengths and weaknesses.
- ⇒ Please provide 1-2 examples of initiatives.

### **4 Promoting workplace learning**

- ⇒ Describe policy progress in your country focused on promoting workplace learning. Please discuss, for example, promoting cooperation and partnerships between educational institutions and the world of work, apprenticeship (including at higher level) and similar alternance training models, continuing vocational training and adult training at the workplace (in particular SMEs).
- ⇒ Discuss the strengths and weaknesses.
- ⇒ What are the opportunities for individuals in your country to have their knowledge, skills and competences acquired through learning at the workplace validated?
- ⇒ Please provide 1-2 examples of initiatives.

(Cedefop, ReferNet questionnaire, 2009)

As for the subquestion on providing examples of policy initiatives the following template is to be filled in to get more detailed information at the operational level.

**Figure 5: Template of ReferNet questionnaire**

<p>1. Title of the policy/ measure (indicate also the year when it was introduced/adopted/implemented):</p>
<p>(a) Rationale (why was the policy/measure introduced? Explain (inter)national context, reason).</p> <p>(b) Objectives/measurable targets (qualitative or quantitative objectives).</p> <p>(c) Target groups (high achievers, young people with learning difficulties and/or disabilities, people with migrant background and ethnic minorities, older workers, people with low skill levels, the long-term unemployed); or educational level(s)/ educational sector(s) (IVET, CVET; VET in upper secondary, post-secondary and tertiary level, etc.).</p> <p>(d) Status of implementation, including statistical data, if available (in the planning stage, pilot project, a measure implemented all over the country/mainstreamed, etc.).</p>
<p>2. Policy/measure operation and delivery:</p> <p>(a) Level of operation (national, regional, local, or sectoral level).</p> <p>(b) Is it an isolated policy/measure or part of a larger (education/ training/ employment/ social) policy approach?</p> <p>(c) Key actors involved (main institutions involved in adopting and carrying out the measure).</p>
<p>3. Evaluation:</p> <p>(a) Assessment of effectiveness, efficiency or impact (including results of evaluation studies, if available).</p> <p>(b) Indicators of success (e.g. high take-up, employment success, low deadweight, etc.).</p> <p>(c) Integration of outcomes into (a) larger (national, sectoral, regional, local) policy/ies.</p>
<p>4. Conclusions:</p> <p>(a) Obstacles encountered (what did not work and why?); measures (planned) to overcome them.</p> <p>(b) What issues remain to be addressed?</p>
<p>5. Source, legend</p>

The purpose of this template is to collect more detailed information on progress in the Copenhagen process of VET since 2002, to identify good initiatives, and gather visions for the future in a national VET policy report. The information provided should be validated by the national ReferNet representative.

### **Results and perspectives for further development**

For the described approach as such methodological concerns of this policy review exercise range from meta-analysis and system's review, to the debate on qualitative reviews vs. quantitative and experimental studies. Despite the disagreements among the proponents to such secondary analysis, there is a widespread consensus that bringing together evidence from across many evaluations is an important guarantor of the reliability of findings (Cedefop/Stern, 2005).

As for potentials and challenges for the use for international comparisons it shall be further researched into evidence-based VET policy by examining important questions about the extent to which policy can be derived from research, how usable knowledge for VET decision making can be created, and the type of evidence that should inform policy. It may be also necessary to critique some of the narrower conceptions of the evidence that might inform policy, and then examine the gaps between what can be revealed by research and the political requirements of policy, and the claims of some of the more traditional educational research traditions to inform policy. This also requires understanding of the policy process and how evidence is disseminated and incorporated into practice. The overall message is the need for a more subtle understanding of the ways in which different forms of enquiry may inform policy and practice, and the recognition and use of insights provided by a variety of the vocational educational research traditions available, and thus provide a multi-faceted policy review of European VET.

In the newly emerging global economic order governments and policy makers are keen to seek ideas from other countries and recognise the importance of looking comparatively. This expansion of interest in comparative VET policy analysis brings new challenges for the discipline: research may be undertaken by non-specialists (by consultants and politicians or educationists from quite different backgrounds); the short lifespan of democratically elected governments may result in 'quick-fix' solutions; statistics and data may be decontextualised. Added to these challenges there is the worldwide proliferation of education providers outside state control and the transformation of teaching and learning brought about by the new information and communication technology. In the near future it might be necessary to rethink the role of comparative VET policy analysis in the light of these changing circumstances and look at the new opportunities they bring.

## References

Cedefop, Tessaring, M., Wannan, J. 2004, Vocational education and training. key to the future Lisbon-Copenhagen-Maastricht: mobilising for 2010, Publications Office: Luxembourg.

Cedefop, Stern, E. 2005, The foundations of evaluation and impact research. In: Cedefop, Descy, P./Tessaring, M. (eds.). The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report, vol.3, pp. 9-50, Luxembourg: Publications Office. (Cedefop Reference series, 58)

Cedefop, Lipinska, P. et al. 2007, Zooming in on 2010. Reassessing vocational education and training. Publications Office: Luxembourg.

Cedefop. 2009, Consolidation, continuity and change. Towards a European era of vocational education and training. Publications Office: Luxembourg (Cedefop Reference series, 73)

Copenhagen Declaration. 2002, Declaration of the European Ministers of Vocational Training and Education and the European Commission on Enhanced European cooperation in Vocational Training and Education. Copenhagen, 29-30 November.

Communiqué of the European Ministers for Vocational Education and Training, the European social partners and the European Commission to review the priorities and strategies of the Copenhagen process. Bordeaux (26 November 2008).

Conclusions of the Council and Representatives of the Governments of Member States, meeting within the Council, of 12 May 2009 on Enhancing partnerships between education and training institutions and social partners, in particular employers, in the context of lifelong learning. 2941. Education, youth and Culture Council meeting Brussels (12 May 2009).

Council conclusions on the future priorities for enhanced European cooperation in vocational education and training (VET), Doc. 16459/08, Brussels (21 November 2008).

Council conclusions on a strategic framework for European cooperation in education and training (ET 2020). Brussels (12 May 2008).

Esping-Andersen, G. 1990, The three worlds of welfare capitalism, Cambridge.

European Commission, 2008, Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. 'New skills for new jobs. Anticipating and matching labour market and skills needs'. Brussels, European Commission, 2008 (COM (2008) 868 final).

European Commission, 2009a, Progress towards the Lisbon objectives in education and training. Indicators and benchmarks 2009 (staff working document), Brussels, 23.11.2009 SEC(2009) 1616 final.

European Commission, 2009b, Commission Working Document Consultation on the Future "EU 2020" Strategy. Brussels, 24.11.2009 COM(2009)647 final.

Hall, P.A. & Soskice, D. (Eds.) 2001, *Varieties of capitalism. The Institutional Foundations of Comparative Advantage*, New York: OUP.

Official Journal of the European Union. 2009, Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ('ET 2020') (2009/C 119/02) 28.5.2009, 119/2-119/10.

Weimer, D.L. & Vining, A. R. 2005, *Policy Analysis: Concepts and Practice*, 4<sup>th</sup> edition, Upper Saddle River: Prentice Hall.

# Results Measurement and Programme-based Approaches (PBA)

*Stefan Thomas*

**Abstract:** Since the adoption of the United Nation's Millennium Development Goals (MDGs) in the year 2000, results measurement appears to have taken on particular importance in the area of development cooperation. The Paris Declaration on Aid Effectiveness (2005) proposed five principles to make development more efficient: ownership, alignment, harmonisation, managing for development results and accountability. These principles affect various stakeholders, among them practitioners in the field, staff at donor headquarters, implementing organisations, governments and partners. In addition, development goals in the field of Technical and Vocational Education and Training (TVET) are often diverse, increasingly focusing on systemic change.

**Keywords:** Results Measurement, Programme-based Approaches (PBA), International Development Co-operation

## 1 Introduction

Against the backdrop of an increasingly varied donor landscape and programme-based approaches, the focus is increasingly on the contribution made by a development intervention or programme to changes in the vocational training sector or system. In order to aggregate results across programmes or interventions, there is a need for joint standards. One possible solution are so-called universal indicators that are integrated into the results chains of programmes or interventions. Together with minimum requirements for results chains, the measurement of indicators and the attribution of results, universal indicators can help to achieve a joint understanding of quality as regards the assessment of programme and intervention results. At the same time, it can be presumed that the adherence to minimum standards when measuring results will regularly provide useful information for programme management.

This paper starts by examining the political background and emergence of programme-based approaches. It explains the various PBA instruments and the importance of PBAs in the TVET sector. It goes on to outline the challenges in terms of results measurement and presents possible universal indicators for cooperation in the TVET sector, and minimum requirements for a results measurement standard.

### 1.1 Background

Several international development and research organisations have observed that certain deficits in development cooperation reduce the effectiveness of the work conducted. This in turn leads to development results that are poorer than anticipated by both developing and donor countries. The three main deficits identified were as follows: (1) Governments in developing countries often lack ownership of their own development strategies and development projects. (2) Implemented pro-

jects are often weakly aligned with the developing countries' own strategies and programmes. (3) Donor countries/organisations do not coordinate activities with each other as intensively as they could do.

In this context, the Paris Declaration on Aid Effectiveness (PD) in 2005, which was signed by numerous countries from both OECD and developing countries, named five principles on which development cooperation should be based:

- (1) **Ownership:** Developing countries set their own strategies, improve their institutions and tackle corruption.
- (2) **Alignment:** Donor countries align themselves with these objectives and use local systems.
- (3) **Harmonisation:** Donor countries coordinate action, simplify procedures and share information to avoid duplication.
- (4) **Results:** Developing countries and donors shift focus to development results and results are measured.
- (5) **Mutual accountability:** Donors and partners are accountable for development results.
- (6) This aid effectiveness process was continued and accelerated by the Accra Agenda for Action (AAA) in 2008, which followed up on the results of Paris 2005.<sup>10</sup>

## 1.2 Programme-based Approaches (PBA)

Programme-based Approaches (PBA) (French: '*Approches fondées sur des programmes*', '*Approches programmes*' for short) can be seen as the strategic mechanism to translate the Paris Declaration into practice. PBA are defined as 'a way of engaging in development co-operation based on the principle of coordinated support for a locally owned programme of development, such as a national poverty reduction strategy, a sector programme, a thematic programme or a programme of a specific organisation.' More specifically, PBAs share the following features that must all be met in order for a project to count as a PBA:

- (1) Leadership (ownership) by the host country or organisation
- (2) Single comprehensive programme and budget framework
- (3) Formalised process for donor coordination and harmonisation of donor procedures for reporting, budgeting, financial management and procurement
- (4) Efforts to increase the use of local systems for programme design and implementation, financial management, monitoring and evaluation.<sup>11</sup>

---

<sup>10</sup> For more information on the Paris Declaration and the Accra Agenda for Action, go to: [http://www.oecd.org/document/18/0,2340,en\\_2649\\_3236398\\_35401554\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/18/0,2340,en_2649_3236398_35401554_1_1_1_1,00.html).

<sup>11</sup> OECD (2006): Harmonising Donor Practices for Effective Aid Delivery Volume 2, DAC Guidelines and Reference Series, <http://www.oecd.org/dataoecd/53/7/34583142.pdf>, p. 37



While PBA is the overall strategic mechanism, the actual implementation of PBA is realised via different instruments (see Chart 1).



Chart 1: Programme-based Approaches and their instruments

**Budget support** (French: appui budgétaire) is a direct transfer to a country's budget supporting a poverty reduction/development strategy or a poverty-related sector strategy (e.g. in the sectors of education or health). One of the main features is the political dialogue between governments of developing and donor countries. Amongst other issues, they discuss and formulate joint policy goals (policy matrix, targets, etc.) and agree upon a Memorandum of Understanding (MoU). As a precondition for disbursement of funds, certain indicators that were agreed upon in the policy dialogue have to be met. Joint reviews and joint evaluations (e.g. using PEFA) are used to ensure that goals are met and are again preconditions for following budget support tranches. There is no earmarking of funds for general budget support. When providing sector budget support, however, transferred funds are earmarked for a specific sector.

A **basket fund** (French: fonds commun) is in the majority of cases a special account into which donors transfer a certain amount of funds. It is used to support concrete activities in a poverty-related sector (earmarking). The fund can be managed by the partner government, a mandated donor (organisation) or an external organisation that is not supporting the fund financially. Decisions about activities are mostly made by a coordination body consisting of representatives of the partner government and participating (financing) donors. One special case of basket funding is the pooling of technical assistance (known as TA pooling). Here, TA resources of various donor and implementing organisations are pooled to support certain TA activities (e.g. studies, training, consultancy services). These resources can be financial, and/or in kind (human and/or material resources).

**Coordinated projects** within the framework of PBA are part of and support an overarching, coordinated partner country programme. They are also harmonised clearly in terms of strategy and content with national programmes and with other donors' strategies. The projects may be (1) financial cooperation projects (parallel financing and accompanying measures) or (2) technical cooperation projects (advisory services measures and financial contributions). As with the other two instruments, budget support and basket funding, projects within the framework of PBA have to adhere to all four 'PBA criteria'.

## **2 PBAs and the TVET sector**

PBAs can be implemented by means of all three of the above-named instruments in the TVET sector. This offers a wide range of action, particularly in the low income countries, for promoting capacity development together with partners. The focus is on measures that support partners in developing strategies, steering and implementing strategies and programmes, and monitoring and evaluating results. At the same time, interventions in the field of initial vocational training and in some parts of the active labour market (teacher training, school infrastructure, workshop equipment, employment promotion programmes) require considerable amounts of funding and presuppose that the state plays an active role. Their implementation within the scope of PBA is therefore particularly expedient.

Within German development cooperation, the TVET sector is dominated by coordinated projects in which technical and financial cooperation measures are interlinked. PBAs are of special relevance in low income countries with high donor dependency and a recognisable will to reform, but also in threshold countries. Often, TVET components are embedded in employment promotion and economic development programmes in these states. National employment strategies in the partner countries increasingly constitute the framework for these employment promotion and economic development programmes. The regional priority area of the PBAs supported by the German Federal Ministry for Economic Cooperation and Development (BMZ) is sub-Saharan Africa. The TVET components of these programmes, for example in Ethiopia or South Africa, cover a wide variety of intervention areas, including TVET policy and system development; standardisation, testing and certification; training and upgrading of technical teachers, instructors and learning facilitators; support of public and non-public training providers; promotion of cooperative and in-company training; promotion of training in rural areas and the informal sector.

In future, an increased demand for complex capacity development programmes in the TVET sector is to be expected from partner governments. The importance of education and training for poverty reduction has prompted many donors, especially in sub-Saharan Africa, to pursue sectoral approaches. In 21 countries in Africa alone, large-scale donor investments are being made to expand the basic education system of the partner countries via the Education for All Fast Track initiative. In many of these Education for All countries, the first large cohorts of school-leavers are now ready to

make the transition from school to work or further training. As a rule, the states concerned are inadequately prepared for this situation, which means that the theme of post-primary education and training (PPET) is becoming increasingly important for donors.

### 3 PBAs in the TVET sector and results measurement

Programme-based approaches therefore alter the project landscape in development cooperation. The coordinated cooperation of various donors in a sector enables greater and more complex projects by means of which highly aggregated results as defined by the Millennium Development Goals can be more realistically achieved.

However, different donors, and the partners involved in each case, frequently pursue different objectives, and there are often diverse expectations of and approaches to results measurement. This concerns the selection of indicators, the requirements to be met by indicators, their measurement and attribution, i.e. the causal attribution of results to concrete development interventions. The different approaches to results measurement may lead to divergent or contradictory results, as well as generating high costs. In addition, partners are often overtaxed by the various demands of individual donors as regards results measurement, since they lack the relevant staff and financial resources.

To enable better data comparability, avoid redundancy and raise the effectiveness of programme-based approaches, it would appear expedient to standardise the assessment measures. This may also facilitate learning across institutions, which in turn would improve the effectiveness of measures in line with the Paris Declaration.

Jointly agreed minimum requirements for results chains, the measurement of indicators and results attribution could be the first step towards standardisation. Beyond this, so-called universal indicators (jointly agreed indicators that measure several dimensions of project success) may be helpful.<sup>12</sup>

In our opinion, the following universal indicators in the TVET sector can make an important contribution to substantiating project results at aggregated level:

1. **Employment ratio:** How many people in the target group have found employment (in %)?
2. **Income:** Average increase in % after the measure?
3. **Scale: Education and advisory staff:** How many intermediaries are reached by the project/programme activities, e.g. teachers, trainers, career advisors (if appropriate in a specific region or sector)?

---

<sup>12</sup> Initial approaches to a Results Measurement Standard have been developed for the field of Private Sector Development by the Donor Committee for Enterprise Development (DCED). Cf. Tanburn, Jim: Measuring and Reporting Results, The 2008 Reader on Private Sector Development, International Training Centre of the International Labour Organization, 2008

4. **Trainees:** How many people in the target group are reached by the project/programme activities, e.g. people who benefit from the measures directly (e.g. in pilot classes) or indirectly (e.g. via new curricula, examination regulations etc.)?
5. **Institutions:** How many institutions are reached, e.g. schools, examination centres, training centres, career advice agencies, employment agencies, companies providing continuing training?

The results chains of vocational training and education programmes should be aligned with at least one of the universal indicators. The 'employment ratio' and 'income' universal indicators are in principle suitable for mapping the performance and demand orientation of a TVET system or parts thereof. The 'scale' indicator focuses more on the supply side of the TVET system or its constituent parts and enables conclusions to be drawn with regard to the efficiency of the interventions.

However, the standardised measurement of results using these universal indicators implies new challenges too. In terms of methodology, the question is how to prove unequivocally the causal relationships between an intervention and potential results. In evaluations, an attempt is made to clarify these relationships by means of the counterfactual approach, whereby an event is seen as the cause of an effect if the effect would not have occurred without the event. According to this, in order to establish the actual result of a measure it must be established what would have happened without this intervention. The required comparable values that can be generated using rigorous methods are difficult to determine even in individual projects. In the case of 'top-down' projects in connection with programme-based approaches, it is even more demanding to determine the counterfactual situation. Whereas experiments can be used in individual projects that concern the micro and meso level, this is often only possible to a limited extent with complex, 'systemic' interventions in the framework of PBAs, because structural results are difficult to prove by means of experiments. The question then is how the counterfactual situation can be sufficiently determined in PBAs. When using experiment-based designs, there are also ethical reservations if the chance for people to improve their living situation is made dependent on a random selection when determining the control group. As an alternative, the use of so-called 'market player opinions' can be discussed. In these, the assessments of important key people at the project are collected to measure the success of the project. Although this qualitative approach does reveal specific causalities, it does not permit representative statements to be made. It is therefore doubtful whether the use of this method alone is sufficient to substantiate project results to the necessary degree.

At political level, the question is to what extent donors are prepared to renounce 'national' interests in favour of a jointly operated results measurement system that focuses on substantiating aggregated results. Frequently, proof of a 'German' or 'French' contribution to the success of a measure is required for the political legitimisation of development cooperation projects and programmes.

Summing up, it can be said that the greater focus on cooperation within programme-based approaches also calls for a joint understanding of how project results should be measured. In our opinion, minimum standards for results measurement are required. A first step has been taken by proposing suitable universal indicators. However, standardised elements of results assessment and the related problems need to be further discussed in depth. Only then can the objective of effective and efficient cooperation between all actors involved in development cooperation be realistically achieved.

# Integrating VET into other policy domains: some thoughts about monitoring and evaluation

Noela Eddington and Ian Eddington<sup>13</sup>

**Abstract:** For two decades there have been calls for the Australian VET system to be more efficient, responsive, industry-driven, and simplified. The responses from governments of all political persuasions have generally been incremental and within the traditional VET rubric of supplying skills to the labour market.

In 2002, the State of Queensland began experimenting with different models of interaction with the VET system and its stakeholders. Since then a range of alternative industry engagement mechanisms has been trialled within an integrated approach to skills formation spanning a number of policy areas. In general, and in speculation about possible future VET systems, each of these mechanisms was initially predicated on tracking of emerging economic, social and environmental challenges faced by western democracies. From these initial experiments, Queensland is now developing an alternative holistic VET system model for 2020 which hopefully will be more able to cope with the changing nature of occupations, work, and the requirements of a carbon-constrained economy. Two early drafts of some tools and techniques being considered for managing and monitoring the system are discussed: (i) Monitoring and Performance Framework, (ii) Capability Scales

**Keywords:** Contextualising VET, Policy Integration, VET 2020

## Introduction

Since 2002, Queensland, Australia, has been questioning the assumption that a highly qualified workforce alone is sufficient to increase profitability, productivity and economic growth. It is contended that the full contribution of a skilled workforce to the economy, industry sectors and individual firms is not realised unless employers cogently address demand-side factors. Such demand-side activity on the part of employers must include responsibility for integrating attraction, development, effective utilisation and retention of skills into their people management practices within the context of a sustainable business strategy.

In trialling this contention, skills, work and industry development policies are being integrated in a pilot program in the manufacturing sector. This policy integration is, in effect, relying on collaborative governance which is actively testing the capability of the bureaucracy to operate in networks.

The pilot is considering how industry, work and skills policy might be designed to support a 'high skill equilibrium' capable of providing decent and sustainable work within a just transition under conditions of carbon-constraint.

---

<sup>13</sup> This paper represents the views of the authors: affiliations are provided for identification purposes only.

Over the seven years from 2002, Queensland has been experimenting with skill ecosystems, and other new forms of industry responsibility for skills, and has also piloted public policy integration of various kinds. Some policy thinkers now believe that VET 2020 must be closely integrated with sectoral and regional responses to changing economic conditions. The Industry Skills Policy framework must encourage all industry sectors and regions to manage sustainable skill ecosystems. Skills, as lower order issues, must be aligned to sustainable industry strategies and good workplace management practices.

The current stage of the Queensland collaborative experiments involves demonstrating effective translation of skills into productive outcomes, and identifying roles, responsibilities, systems and processes that underpin sustainable production and good jobs. The monitoring and evaluation process has provided insights for questioning the rationale and practice of traditional skills policy. We are now seeking to monitor the impact of integrated interventions on business outcomes, as opposed to measurements based solely on individual program outputs such as qualifications.

As a result of these experiments, Queensland is on the verge of adopting a Dual Skills Policy Framework with two distinct strands: Industry Skills Policy and Skills Policy for Individuals. This paper is primarily about the former, and more specifically about Industry and Government Capability Scales and a Monitoring and Reporting Framework for Industry Skills Policy in a carbon-constrained economy. The term 'industry' in the context of the emerging Industry Skills Policy also refers to and incorporates regions and communities, as it applies equally to the skill ecosystems in these contexts.

### **Targeted Problem**

The targeted problems of the on-going Queensland research and action learning processes are (i) elimination of wastage in a supply-driven VET system and (ii) optimising the value of skills to industry and the economy. We believe that, within the Queensland context: (i) centralised processes based on workforce planning, forecasting and identifying future skills needs alone are unreliable and insufficient, and (ii) skills supply focused systems abrogate industry responsibility to maintain skill attraction, development, effective utilisation and retention processes (Queensland Department of Education, Training and the Arts 2008).

Queensland's developing Industry Skills Policy is aimed at leveraging industry ownership and responsibility for sustainable skill ecosystems within sustainable businesses. Monitoring and evaluation within this context is complex and difficult for VET agencies because it requires a different set of indicators. In addition to measuring quantitative data on skills supply, we are assessing the impact of multiple programs on productivity, profitability and employment issues in a people and planet context. We contend that qualification levels alone are not a sufficient indicator of the value of skills to an economy (Scottish Government, 2007; UK Commission for Employment and Skills, 2009).



## Methodology

Queensland operates within a complex Australian VET system which incorporates (i) shared responsibility between the national and state governments; (ii) regulated public VET systems; (iii) a national qualifications framework; (iv) competency based training with some 1400 national qualifications; (v) registered training organisations operating in a 'training market' and (vi) regulated occupations in the trades.

Effecting significant change quickly within the entrenched VET system is well nigh impossible in Australia because of the deliverables imposed on the states in return for part funding of training initiatives. Accordingly, Queensland has been addressing this impasse by making small changes to components of the VET System on an incremental and politically acceptable basis. At the same time, policy strategists have been developing a new holistic mental model for VET to guide the general direction of incremental action learning based initiatives over time. Without this holistic vision for VET 2020, there is a tendency for subsequent administrators to continually meddle at the edges of the existing system. (And, as noted by Keep 2009, all the worthwhile changes at the edges have been made to the existing System.)

The holistic model for VET 2020 covers the following components of the VET System: role and purpose of VET, VET institutions, training product, pedagogy and professionalism, linkages, pathways, governance, funding, culture, regulation, accountability and impacts. The focus is on designing each of these components and combining them into a VET System 'jigsaw' in such a way so as to ensure that the integrated drivers of the System support the stated role and purpose of the VET program<sup>14</sup>. This holistic model is being researched and framed incrementally in line with the outcomes of a Ministerial Forum on VET of the Future held in Queensland in 2008 (ibid, 2008). It is also heavily influenced by the authors' involvement with international researchers working in a range of other contemporary economic, social and environmental issues such as education, sustainability, climate change, workforce development, consumption, production, and equity.

In relation to action research initiatives related to Queensland's emerging Industry Skills Policy, we now have the benefit of experience with excess of fifty-two (52) skill ecosystems, and sixteen (16) industry centres of excellence and other industry/government alliances and arrangements linking skills to workplace management and sustainable strategic business directions. These integrated policy regimes are increasingly removing the need for forecasting and planning for skills in a labour market context where supply and demand dominate the rhetoric. Instead, the State's new forms of industry engagement allow the contextualization of skills within a sustainable business development debate. That is, skills can be contextualized and managed in a realistic context where they are utilized, influenced and owned by industry, regions or communities (all of which can develop sustainable skill ecosystems able to respond as economic and social issues vary.) We (the au-

---

<sup>14</sup> We would argue that the role and purpose of VET should be measured at four levels, namely the economy, industry, enterprises and individuals but within the general context of sustainable development.



thors) also believe that this industry development approach is the key to stimulating employer demand for quality jobs, thereby creating a demand-pull for skills. As industry capacity develops in this regard, skills are more likely to be available where they are needed and be more effectively utilised.

Queensland's current action research is exploring policy coordination in the manufacturing sector. It is known as the Workplace Partnership and Productivity (WP&P) pilot. Three agencies have combined their industry development programs to support a sustainable manufacturing sector. Industry development, work and skills policies are being coordinated to deliver an holistic action plan that industry undertakes to implement. Business reviews, sustainable business strategies, efficient operating systems and effective people management in safe and decent 'green jobs' are being encouraged through 'partnership' negotiating processes. The Australian industrial relations environment has a recent history of excluding unions from workplace negotiations. This heritage of exclusion is currently being dismantled and one aim of the current WP&P pilot is to demonstrate improved business performance and productivity through partnerships as reported by Black and Lynch (2003).

The integrated government activity in the WP&P pilot is challenging to public agencies which are generally inexperienced in operating within client driven networks. Our research suggests that government agencies are accustomed to working in *state* and *market* modes of governance where the drivers focus on numbers/quantity and efficiency. However, *network* modes of governance in coordinated policy scenarios are difficult for government agencies (Keast et al, 2004). Networks are based on relationships, a collective sense of mutual responsibility, trust and power sharing: consequently, agency accountability in networks needs careful mentoring and attention. We (as authors) go so far as to suggest that government and industry capability to operate effectively in networks needs to be developed, and Capability Scales to identify behaviours requisite to operating within networks are being constructed.

In addition, the action research being undertaken in the WP&P pilot also seeks to refine (i) a Monitoring and Performance Framework for the holistic 'industry development' process and (ii) Industry and Government Capability Scales. These are discussed in the next section.

## Results and Perspectives for Further Development

### *For the Described Instruments*

- Monitoring and Evaluation Framework

The draft Monitoring and Evaluation framework presented here is intended to provide a starting point for negotiations on a specific framework for the industry, region or community involved. It attempts to provide indicators that are acceptable across all three governance modes for baseline activity, facilitative activities and program effect data. The *baseline data* (see Table 1) will support the graphing of trend lines in *facilitative* and *effect* data as the pilot progresses. The *facilitative data* is intended to measure a range of context, process and learnings dimensions such as (i) how well the stakeholders are collaborating, (ii) their developing capability levels and (iii) learnings from the processes. *Effect* data will enable the development of trend lines resulting from the integrated service delivery process on outputs, outcomes, impacts and business performance. Table 1 contains examples only of the type of indicators and measures that might be agreed upon by stakeholders in the Manufacturing WP&P pilot.

	Indicator Type	Function	Indicator Example	Measures
EFFECT	Output	To assess outputs such as workforce management tools, number of qualifications, number of awareness sessions, and number of firms committed.	(i) Work policies that support attraction and retention (participation) e.g. work/life balance, career paths, health and safety, negotiated pay scales, job redesign, high performing work practices such as commitment to learning, open-mindedness and shared vision.	(i) Specific 'demand-side' workplace management practices in place e.g. partnerships process that delivers improvement in attraction, development, effective utilisation and retention of workers e.g. work/life balance, career paths, health and safety, negotiated pay scales, job redesign, high performing work practices such as commitment to learning, open-mindedness and shared vision.
			(ii) Skill utilisation status	(ii) (a) Story: outlining how the sector creates and uses knowledge to enhance competitive advantage. (b) Employee/Employer surveys: Opinion of % skill utilisation (c) QMI utilisation measures
			(iii) Employees with a Certificate III	(iii) % of staff with a Cert III or above in the following categories: • Professional / managers • Skilled Trades • Intermediate skilled workers • Elementary skilled workers
	Outcome	<ul style="list-style-type: none"> <li>To assess outcomes related to trends in changes or improvements that result from interventions</li> <li>To define a key outcome: Enterprise bargaining agreement developed through partnership process</li> </ul>	(i) Industry adopts sustainable business practices – people, profit, planet (ii) Industry utilising the social partnership process.	(i) Evidence of industry capability improving (see thermometer chart) (ii) QMI measures re productivity (iii) Negotiated agreement around improved workplace management practices aligned to strategic and operational needs of the firm or network.
	Impact	To assess 'sustainable' impacts that result from interventions	Trends in industry profits, labour productivity, people management and carbon footprints	(i) Trends in: <ul style="list-style-type: none"> <li>Business viability or profits, labour productivity</li> <li>Employee satisfaction</li> <li>Labour utilisation</li> <li>Improved WH&amp;S performance</li> <li>Retention/ staff turnover</li> <li>Primary carbon footprint/ reduced waste/ emissions</li> </ul> (ii) Story: Outline how the sector supports participation of disadvantaged groups and how VET Investment has supported this participation (iii) Stories: Case studies of how VET Investment is supporting good workplace practices that impact on individual firm or industry performance across profit, people and planet issues.
	Performance	To assess the capability of the manufacturing sector to remain sustainable in variable economic cycles.	Capability of the manufacturing firms/ sector to align strategy, operations and workplace strategies, including operating effectively in an industry-led industry-driven ecosystem through variable economic cycles.	Story: summarising the original context, what happened, what worked /did not work, industry performance in managing industry-led, industry-driven government interventions (industry development, work, skills) geared to support sustainable profit, workforces and skills. Comment should be made on a range of economic, social and environmental indicators e.g. technology diffusion, improved business processes, labour productivity, total factor productivity, workplace practices, and waste / carbon management, adaptation / mitigation strategies, trends in competitive advantage, how employer responses to changing business environments are shaping labour demand, contribution to formal and informal training etc

## Notes:

Firms / clusters must commit to providing the agreed business data. This needs to be a funding requirement. The Most Significant Change story telling process, or an adaptation of it, provides rich data on changes that occur. Stories are sometimes considered more useful in reporting some indicators than other mea

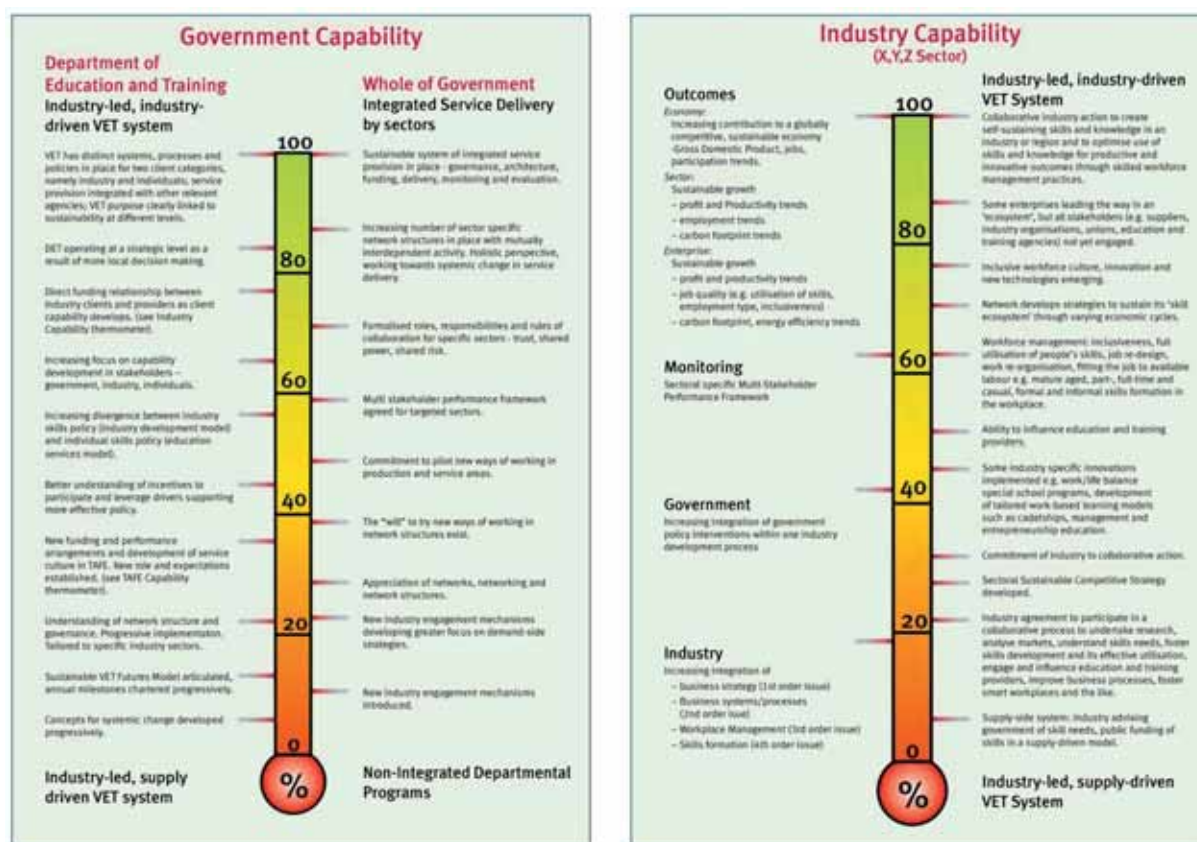
The Framework is being trialled in 2010 - 2011 in the WP&P pilot. Industry, government and unions will negotiate a set of indicators in each of the three categories that is relevant to the specific skill ecosystem. Quantitative data and qualitative data in the form of stories will be mixed and matched to provide both industry and government with rich information on base-line, facilitative and impact data. Specific indicators will be developed through the 'partnership' process being used in this pilot but could equally be developed using alliancing principles or some other form of collaborative arrangement.

The indicators will generally reflect the issues for improvement identified in business and people management diagnostic processes (please see Table 1).

- Industry and Government Capability Scales

The Industry and Government Capability scale tools (Figure 1) will also be refined in the WP&P project. They articulate the types of behaviours that need to be developed incrementally by these specific stakeholders in order to optimise value from integrated policy designed to support economic, social and environmental outcomes.

**Figure 1: Capability Scales**



Our experience is that, unless the ‘capability’ issue is addressed, stakeholders tend to follow ‘business as usual protocols’ and collaborative networks are ‘business as unusual’ (for both government and industry) from multiple perspectives: governance, roles, responsibilities, accountability, monitoring and reporting, employment and just transition goals and a high skill sustainable production function. These capability scales could be used to guide funding directly to industry as its ability to manage demand for and utilisation of skill increases. Funding could be conditional upon the development of demand-side factors.

### **Potentials and challenges for the use in/for international comparisons**

The potential for the use of the tools outlined above relies on the underpinning philosophy of the role and purpose of VET in defining country specific skills policy. They would only be of value in regimes that subscribed to industry ownership and responsibility for their own skill ecosystems, where integrated policy environments leveraged demand-side support for attraction, development, effective utilisation and retention of skills through good leadership and people management practices, and where skills policy was clearly linked to higher order sustainability goals.

Internationally, some countries are moving towards this skills policy scenario, particularly where large investments in skills in recent decades have lead to little or no improvement in comparative labour productivity. For example, the UK, the OECD Local Employment and Economic Development program, New Zealand and Australia generally recognise that skills alone are not enough; skills need to be effectively utilised in order to transform their value into economic benefit. Integrated Industry Development Skills Policy has the potential to create a demand-pull for skills which in turn supports employment policies.

The Global Financial Crisis has also prompted some countries to consider the value to their economies of traditional supply-driven skills policy. There is growing recognition of the complexities that circumscribe and potentially restrict the value of skills in workplaces. There is growing popularity of ‘workforce development’ strategies being used in conjunction with traditional skill supply policies. However, the limitations of the latter used in isolation, generally in the context of labour market rhetoric, are increasingly being recognised. The more contemporary context for skills policy is in workplaces, regions and communities where the influences on skills can be more effectively managed.

## References

Black, S.E. and Lynch, L.M. 2003, 'The new economy and the organization of work' in The Handbook of the New Economy, ed. Jones, D. New York: Academic Press.

Black, S.E. and Lynch, L. M. 2004, 'What's driving the new economy? The benefits of workplace innovation', Economic Journal, February.

Keep, E. 2009, Workforce improvement strategies in the UK presented at Workforce Development Forum, Department of Employment, Economic Development and Innovation, Brisbane.

Keep, E, Mayhew, K, and Payne, J, 2006, From skills revolution to productivity miracle – not as easy as it sounds? Oxford Review of Economic Policy vol.22 no. 4.

Queensland Department of Education Training and the Arts, 2008, The role and purpose of VET, VET of the Future Ministerial Forum, Brisbane, May.

The Scottish Government, 2007, Skills for Scotland, a lifelong skills strategy, Edingburgh.

United Kingdom Commission for Employment and Skills, 2009, Ambition 2020: world class skills and jobs for the UK.

# The Study on the Evaluation of Secondary Vocational School in China

*Dr. Xiaorong Nie, Mr. Ziyuan Wang and Mrs. Zhibao Pan*

**Abstract:** The functions of the evaluation of Secondary Vocational School in China are two: guiding the school to the policies of national vocational education and to provide identification of accountability for the improvement of the quality in secondary vocational education. The study, based on the survey from the Secondary Vocational Schools in China, tries to find out to what degree the practitioners accept the National Evaluation System of the Secondary Vocational School in China as the instrument for school improvement. Besides, the study also analyzes the impact of the evaluation of the perspectives of the practitioners on the effectiveness of the schools, and the relevance between the background of the practitioners and school evaluation and effectiveness.

**Keywords:** acceptance of the criteria, school effectiveness, co-efficient

## Introduction

The National Evaluation System of the Secondary Vocational School in China has been conducted since 2004 (China National Education Department, 2004), aimed at developing strategies to synthesize fragmented efforts around a familiar framework, and at encouraging vocational schools to take on new roles to create renewed conditions and to involve themselves in pursuing the purposes (China National Planning Commission, 2006). The criteria of Evaluation cover three main dimensions with 41 items in details. The three dimensions are: schooling direction and the quality of education; basic conditions of school and rational utilization; and management of reform and innovation. The criteria in details are presented in Table1. The study will check if the evaluation has reached its aims, as a result of school improve.



**Table 1: Criteria of National Evaluation System of the Secondary Vocational School in China**

1. guidance and concepts of schooling
2. aims of education and targets of employment orientation
3. number of attendants for education diploma
4. number of trainees
5. rate of employment
6. rate of double certificate (diploma and profession)
7. achievement of academic research
8. training model for the shortage of professional talent
9. school honor
10. campus layout
11. building area and its reasonable utilization
12. ratio of teachers-students
13. educational background of Full-time teachers
14. professional structure of full-time teachers
15. exercitation equipment and experiment
16. internship base outside school
17. internetwork construction
18. computers equipment for teaching
19. multimedia classroom
20. information facility and its utilization in teaching
21. books and data in library and frequency of borrowing
22. sports facilities
23. health facilities
24. funding sources and use
25. multi-level form of schooling
26. joint school



27. cooperation of school and industry
28. flexible academic structure
29. leadership team building
30. management team building
31. teaching faculty building
32. moral education
33. school culture
34. the vocational and professional guidance
35. specialty building
36. curriculum reform
37. innovation of teaching methods
38. teaching materials selection and management
39. system construction and operation mechanism
40. internal management system reform
41. innovation of teaching management

### Targeted Problems

The study focuses on: 1) to what degree the participants accept “the System of the Evaluation of the Secondary Vocational Schools”; 2) if the evaluation has impact on the school effectiveness and; 3) if the backgrounds of respondents have the effects on their acceptance to the evaluation and school effectiveness.

**Table 1: Criteria of National Evaluation System of the Secondary Vocational School in China**

1. guidance and concepts of schooling
2. aims of education and targets of employment orientation
3. number of attendants for education diploma
4. number of trainees
5. rate of employment
6. rate of double certificate (diploma and profession)

7. achievement of academic research
8. training model for the shortage of professional talent
9. school honor
10. campus layout
11. building area and its reasonable utilization
12. ratio of teachers-students
13. educational background of Full-time teachers
14. professional structure of full-time teachers
15. exercitation equipment and experiment
16. internship base outside school
17. internetwork construction
18. computers equipment for teaching
19. multimedia classroom
20. information facility and its utilization in teaching
21. books and data in library and frequency of borrowing
22. sports facilities
23. health facilities
24. funding sources and use
25. multi-level form of schooling
26. joint school
27. cooperation of school and industry
28. flexible academic structure
29. leadership team building
30. management team building
31. teaching faculty building
32. moral education
33. school culture

34. the vocational and professional guidance
35. specialty building
36. curriculum reform
37. innovation of teaching methods
38. teaching materials selection and management
39. system construction and operation mechanism
40. internal management system reform
41. innovation of teaching management

## **Methodology**

### *Sample*

For case study 65 Secondary Vocational Schools in Jiangxi province were chosen as the main objects of the survey. The chosen schools with different size from less than 1000 students to more than 2000 were located in such areas as urban, suburban and rural. The respondents of the survey were practitioners, mainly principal, vice principal or dean since they were experiencing evaluation and familiar to it.

### **Methods of survey**

Survey questionnaires with Likert scale were used to measure to what degree the schools accept the evaluation criteria. There were five independent variables and 41 dependable variables in questionnaires. The independent variables were: gender, position, school type, school location and school size; while the dependable variables were the 41 criteria items as shown in Table 1. Five scales were set as: very important, important, general, unimportant and unimportant at all. Also the schools were asked to measure the effectiveness of their school by self-evaluation. Five scales were: very efficient, efficient, general, not efficient, not efficient at all.

The survey paper were mailed to 65 Secondary Vocational Schools, of which, 15 returned the filled paper, rated 23%. After inspection, the paper with all the criteria scored “very good” was regarded as null, and the other 14 respondents’ filled papers met the survey requirements, pass rated 93%.

## **Data Analysis**

The survey data were analyzed using the SPSS software. Primarily means were comparisons, bi-variate correlation analysis and factor analysis ANOVA.

## **Results and Perspectives of further Study**

### *Respondents' acceptance to criteria of the school evaluation*

The statistics results show the scores of the importance of evaluation criteria from 4.64 to 3.29 with standard deviations from 0.99 to 0.47, and the overall average is scored 3.91. Of 41 items 18 items are considered "very important" (see Table 2).

By Pehkonen's the standard of the calculation of the acceptance, the nine items are considered very positive, namely 90% of the respondent thought as very important (Pehkonen, 1993). The other 32 items are also considered positive, and no items considered negative (see Table 3).

**Table 2: The acceptance degree of the participants to the criteria of Evaluation**

<b>Criteria</b>	<b>Means</b>	<b>S. D.</b>
1. guidance and concepts of schooling	4.64	0.54
2. aims of education and targets of employment orientation	4.5	0.65
7. achievement of academic research	4.36	0.63
5. rate of employment	4.29	0.73
29. leadership team building	4.21	0.8
31. teaching faculty building	4.21	0.58
9. school honor	4.14	0.66
12. ratio of teachers-students	4.14	0.53
15. exercitation equipment and experiment	4.14	0.77
18. computers equipment for teaching	4.14	0.53
24. funding sources and use	4.14	0.53
35. specialty building	4.14	0.66
37. innovation of teaching methods	4.07	0.47
11. building area and its reasonable utilization	4	0.68
13. educational background of Full-time teachers	4	0.55
36. curriculum reform	4	0.68
39. system construction and operation mechanism	4	0.55
41. innovation of teaching management	4	0.55
16. internship base outside school	3.93	0.62
8. training model for the shortage of professional talent	3.92	0.73
32. moral education	3.86	0.66
33. school culture	3.86	0.53
30. management team building	3.82	0.73
6. rate of double certificate (diploma and profession)	3.79	0.7
10. campus layout	3.79	0.8

17. internetwork construction	3.79	0.58
19. multimedia classroom	3.79	0.58
14. professional structure of full-time teachers	3.71	0.61
22. sports facilities	3.71	0.61
23. health facilities	3.71	0.61
25. multi-level form of schooling	3.71	0.73
27. cooperation of school and industry	3.71	0.99
38. teaching materials selection and management	3.71	0.73
40. internal management system reform	3.71	0.73
3. number of attendants for education diploma	3.64	0.63
34. the vocational and professional guidance	3.64	0.63
20. information facility and its utilization in teaching	3.57	0.65
26. joint school	3.57	0.76
21. books and data in library and frequency of borrowing	3.5	0.65
4. number of trainees	3.37	0.77
28. flexible academic structure	3.29	0.73

**Table 3: The rate of the acceptance to the criteria**

<b>Criteria</b>	<b>Negative 1-2</b>	<b>General 3</b>	<b>Positive 4-5</b>	<b>Positive Rate %</b>
1. guidance and concepts of schooling	0	0	14	*
2. aims of education and targets of employment orientation	0	1	13	*
3. number of attendants for education diploma	1	8	5	
4. number of trainees	0	2	12	
5. rate of employment	0	2	11	
6. rate of double certificate (diploma and profession)	0	5	9	
7. achievement of academic research	0	1	13	*
8. training model for the shortage of professional talent	0	4	10	
9. school honor	0	2	12	
10. campus layout	0	6	8	
11. building area and its reasonable utilization	0	3	11	
12. ratio of teachers-students	0	1	13	
13. educational background of Full-time teachers	0	2	12	
14. professional structure of full-time teachers	0	5	9	
15. exercitation equipment and experiment	0	3	11	
16. internship base outside school	0	3	11	
17. internetwork construction	0	4	10	
18. computers equipment for teaching	0	1	13	*
19. multimedia classroom	0	4	10	
20. information facility and its utilization in teaching	0	7	7	
21. books and data in library and frequency of borrowing	0	8	6	
22. sports facilities	0	5	9	
23. health facilities	0	5	9	

24. funding sources and use	0	1	13	*
25. multi-level form of schooling	0	6	8	
26. joint school	1	5	8	
27. cooperation of school and industry	1	6	7	
28. flexible academic structure	1	9	4	
29. leadership team building	1	0	13	*
30. management team building	1	0	13	*
31. teaching faculty building	0	1	13	*
32. moral education	1	1	12	
33. school culture	0	3	11	
34. the vocational and professional guidance	0	6	8	
35. specialty building	0	2	12	
36. curriculum reform	0	3	11	
37. innovation of teaching methods	0	1	13	*
38. teaching materials selection and management	0	6	8	
39. system construction and operation mechanism	0	2	12	
40. internal management system reform	0	6	8	
41. innovation of teaching management	0	2	12	

Note \* means 90% of participants holding positive attitude

### **Respondents' self-assessment of their school effectiveness**

The respondents were asked to assess the effectiveness of their schools by the criteria of the evaluation. The results show the total average score 3.75, the highest 4.36 and the lowest 3.21. Of 41 items the average score of 5 items are above 4.00. They are: rate of employment, guidance and concepts of schooling, aims of education and targets of employment orientation, computers equipment for teaching, leadership team building. There are no items scored less than 3.0 as average (see Table 4).



**Table 4: School effectiveness by self evaluation**

<b>Criteria</b>	<b>Means</b>	<b>S. D.</b>
5. rate of employment	4.36	0.63
1. guidance and concepts of schooling	4.29	0.49
2. aims of education and targets of employment orientation	4.07	0.47
18. computers equipment for teaching	4	0.68
29. leadership team building	4	0.68
3. number of attendants for education diploma	3.93	0.47
9. school honor	3.93	0.65
15. exercitation equipment and experiment	3.93	0.73
24. funding sources and use	3.93	0.73
31. teaching faculty building	3.93	0.62
36. curriculum reform	3.93	0.73
41. innovation of teaching management	3.93	0.62
6. rate of double certificate (diploma and profession)	3.86	0.66
7. achievement of academic research	3.86	0.66
12. ratio of teachers-students	3.86	0.66
13. educational background of Full-time teachers	3.86	0.36
30. management team building	3.86	0.53
32. moral education	3.86	0.53
35. specialty building	3.86	0.53
16. internship base outside school	3.79	0.43
25. multi-level form of schooling	3.79	0.58
26. joint school	3.79	0.58
37. innovation of teaching methods	3.79	0.58
39. system construction and operation mechanism	3.79	0.43
38. teaching materials selection and management	3.71	0.61

11. building area and its reasonable utilization	3.64	0.63
14. professional structure of full-time teachers	3.64	0.5
23. health facilities	3.64	0.5
33. school culture	3.64	0.5
10. campus layout	3.57	0.76
19. multimedia classroom	3.57	0.51
34. the vocational and professional guidance	3.57	0.51
40. internal management system reform	3.57	0.76
8. training model for the shortage of professional talent	3.5	0.65
17. internetwork construction	3.5	0.76
22. sports facilities	3.43	0.65
27. cooperation of school and industry	3.43	0.85
21. books and data in library and frequency of borrowing	3.36	0.5
4. number of trainees	3.29	0.73
28. flexible academic structure	3.29	0.73
20. information facility and its utilization in teaching	3.21	0.58

### **The relevance of evaluation and school effectiveness**

The study did analysis on the relevance co-efficiency of evaluation and school effectiveness. The results show that school effectiveness is very closely related to evaluation. The correlation co-efficiency of 25 items is above 60%. The top 4 items over 0.80 are: rate of employment 0.96, campus layout 0.85, computers equipment for teaching 0.85, exercitation equipment and experiment 0.84. But there are still two items whose co-efficiencies are less than 0.20. They are: moral Education 0.16 and the number of attendants for education diploma 0.17 (see Table 5).

**Table 5: Co-efficiency of evaluation and school effectiveness**

<b>Criteria</b>	<b>Importance</b>	<b>effectiveness</b>	<b>coefficient</b>
5. rate of employment	4.29	4.36	0.96
10. campus layout	3.79	3.57	0.852
18. computers equipment for teaching	4.14	4	0.847
15. exercitation equipment and experiment	4.14	3.93	0.84
27. cooperation of school and industry	3.71	3.43	0.791
25. multi-level form of schooling	3.71	3.79	0.758
40. internal management system reform	3.71	3.57	0.741
22. sports facilities	3.71	3.43	0.723
28. flexible academic structure	3.29	3.29	0.708
29. leadership team building	4.21	4	0.706
19. multimedia classroom	3.79	3.57	0.702
31. teaching faculty building	4.21	3.93	0.694
20. information facility and its utilization in teaching	3.57	3.21	0.676
34. the vocational and professional guidance	3.64	3.57	0.676
38. teaching materials selection and management	3.71	3.71	0.668
33. school culture	3.86	3.64	0.662
26. joint school	3.57	3.79	0.653
23. health facilities	3.71	3.64	0.651
39. system construction and operation mechanism	4	3.79	0.651
2. aims of education and targets of employment orientation	4.5	4.07	0.623
24. funding sources and use	4.14	3.93	0.62
30. management team building	3.82	3.86	0.62
36. curriculum reform	4	3.93	0.62
37. innovation of teaching methods	4.07	3.79	0.62

17. internetwork construction	3.79	3.5	0.612
21. books and data in library and frequency of borrowing	3.5	3.36	0.595
8. training model for the shortage of professional talent	3.92	3.5	0.567
11. building area and its reasonable utilization	4	3.64	0.536
16. internship base outside school	3.93	3.79	0.524
4. number of trainees	3.37	3.29	0.508
7. achievement of academic research	4.36	3.86	0.497
35. specialty building	4.14	3.86	0.496
1. guidance and concepts of schooling	4.64	4.29	0.471
6. rate of double certificate (diploma and profession)	3.79	3.86	0.427
9. school honor	4.14	3.93	0.404
14. professional structure of full-time teachers	3.71	3.64	0.398
13. educational background of Full-time teachers	4	3.86	0.382
12. ratio of teachers-students	4.14	3.86	0.279
41. innovation of teaching management	4	3.93	0.225
3. number of attendants for education diploma	3.64	3.93	0.165
32. moral education	3.86	3.86	0.155

### ANOVA analysis

By ANOVA analysis through SPSS software the study checked if the backgrounds of respondents had the effects on their acceptance to the evaluation and school effectiveness. The data show there is not significant difference among the respondents with different backgrounds (see Table 6 and Table 7).

**Table 6: ANOVA importance**

background	group	No.	means	S. D.	F-ratio	Sig
gender	female	1	3.34	0	3.11	0.49
	male	13	3.96	0.63		

position	principal	4	3.90	0.52	0.85	0.61
	Vice-principal	5	3.93	0.83		
	dean	5	3.91	0.49		
School type	Vocation high school	2	3.87	0.40	1.01	0.54
	Vocation school	12	3.92	0.65		
location	urban	8	3.87	0.72	0.68	0.58
	suburban	6	3.97	0.54		
	Rural	0				
School scale	➤ 2000人	11	3.98	0.66	1.20	0.49
	1000-2000	3	3.67	0.49		
	< 1000 人	0				

**Table 7: ANOVA effectiveness**

background	group	No.	means	S. D.	F-ratio	Sig
gender	female	1	3.51	0	1.42	0.45
	male	13	3.77	0.6		
position	principal	4	3.74	0.64	0.92	0.59
	Vice-principal	5	3.81	0.62		
	dean	5	3.70	4.75		
School type	Vocation high school	2	3.84	0.60	0.90	0.54
	Vocation school	12	3.73	0.57		
location	urban	8	3.70	0.59	0.74	0.50
	suburban	6	3.81	0.60		
	Rural	0				
School scale	➤ 2000人	11	3.80	0.60	1.83	0.43
	1000-2000	3	3.60	0.48		
	< 1000 人	0				

## **Conclusion and further study**

Data analysis shows that the respondents gave a high priority in implementing education policy and educational objectives to develop professional talents, as well as employment rate. School evaluation as a means of school improvement has a positive impact of its school effectiveness, helping schools to implement the education goals and to develop profession talents. Research also finds that respondents felt that the educational funding resource and facilities were important as well as leadership and teacher team building. They also had a direct relationship to educational effectiveness. Study recommends that the improvement of the conditions of vocation schools will be the important work ahead.

The research group sent out paper of questionnaires to 65 schools for survey, but only 15 respondents, rated 23%, filled the paper, resulting in too small sample, which affects whether the study findings reflect the actual situation. Therefore the study will verify the reliability and validity in the follow-up study.

Evaluation is a complex work, which cannot reflect solely in scores. Moreover, the conditions of all the schools are not on the same level, therefore scoring by the same standard can not fully tell out the work in different schools. Future study is needed to do on the ground survey and qualitative analysis to complement the single-sum of the deficiencies.

## **References**

- China National Education Department, May 2004. "The Criteria of Evaluation of the Key Secondary Vocational Schools in China".
- China National Planning Commission, March 2006. 10<sup>th</sup> Five-Year Plan of National Economical and Social Development.
- Pehkonen, E. What are the Finnish Teacher Educators' Conceptions About Teaching of Problem-solving in Mathematics. *European Journal of Teacher Education* 1993. 16. pp. 237-256.

# Monitoring of VET systems – The perspective of comparative education

*Dr. Uwe Lauterbach*

**Abstract:** This paper looks at the contribution that comparative education (CE) as a developed research perspective can make to questions of monitoring and benchmarking of VET systems. A plea is made for overcoming the separation between CE and comparative research in TVET (VBBF). Some lessons can be learned: Monitoring of national TVET systems needs a combination of the internal and external view of the researchers. Ideally this is a combination of a native and a foreign researcher. Only the combination of qualitative and quantitative will elicit full pictures of Vet systems' performance and the monitoring of national TVET systems or of special aspects is to integrate in the national, regional, etc. context of the society, economy etc.

**Keywords:** Comparative Education, Comparative Research in Technical and Vocational Education and Training

## 1. Introduction

For decades, *Vergleichende Berufsbildungsforschung* [Comparative Research in Technical and Vocational Education and Training] – *German abbreviation: VBBF* – eked out a shadowy existence in Germany. Due to the changing general conditions, which could be summarized with the slogans *globalization* and *European integration* and which lead to an intensive international and supranational cooperation, the *VBBF*, which accompanies this change with studies of the conditions, processes, and consequences of the acquisition of qualifications and their general settings (e.g. culture, society, politics, history, economy), has again been attracting more attention during the past few years. A quickly increasing need for research is noticeable.

When analyzing selected results of the research within *VBBF*, questions arise in particular regarding the research interest, the objects, the theories, as well as the paradigms and methodologies applied. The development of generally accepted standards has previously quite obviously failed in this regard. There is need for action in order to systematize the research studies and to analyze the theories and methodologies of the *VBBF* and the related social sciences.

The *Vergleichende Erziehungswissenschaft (VE)* (German expression of *Comparative Education*) – and the *Comparative Education (CE)* present themselves as direct reference *disciplines* due to the fact that their primary educational interest in knowledge and the interlinking of general and vocational education form an excellent basis for this purpose. The *CE*, which is especially well-established in the USA and in Great Britain, naturally deals with objects of

research referring to the general and vocational education. The 'first' German (comparative) TVET educationists Kerschensteiner and Fischer were also embedded in the *scientific community* of *CE/VE* and participated in its worldwide discourse up to the 1930's.

The researchers of *CE/VE* are able to base on a wide and deep basis of theories and methodologies that has been developed in the course of a frequently controversial dialog lasting for almost two centuries. In spite of this seemingly clear initial position, the results of the research work must be introduced into a scientific community in which a constant discourse on the leading issues of knowledge, the theories, and the methodological approaches takes place which is mostly fruitful and leads further but which, apart from that, was and still is fierce, viz. especially because researchers of other social sciences and arts, such as pedagogy, sociology, psychology, economics, history, or comparative cultural research introduce their concepts as 'side-intruders'. This means that the *CE/VE* has been forced to get involved in debates about fundamental academic issues on theory ever since the beginning of 1800.

Despite the turbulent debates that took place three decades ago and clarified the positions between 'empiricists' and 'hermeneuts', even *CE/VE* cannot present any indisputable results referring to a closed theory of the comparison of educational systems (included TVET). The fierce discourse on the leading interests and theories of knowledge rather indicates that the theories and methodologies of the international comparison are in an open development process. It does not therefore come as a surprise that there is no standard work containing an overview of the development, the state of theory and methodology, and of important results of the research (Glowka 1997, 55). As early as in 1981, *Kelly/Altbach* (1981, 1) brilliantly commented on this situation by indicating that, in contrast with history, economics, sociology, and psychology, the *CE/VE* neither have discipline-related methods of research at their disposal nor is there any mutual basic knowledge. In spite of this plurality, a 'minimum standard' exists in contrast to the *VBBF*. On the other hand the non-existing connection to the international standard, criticized by *Czycholl* (1975, 15), still applies to the *VBBF* today with reference to research interest, theories, paradigms, and methodologies. Moreover in the most countries there is no differentiation between *VBBF* and *CE*.

Because in the most countries there is no differentiation between *VBBF* and *CE* we (1) check the reasons of the separation in Germany and (2) draw conclusions for the *VBBF* and the monitoring of TVET.



## 2. Phases of development of the CE/VE

Five phases of the CE/VE can be described up today:

a. *The stock-taking of relevant educational information and data on national or regional education systems*

This was made in order to obtain the principles and constituent rules for certain regularities (*nomothetic approach*) and, at the same time, to serve as a 'model' for the development of the education system itself (*melioristic approach*) information was obtained by means of specific 'commissioned' itineraries which were then, brought in' to each situation. This 'need' came up between the end of the 18th till the middle of the 19th century when *national education systems* were established. The first representative mentioned is *Marc Antoine Jullien de Paris* who combined these purpose-oriented descriptions with a *nomothetic theoretical approach* and, in addition, made fundamental remarks on the establishment of a comparative educational science.

b. *The typical features of national education systems*

This search for the cross-system elements was countered with the analysis of the driving forces and determining factors making up the typical features of national education systems by *Michael Sadler* around 1900 by means of historical and education philosophical studies (*idiographic approach*).

c. *Integration of the phases 1 and 2 from the 1920's onward*

E.g. by the attempt of an overall analysis of the education systems in the individual countries (*global analysis*) with the core: nature and main focuses of national education systems, education system and pedagogy as the product of national character and the nation as an integrated intellectual whole, historical orientation with mostly genetic, historico-humanistic methods, use of 'complementary sciences' in order to adequately research into the evolution of the education systems in comparative analyses (*evolutionistic approach*). Typical representatives are *Kandel Schneider*, and *Hans*.

d. *Empirical approaches*

The first attempt to obtain an empirical quantity of the *evolutionistic approach* of the genesis of national systems and of the transnational comparison was made by *Bereday* who wanted to extract hypotheses on the basis of empirical data and to gain, by means of induction, the 'world formula' of the CE/VE which was meant to show the important regularities between society and education system. His followers *Noah* and *Eckstein* left the qualitative basis and set it against the *empirico-quantitative paradigm*. In consequence, the *dispute of paradigms* arose between the metrists and hermeneuts in the 1960's and 1970's.

The empirical dimension of the comparison was further developed ever since for pragmatic reasons because the evolutionistic global analyses of education systems marginalize important single problems and because their more in-depth handling by means of sociological methods is not provided for. The problem approach, which deals with functions of partial aspects of education systems on the basis of *empirical analyses* by way of comparison, was therefore set up as a deliberate alternative plan.

e. *Resolving the dichotomy between arts and social sciences*

The dichotomy between arts and social sciences has nowadays been resolved in favor of an integrating view. Both notions have their merits but, on the other hand, they are not able to deal with the complex subject of education system as a whole. *Functional analyses* are therefore carried out. The *historical approach* and the *sociological approach (empirico-analytical approach)* are used in order to reveal the complex subject. While the methods of the *historical approach* are undisputed, the results nevertheless often not being relevant to practice, there is still need for discussion between the hermeneutic qualitative and the empirical quantitative research approaches. Results of comparative empirical studies, whose research design was developed without any reference to the hermeneutically oriented qualitative social science and to arts, virtually promote the misinterpretation.

This is the reason why present-day researchers who want to include the whole dimension of the complex object of research of education system in the corresponding context use the *empirico-analytical approach* in combination with a humanistic-hermeneutic dimension (*pragmatic approach*) as *multi-level analysis*.

### **3. Interests in knowledge and comparative theories**

The statement commonly made over and over again within the *VBBF* to the effect that „it cannot be compared because it is not the same“ is confronted with the fundamental theoretical principle of science that *scientific interest in knowledge* can be implemented by way of *comparison* as a means of analysis of „something else“. The comparison has a central meaning for the *social sciences* (Schriewer 2000, 5). Besides this fundamental meaning, there are special trends in individual disciplines which have made the comparison of systems or subsets of these units the central object of their research work. The position that the *VBBF* defines itself on the basis of the method of comparison can therefore not be maintained either. It's rather the objects of research and the research interest that are constituting factors for the *VBBF*. The international and intercultural comparison of general and vocational education systems includes its elements by means of defined and well-founded system categories. The objectives of research directly related to that refer to the categories of the comparison with the *range of results of the research from correspondence to differences*, with many

variants inbetween. Research studies which are apparently not comparative in their design, such as idiographic monographs or problem analyses, should be included since they likewise follow this concept as *implicite comparison*. The interest in knowledge of comparative studies, their reference framework in the social reality, their usefulness and scope are methodically presented by means of *tertium comparationis*.

The *CE/VE* as practice-related research provides the public, the actors involved, the education planning and administration, as well as the politics with categorized patterns of interpretation for the structuring of reality, and consequently it prepares solution concepts for pending reforms. The definition of the practical objectives of research is closely connected with the controversial issue as to the extent to which the *CE/VE* could make its *contribution to improving the education system and present measures for the planning of instructions for actions* in practice. This *functionalistic paradigm* was a constituting and dominating element of theories from the very beginning until the 1970's. The *functionalism* itself shows the close correlation between the current research subjects and the general political settings in each case.

The *research practice-oriented paradigms* have been connected with the presented *fundamental theories and paradigms* that refer to a certain concept of the theory of science, e.g. to the positivism, to evolutionary theories, to the function of the experiment in social research, to the induction theory, etc., ever since the constitution of the *CE / VE*. Since the *CE/VE* as integration science receives and adapts both humanistic and sociological theories, the discourse that existed or still exists in the corresponding sciences is just as due to be dealt with in the *CE/VE* and in the *VBBF*. Several levels are mostly encountered in this regard. If the *melioristic paradigm* is the basis, then further analyses regarding the validity of additional or corresponding theories, such as *borrowing*, *lending*, *best practice*, or generalizing regularities between education system and society must be carried out and, moreover, put in the global context of educational cooperation with the third world and the plausible theory of global system there, with the reference to the *evolutionistic paradigm* then becoming apparent.

The sciences giving the impetus for the four comparative functions brought out are empirically oriented ones, such as economics, sociology, and psychology, in particular regarding the theories of (c) *evolutionistic function* and (d) *quasi-experimental function*, whereas the (a) *idiographic function* and (b) *melioristic function* rather refer to arts, especially to the science of history and the humanistically oriented pedagogy, and their hermeneutic methods: The four functions represent at the same time the phases of development of the *CE/VE*. Whilst the melioristic function can be directly derived from practice-based motivations of comparative research, the three other functions are closely connected with the *scientific theoretical*

*position* assigned to the *VBBF* or to the *CE/VE*. The paradigms and theoretical models likewise represent certain methods of research, such as hermeneutics or metrics.

Even if the interest in knowledge of comparative analyses were only limited to one of the named *functions*, these studies can hardly be carried out in this way. The complexity of the object of research, viz. education and training leads to the combination of functions. The integration into the *multi-level analysis* recommended by *Schriewer* seems to be an appropriate way for it.

The estimation of *Arno*ve / *Altbach* / *Kelly* regarding the development of the interest in knowledge, the theories and paradigms of the *CE/VE* with the motivating force “change for the better” groups around the melioristic paradigm and can be used as a hypothesis for the development of the *VBBF* in combination with further theories and paradigms.

#### **4. Fields and types of comparative research studies**

The theoretical concepts and the methodical implementations are decisive for the relevance of the results of the research in the analyzed *comparative research studies* standardized according to their main focuses. The hypothesis of the *primacy of theory* and the leading issues of knowledge with respect to the methodical decisions was substantiated. Both the hermeneutic and the metric comparative methods are – if only implicitly in part – connected with fundamental leading issues of knowledge.

As concerns the fields and objects of the comparison, there was a differentiation in many respects in the course of the development of the *CE/VE*. While geographic/political units had a determining influence on the fields of the comparison at the beginning, defined units, such as family, regions, or other functional units, such as companies or associations were later turned to, often with reference to special issues within the framework of the *problem approach*. Mostly, these examples already distinguish themselves during the process of establishing and delimiting the fields and objects by the fact that theory and leading issues of knowledge are the determinants for the research design. These interrelations were particularly apparent during the dispute over hermeneutic or sociological paradigms in the 1960's and 1970's.

The turning to special geographic/political units or *problem approaches* finds its expression in the type of the research study. Representatives of the *problem approach* claim over and over again that regional studies (*total approach*) would be pure description following the itineraries. Our study on this does not confirm this hypothesis. The truth is rather that researchers who have taken up this assignment have incorporated the *problem approach* – though not using quantitative methods – in their ‘description’.

In order to bring out the philosophies, peculiarities and system structures that stand for a special national, regional or operational stamping of general and vocational education, the factors and development trends that led there have to be examined. This type of comparative study, the *development study*, was dominating up to the sixties.

### **5. Methodologies, methods and interest in knowledge**

Methodological questions cannot be separated from those of forming a theory or decision on paradigms. The hypothesis of the *primacy of theory* (1) and of the leading issues of knowledge (2) with respect to the methodical decisions (3) was confirmed. The choice of paradigms (4) must still be added to it. The method (3) is in the interdependent field of tension between (1), (2), and (4). They, in turn, have mutual influence on each other.

The assumption that the method is directly linked as a fourth step after the objective of research in combination with the interest in knowledge for a theory and the paradigm that influences everything, is often not compatible with the research reality. The study conducted by Rust i.a. (1999) or the assessment by Nohlen (1994) for the political science, likewise an integration science with open methodical concepts, reach the conclusion that researchers care little about theory- and paradigm-related principles in their studies and that they only sporadically 'divulge' any methodical concepts, paradigms, theories, and research interest. Nothing is mostly stated in this respect. It seems that there is only *little* interest in making theory- and methodology-oriented plans and that it's more popular to design the studies 'intuitively'.

On the other hand, these interdependences between (1)-(4) appear in their mutual dependence in a well-ordered – partly also sound – connection in the *methodologies* of Hilker, Bereday, Robinsohn, Holmes, Röhrs, and Garrido. The decisions on (1)-(4) are directly substantiated, or they may be indirectly gathered from the methodologies and from the additional practice-related examples. These *methodologies* express a special phase of development, a certain epoch of the *CE/VE*, Bereday for example stands for the *evolutionistic paradigm*, for the inductive method, the nomothesis and qualitative methods of the data collection in the framework of a *total analysis*.

*Robinsohn* is a typical representative of the *sociological paradigm* with pluralism of methods in the data collection, the intercultural comparison as *total analysis* or *problem approach*, with the intention of putting these results in defined correlations and regularities with regard to each other (integration theory) in order to thus gain knowledge of possible development trends and the educational planning.

## 6. VBBF as delimitation with regard to the CE/VE?

The hypothesis that the VBBF would have independent roots with regard to the CE/VE and that the basic development would be characterized by a separation with regard to the CE/VE, is mainly based on the assessments of Lipsmeier (1969a and Dörschel (1960/19754), and on the more recent studies conducted by Münch, Greinert or Georg, as well as on current comparisons of technical and vocational education and training.

The position of *Abel*, who repeatedly established connections between the *VBP* (*Vergleichende Berufspädagogik*) [Comparative vocational education] as core of the VBBF and the CE/VE, appeared to be a non-representative exception. On the other hand, there were indications that the recognized researcher in TVET Fischer and Kerschensteiner already at the beginning of the 20<sup>th</sup> century would do their research work by way of international comparisons without only referring to the technical and vocational education and training.

After the interregnum caused by the National Socialists in the German international CE/VE, this integral tradition of the CE/VE was resumed after the World War. A separation of the VBBF was not on the agenda. The results of the comparative research studies were rather used for supporting the reform efforts and for the reconstruction of the general and vocational education system in the western occupation zones, viz. in the Federal Republic of Germany. Especially in the vocational main field, idiographically and melioristically oriented comparative studies in the meaning of the theories set up by Kandel and Schneider were dominating with their results not only referring to the German situation but also being able to give ideas for changes in the technical and vocational education and training in the USA as was the case with the *Ware report*. Besides these research studies oriented towards the pedagogical theories of comparison, international comparisons with sociological, economic, and psychological backgrounds, were likewise made in the field of technical and vocational education and training without referring to it as a special discipline called VBBF.

An essential reason for giving up the debate about the CE/VE by the vocational educationists in the 60<sup>th</sup> of the 20<sup>th</sup> century seems to be the fact that the vocational education gained status, which became possible by the turning of the teacher training for vocational schools into a graduate profession, and that there was a delimitation with regard to the general education. In addition, the setting up of courses of studies that tied down capacities, the 'low' staffing level, and the hardly existing integration of elements of the CE/VE in the study schemes should be mentioned as possible reasons. It seems as if the CE/VE had been simply forgotten although Lipsmeier, a disciple of Abel, establishes relationships now and then but argues more in the sense of a stand-alone VBP in this connection. The connection, and the analysis of the CE/VE mostly accompanying it, was almost exclusively made by



representatives of the economic education, such as Abraham, Dörschel or Pukas, and in the current discussion by Frommberger, Deißinger, Justin, Schütte or Reinisch.

In the period between the seventies till the nineties of the 20th century, it was no longer the solid comparative analysis that dominated in the *VBP* along with the idiographically oriented studies of foreign systems of technical and vocational education and training but the forming of artifacts to systematize vocational education systems in combination with theories on typologies. This became the typical characteristic of the *VBBF* which saw itself more and more as something separated from the *CE/VE*. A pattern of explanation for this development is the tendency toward cutting oneself off and the devotion of one's attention to oneself, likewise with regard to the social sciences in general because typology is widespread there as well, however, as the result of empirical scientific studies and not of speculative formation of artifacts. Solid empirical comparisons met with less response.

An independent *VBBF* developed whose representatives frequently closed their minds to the development trends within the *CE/VE*, or did not register them, as opposed to Kerschensteiner, Fischer, Röhrs, Abel or Blankertz who concentrated their interest on the general and vocational education. Let us only refer to the integration of empirical methods, the intercultural comparison, the problem approach, or the multi-level analysis in this connection.

Not until the last few years, an approach toward the *State of the art* of the *CE/VE* has again been ‚coincidentally‘ taking place – as in the case of Georg – indirectly via the comparative cultural psychology and organizational sociology. This path is typical since scientists from other sociological disciplines are mostly active as comparative ‚vocational education‘ researchers due to the increased need for international comparative research within the technical and vocational education and training, e.g. within the EU or the development cooperation, and due to the low level of staffing of the ‚original‘ vocational educationists. In their original disciplines, comparative research studies are an assignment which is methodologically included in the *State of the art* of the corresponding discipline and consequently something normal to deal with.

The result will therefore not only be an extended interdisciplinary cooperation in international comparative projects but also ceasing of the state of halt in the *VBBF*, which is still partly characterized by the formation of artifact-like models, in favor of comparative research studies that combine the internal and the external view in the research interest. The international comparison will then represent a theoretical concept in the vocational educational research which is quite ordinarily likewise used for dealing with ‚internal‘, i.e. German problem areas. This would constitute an approach toward the international standard in the sociologically oriented research in TVET in which this field of research is not exclusively claimed by one sin-

gle discipline. The communication platforms for the documentation of the results of the research and for the promotion of the discourse among the researchers and with the educational practice are essential in view of supporting this process.

As concerns the question of the particular feature of the *VBBF*, this is not to be found in the paradigms, theories, methodologies, and methods of comparison, but in the concentration on the field of research, viz. technical and vocational education and training. The interest in knowledge of the comparative vocational education researchers is generated from the field of technical and vocational education and training, nevertheless on the basis of (vocational) education theories or special problem areas. Comparative research studies analyzing the equivalency of general and vocational education therefore still remain an essential research issue today, also on the level of the European Community. Apart from these (vocational) educationally oriented comparisons, which refer to *CE/VE* as reference discipline due to the pedagogical orientation, there are parallel comparative research studies in the field of technical and vocational education and training with sociological, psychological, or economic, etc. orientation, which are embedded in their corresponding disciplines. These disciplines are combined in the field of the *VBBF* in a multi- or interdisciplinary way.

## 7. Conclusions

The comment by Kelly / Altbach (1981, 1) on the worldwide state of the *CE/VE*, when the pieces were picked up, arranged, and patched up again after the *war of paradigms* in the 1970's – between the hermeneuts and metrists, between the supporters of the *total analysis* and of the *problem approach*, would be an excellent initial position for the necessary dispute over the contents of research studies in the *VBBF*, with the recollection of the roots in the *CE/VE* and less of the existing methods elaborated in the social sciences.

There is reason for scepticism since although the estimation of Kelly / Altbach also applies to the *VBBF*, the inevitable analysis of the *State of the art* was previously not implemented as a basis of the discourse for the *VBBF* in the same way as it was continuously done in the *CE/VE*. The few current attempts at the theoretical and positional discussion must first prove their continuity. Not until after that – as the experience in the *CE/VE* shows – will it be possible to start a debate about the interest in knowledge, the theories and methods, as well as about the fields of research, which has previously not taken place in the *VBBF* despite the attempts of *Czycholl*. These debates only took place in part within the general vocational and economic education (cf. Beck 1995, 457 ff.; Pätzold 1999a, 124 ff.) without having any considerable effect on the *VBBF*.

One result can nevertheless be recorded as certain after this analysis of the historical development of the *VBBF*, viz. that the supranational historico-political contexts were and still are



decisive for the development of the research work and of the discipline. At the moment, the factors increasing the pressure for changes on the national systems of technical and vocational education and training are the political and economic amalgamations, the technical progress, and the social change.

There is a field of research with a vast variety of assignments that should in many cases be handled in interdisciplinary mode for them to be successful due to the mostly complex structures within a research project. This internal view of the *Berufsbildungsforschung – BBF* – [Research in TVET], that means the concentration on the national perspective and its analysis, will be ever increasingly linked with the external view of the *Vergleichende Berufsbildungsforschung – VBBF* – [Comparative research in TVET] in other countries where similar problem situations exist or where the melioristic gain of knowledge comes more to the fore. The international comparison, the intercultural comparison, and the like are available for this purpose as special methodical instruments. The results of international comparisons and the like likewise support the national interest in knowledge with the external-internal view. It helps to transfer knowledge to the different national levels (macro to micro) from the international and supranational point of view. It is obvious that the working of this field of research calls for difficult and complicated scientific processes by the *leading discipline of vocational and economic education* with the main focuses on *BBF* and *VBBF*. Due to the extraordinarily favorable general conditions prevailing at present for the *VBBF*, it should be assumed, apart from the amount of scepticism, that the adoption of the sociological comparative educational research will make it possible to succeed in achieving a more in-depth mode of working.

**For the monitoring of national TVET systems this fundamental remarks mean:**

- The separation between CE and comparative research in TVET (VBBF) is a traditional German phenomena. The named reasons speak against the separation of VBBF. The researchers in TVET should search the co-operation with researchers in education and CE. In this case there is a participation of this researches on the state of the art of the world wide research in education.
- The monitoring of national TVET systems is a combination of the internal and external view of the researches. Ideal is the combination of the co-operation of a native and a foreign researcher.
- The problem approach is valid if integrated in the context of the current education system (included the TVET system). There for look for the combination of qualitative and quantitative analysis.
- The implicit comparison (often called as secretly comparison = the own national education system is the scale) when monitoring foreign TVET systems is a permanent

challenge. To avoid this the whole dimension of the complex object of education systems is to integrate in the corresponding context. There for today researchers use the *empirico-analytical approach* in combination with a humanistic-hermeneutic dimension (*pragmatic approach*) as *multi-level analysis* with the context of the society, the economic and the historic dimension.

- For the monitoring of TVET this means great openness. In each cultural scope the understanding of TVET is different. As a demonstration of the difference the comparison of the German “Berufsbildung” with the English or American terms “VET” or “TVET” is an example. In general and cross national the fixing of the disassociation between TVET and general education or between TVET and higher education is not possible. There for the monitoring of national TVET systems or of special aspects is to integrate in the national, regional, etc. context of the society, economy etc.

## References

*Abel, Heinrich* (1962) Die Bedeutung der vergleichenden Erziehungswissenschaft für die Berufspädagogik. In: Die Deutsche Berufs- und Fachschule, 58 (1962) 4, S. 241-253.

*Abel, Heinrich* (1963) Vergleichende Berufspädagogik. In: Politik und Bildung. Hrsg. vom Internationalen Arbeitskreis Sonnenberg. Braunschweig 1963, S. 208-216.

*Abel, Heinrich u.a. (Bremer, Wilhelm / Hobbensiefken, Günter / Laube, Horst / Greinert, Wolf-Dietrich)* (1965) Vorbericht über das Forschungsvorhaben über den Einsatz Programmierter Instruktion in der Berufsbildung in den Vereinigten Staaten. Darmstadt: Lehrstuhl für Berufs-, Wirtschafts- u. Arbeitspädagogik a. d. TH Darmstadt. S. 95. + Anhang (nicht paginiert).

*Abraham, Karl* (1962) Die Aufgaben der vergleichenden internationalen Wirtschaftspädagogik. In: Zeitschrift für Pädagogik, 8 (1962) 2, S. 166-178.

*Archer, Margaret S.* (1979) Social origins of educational systems. London: Sage. S. 816

*Arnone, Robert F. / Altbach, Philip G. / Kelly, Gail P. (Ed.)* (1992) Emergent Issues in Education. Comparative Perspectives. Albany, N. Y.: State University of New York Press. XI, S. 363 (Frontiers in Education)

*Baumann, Ulrich* (1981) Historik und Komparatistik – Parallelen zwischen einer historischen und einer komparativen Sozialwissenschaft. In: *Baumann, U. u.a. (Hrsg.)* Vergleichende Erziehungswissenschaft. Festschrift für Hermann Röhrs zum 65. Geburtstag. Wiesbaden: Akademische Verlagsgesellschaft, S. 3-11.

*Beck, Klaus* (1995) Theorieansätze. In: *Arnold, R. / Lipsmeier, A. (Hrsg.)* Handbuch der Berufsbildung. Opladen: Leske und Budrich, S.457-464.

*Bereday, George Zygmunt Fijalkowski* (1964) *Comparative Method in Education*. New York u.a.: Holt, Rinehart and Winston. XVI, S. 302

*Berufsbildungsforschung an den Hochschulen der Bundesrepublik Deutschland* (1990) Situation, Hauptaufgabe, Förderungsbedarf. Hrsg. von der Senatskommission für Berufsbildungsforschung der Deutschen Forschungsgemeinschaft. Weinheim u.a.: VCH Acta Humaniora. S. 125

*Blättner, Fritz / Münch, Joachim* (1965) *Pädagogik der Berufsschule*. Heidelberg: Quelle & Meyer. S. 201

*Bray, Mark / Thomas, Robert Murray* (1995) Levels of Comparison in Educational Studies. Different Insights from Different Literatures and the Value of Multilevel Analyses. In: *Harvard Educational Review*, 65 (1995) 3, S. 472-490.

*Busch, Adelheid / Busch, Friedrich Wilhelm / Krüger, Bernd / Krüger-Potratz, Marianne* (Hrsg.) (1974) *Vergleichende Erziehungswissenschaft: Texte zur Methodologie – Diskussion*. Pullach bei München: Verlag Dokumentation. S. 309. (Uni-Taschenbücher, 410)

*CEDEFOP / DIPF / Lauterbach, Uwe / Sellin, Burkart* (Hrsg.) (2001) *Vergleichende Berufsbildungsforschung in Europa, Ansätze, Politikbezüge und Innovationstransfer*. (Beiträge, Empfehlungen und Weiterführung der CEDEFOP/DIPF-Konferenz vom Januar 1998 im Wissenschaftszentrum in Bonn). Frankfurt a.M., Thessaloniki (<http://www.cedefop.eu.int/>). S. 492

*Czycholl, Reinhard* (1971) *Vergleichende Wirtschaftspädagogik. Möglichkeiten und Grenzen einer vergleichenden erziehungswissenschaftlichen Disziplin*. Trier: Spee. S. 295

*Czycholl, Reinhard* (1975) Vom Nutzen einer vergleichenden Wirtschafts- und Berufspädagogik, in: *Die Deutsche Berufs- und Fachschule*, 71 (1975) 1, S. 3-16.

Deißinger, Thomas → Schütte, Friedhelm

*Dieckmann, Bernhard* (1970) *Zur Strategie des systematischen internationalen Vergleichs*. Stuttgart: Klett. 192 S. (Texte und Dokumente zur Bildungsforschung)

*Dörschel, Alfons* (1960 / 1975<sup>4</sup>) *Einführung in die Wirtschaftspädagogik*. München: Vahlen. S. 301 (Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften)

*Epstein, Erwin H.* (1983) Currents Left and Right. Ideology in Comparative Education. In *Comparative Education Review*, 27 (1983) 1, S. 3-29.

*Epstein, Erwin H.* (1988) The problem meaning of 'comparison' in comparative education. In: *Schriewer, Jürgen / Holmes, Brian* (Ed.) *Theories and Methods in Comparative Education*. Frankfurt a.M. u.a.: Lang 1988. X, S. 3-23.

*Epstein, Erwin H.* (1994) Comparative and International Education: Overview and Historical Developments. In: *Husén, T. / Postlethwaite, T. N. (Ed.)* The International Encyclopedia of Education. Oxford u.a.: Pergamon, S. 918-923.

*Fischer, Aloys* (1925/1950) Die kulturelle Grundlage der Erziehung. In: *Kreitmeir, Karl (Hrsg.)* Aloys Fischer Leben und Werk. Bd. 2. München 1950, S. 73-133.

*Flitner, Wilhelm* (1931/32) Die Reformpädagogik und ihre internationalen Beziehungen. In: Internationale Zeitschrift für Erziehungswissenschaft, 1 (1931/32) 1, S. 39-57.

*Frommberger, Dietmar / Reinisch, Holger* (1999) Ordnungsschemata zur Kennzeichnung und zum Vergleich von „Berufsbildungssystemen“ in deutschsprachigen Beiträgen zur international vergleichenden Berufsbildungsforschung: Methodologische Fragen und Reflexionen. In: ZBW, 95 (1999) 3, S. 323-343.

*Georg, Walter* (1995) Probleme vergleichender Berufsbildungsforschung im Kontext neuer Produktionskonzepte: Das Beispiel Japan. In: *Dybowski, G. u.a. (Hrsg.)* Berufsbildung und Organisationsentwicklung. Perspektiven, Modelle, Forschungsfragen. Bremen: Donath, S. 67-84.

*Georg, Walter* (1997) Zwischen Tradition und Moderne: Berufsbildung im internationalen Vergleich. In: *Arnold, R. / Dobischat, R. / Ott, B. (Hrsg.)* Weiterungen der Berufspädagogik von der Berufsbildungstheorie zur Internationalen Berufsbildung. Festschrift für Antonius Lipsmeier. Stuttgart: Steiner, S. 153-166.

*Georg, Walter* (2001) Beruf und Berufsbildung in der Entwicklungszusammenarbeit. In: *Koch, T. / Neumann, G. / Stach, M. (Hrsg.)* Aspekte Internationaler Berufspädagogik. Von der Qualitätssicherung über Länderberichte zur Methodologie der Vergleichsforschung. Festschrift für Eberhard Schoenfeldt. Universität Gesamthochschule Kassel, S. 97-112.

*Glowka, Detlef* (1997) Überlegungen zu einer „Einführung in die Vergleichende Erziehungswissenschaft“. In: *Vergleichende Erziehungswissenschaft*. Herausforderung – Vermittlung – Praxis; Festschrift für Wolfgang Mitter zum 70. Geburtstag, hrsg. v. Ch. Kodron, B. v. Kopp, U. Lauterbach, U. Schäfer, G. Schmidt. Köln: Böhlau, S. 52-64.

*Gonon, Philipp* (1998) Das internationale Argument in der Bildungsreform. Die Rolle internationaler Bezüge in den bildungspolitischen Debatten zur schweizerischen Berufsbildung und zur englischen Reform der Sekundarstufe II. Bern u.a.: Lang. S. 584 (Explorationen)

*Greinert, Wolf-Dietrich* (1995) Die Wiederentdeckung der Tradition als Regelungsmuster moderner Berufsausbildung. In: Berufsbildung, 49 (1995) 33, S. 32-34.

*Grüner, Gustav* (1967) Berufsbildung in Stufen. In: Berufliche Bildung in Europa. Viertes bildungspolitisches Gespräch, Dortmund 18. u. 19. November 1966. Bonn, S. 5-24. (Beiträge zur Europakunde, 5)

*Grüner, Gustav* (1972) Die Typen der Berufsausbildung. Ein Beitrag zur Bestimmung der zentralen Aufgaben der Berufsbildungsforschung. In: *bbs*, 24 (1972) 1, S. 15-22.

*Gudykunst, William B.* (1991) Bridging Differences. Effective Intergroup Communication. Sage: Newsbury Park u.a. S. 171.

*Halls, Wilfred Douglas* (1990) Trends and issues in comparative education. In: *Halls, W. D. (ed.) Comparative Education. Contemporary Issues and Trends.* London u.a.: Kingsley u.a., S. 21-65.

*Hans, Nicholas A.* (1951<sup>2 nd</sup> ed) Comparative Education. A Study of Educational Factors and Traditions. London: Routledge & Kegan Paul 1951. X, 333 S.

*Hilker, Franz* (1962) Vergleichende Pädagogik. Eine Einführung in ihre Geschichte, Theorie und Praxis. München: Hueber 1962. 184 S.

*Hilker, Franz* (1963) Was kann die vergleichende Methode in der Pädagogik leisten? In: *Bildung und Erziehung*, 16 (1963) 9, S. 511-526.

*Holmes, Brian* (1971) Comparative Education. In: *Deighton, Lee C.* The Encyclopedia of Education. Vol. 2. New York: Macmillan, S.357-363.

*Holmes, Brian* (1981) Comparative Education. Some Considerations of Method. London u.a.: Allen and Unwin 1981. 195 S. (Unwin Education Books)

*Hörner, Wolfgang* (1993) Technische Bildung und Schule. Eine Problemanalyse im internationalen Vergleich. Köln u.a.: Böhlau. 384 S. (Studien und Dokumentationen zur vergleichenden Bildungsforschung, 52)

*Husén, Torsten* (1994) Research Paradigms in Education. In: *Husén, T. / Postlethwaite, T. N. (Ed.) The International Encyclopedia of Education.* Oxford u.a.: Pergamon, S.5.051-5.056.

*Kandel, Isaac Leon* (1933) Comparative Education. Boston/Mass. u.a.: Houghton Mifflin 1933. XXVI, 922S. (Reprint: London: Harrap 1939. XXVI, 922S.)

*Kandel, Isaac Leon (ed.)* (1930) Educational Yearbook of the International Institute of Teachers College, Columbia University 1929. New York: Bureau of Publications, Teachers College, Columbia University. 560 S.

*Keeves, John P. / Adams, Don* (1994) Comparative Methodology in Education. In: *Husén, T. / Postlethwaite, T. N. (Ed.) The International Encyclopedia of Education.* Oxford u.a.: Pergamon, S.948-953.

*Kelly, Gail P. / Altbach, Philip G.* (1981) Comparative Education: A Field in Transition. In: *Altbach, Ph. G. / Kelly, G. P. / Kelly, D. H.* International Bibliography of Comparative Education. New York: Praeger, S. 1-27.

*Kerschensteiner, Georg* (1901) Beobachtungen und Vergleiche über Einrichtungen für gewerbliche Erziehung außerhalb Bayerns. München: Carl Gerber. XI, S. 245

*Kerschensteiner, Georg* (1931) Der Zeitschrift zum Geleit. In *Internationale Zeitschrift für Erziehungswissenschaft*, 1(1931/1932)1, S. 5-6.

*Lauterbach, Uwe* (1999) Theorien und Methodologien der Vergleichenden Bildungsforschung und der Vergleichenden Berufsbildungsforschung. In: *Lauterbach, Uwe (Hrsg.)* (2001 ff.) Internationales Handbuch der Berufsbildung. Baden-Baden: Nomos, S. VGL 133-240.

*Lauterbach, Uwe* (2003) Vergleichende Berufsbildungsforschung. Theorien, Methodologien und Ertrag am Beispiel der Vergleichenden Berufs- und Wirtschaftspädagogik mit Bezug auf die korrespondierende Disziplin Comparative Education / Vergleichende Erziehungswissenschaft. Nomos: Baden-Baden 432 S. (Bildung und Arbeitswelt, 8)

*Lauterbach, Uwe / Mitter, Wolfgang* (1998) Theory and Methodology of International Comparisons. In: *CEDEFOP (Ed.) Vocational Education and Training – the European Research Field. Background Report. Vol. 2.* Luxembourg: Office for Official Publications of the European Communities, S.235-271.

*Lipsmeier, Antonius* (1969) Technik, allgemeine Pädagogik und Berufspädagogik im 19. Jahrhundert. Ein Beitrag zur Geschichte der vergleichenden Berufspädagogik. In: *Technikgeschichte*, 36 (1969) 2, S. 133-146.

*Lipsmeier, Antonius* (1985) Heinrich Abel – ein Pionier der vergleichenden Berufspädagogik. In: *ZBW*, 81 (1985) 8, S. 719-727.

*Lipsmeier, Antonius* (1997) Herausforderungen an die Berufsbildungsforschung im Prozeß der europäischen Integration. In: *Gronwald, D. / Hoppe, M. / Rauner, F. (Hrsg.)* 10 Jahre ITB. Festveranstaltung u. Berufsbildungskonferenz 21.-23. Februar 1997. Bremen: Donat, S. 50-74.

*Lipsmeier, Antonius / Nölker, Helmut / Schoenfeldt, Eberhard* (1975) Berufspädagogik. Eine Einführung in die bildungspolitische und berufspädagogische Situation und Diskussion von Berufsausbildung und Gesellschaft. Stuttgart u.a.: Kohlhammer. 170 S.

*Loxley, William* (1994) Comparative and International Education: Organizations and Institutions. In: *Husén, T. / Postlethwaite, T. N. (Ed.) The International Encyclopedia of Education.* Oxford u.a.: Pergamon, S. 933-942.

*Masemann, Vandra L.* (1994) Comparative Education Societies. In: *Husén, T. / Postlethwaite, T. N. (Ed.) The International Encyclopedia of Education.* Oxford u.a.: Pergamon, S. 942-948.

*Maurice, Marc* (1991) Methodische Aspekte internationaler Vergleiche. Zum Ansatz des gesellschaftlichen Effekts. In: *Heidenreich, M. / Schmidt, G. (Hrsg.) International vergleichende Organisationsforschung.* Opladen: Westdeutscher Verlag, S. 82-90.



*Mitter, Wolfgang* (1997) Pädagogik, vergleichende. In: *Lenzen, D. (Hrsg.)* Pädagogische Grundbegriffe. Reinbek: Rowohlt, S. 1.246-1.260.

*Münch, Joachim* (1999) Qualifikation als Standortfaktor – Deutschland, USA und Japan im Vergleich. Hochheim a.M.: Neres. S 197.

*Noah, Harold J. / Eckstein, Max A.* (1969) Towards a science of comparative education. New York: Macmillan S. 1969. 222

*Nohlen, Dieter* (1994) Vergleichende Methode. In: *Kriz, J. / Nohlen, D. / Schultze, R.-O. (Hrsg.)* Lexikon der Politik. Hrsg. v. Dieter Nohlen. Bd. 2. München: Beck. S. 507-523.

Nölker, Helmut (1975) → Lipsmeier, Antonius u.a. (1975)

*Pätzold, Günter* (1999) Berufspädagogik. In: *Kaiser, F.-J. / Pätzold, G. (1999)* Wörterbuch der Berufs- und Wirtschaftspädagogik. Bad Heilbrunn: Klinkhardt, S. 124-126.

*Paulston, Rolland Glenn* (1994) Comparative and International Education: Paradigms and Theories. In: *Husén, T. / Postlethwaite, T. N. (Ed.)* The International Encyclopedia of Education. Oxford u.a.: Pergamon, S. 923-933.

*Robinsohn, Saul Benjamin* (1970) Erziehungswissenschaft: Vergleichende Erziehungswissenschaft. In: *Speck, J. / Wehle, G. (Hrsg.)* Handbuch pädagogischer Grundbegriffe. Bd. 1, München: Kösel, S. 456-492.

*Röhrs, Hermann* (1967<sup>2</sup>) Die Pädagogik Aloys Fischers. Heidelberg: Quelle & Meyer. S.252

*Röhrs, Hermann* (1995) Die Vergleichende und Internationale Erziehungswissenschaft. Weinheim: Deutscher Studien Verlag. S. 480. (Gesammelte Schriften, 3)

*Rust, Val D. / Soumaré, Aminata / Pescador, Octavio / Shibuya, Megumi* (1999) Research Strategies in Comparative Education. In: *Comparative Education Review*, 43 (1999) 1, S. 86-109.

*Schmidt, Hermann* (1995) Berufsbildungsforschung. In: *Arnold, R., Lipsmeier, A. (Hrsg.)* Handbuch der Berufsbildung. Opladen, S. 482-491.

*Schneider, Friedrich* (1931/33) Internationale Pädagogik, Auslandspädagogik und Vergleichende Erziehungswissenschaft. Geschichte, Wesen, Methoden, Aufgaben und Ergebnisse. In: *Internationale Zeitschrift für Erziehungswissenschaft*, 1 (1931/32) 1-4, S.15-39, S. 243-257, S.392-407, 2(1932/33)1, S. 79-89.

*Schneider, Friedrich* (1943) Geltung und Einfluß der deutschen Pädagogik im Ausland. München u.a.: Oldenbourg. S. 358

*Schneider, Friedrich* (1961) Vergleichende Erziehungswissenschaft. Geschichte, Forschung, Lehre. Heidelberg: Quelle & Meyer. 218 S. (Vergleichende Erziehungswissenschaft und Pädagogik des Auslands, 1)

*Schoenfeldt, Eberhard* (1996) Der Edele ist kein Instrument. Bildung und Ausbildung in Korea (Republik). Studien zu einem Land zwischen China und Japan. Universität Gesamthochschule Kassel. Institut für Berufsbildung. Kassel: Universitätsbibliothek 1996. S. 383 (Berufs- u. Wirtschaftspädagogik, 22)

*Schriewer, Jürgen* (1982) „Erziehung“ und „Kultur“. Zur Theorie und Methodik Vergleichender Erziehungswissenschaft. In: *Brinkmann, Wilhelm / Renner, Karl (Hrsg.)* Die Pädagogik und ihre Bereiche. Paderborn u.a.: Schöningh, S. 185-236.

*Schriewer, Jürgen (Ed.)* (2000) Discourse Formation in Comparative Education. Frankfurt a.M. u.a.: Lang. X, S. 366 (Komparatistische Bibliothek, 10)

*Schütte, Friedhelm / Deißinger, Thomas* (2000) ‚Bildung‘ und ‚Arbeit‘ im internationalen Diskurs. Zur Aktualität und Relevanz historisch-vergleichender Forschung in der Berufs- und Wirtschaftspädagogik. In: ZBW, 96 (2000) 4, S. 540-555.

*Stratmann, Karlwilhelm* (1979) Berufs- und Wirtschaftspädagogik. In: *Groothoff, Hans-Hermann (Hrsg.)* Die Handlungs- und Forschungsfelder der Pädagogik. Königstein/Ts.: Athenäum. 337 S. (Erziehungswissenschaftliches Handbuch, 5) S. 285-337.

*Tenorth, Heinz-Elmar* (1996) Das Zentralinstitut für Erziehung und Unterricht. In: *Geißler, G. / Wiegmann, U. (Hrsg.)* Außeruniversitäre Erziehungswissenschaft in Deutschland. Köln u.a.: Böhlau, S. 113-135.

*Trommsdorff, Gisela* (1989) Vergleich, interkultureller, intersozietärer. In: Wörterbuch der Soziologie. Hrsg. von Günter Endruweit u. Gisela Trommsdorff. Bd. 3. Stuttgart: Enke, S. 774-777.

*Ware, Georg W.* (1952) Berufserziehung und Lehrlingsausbildung in Deutschland. Hrsg. v. Amt des Hohen Kommissars für Deutschland. Amt für öffentliche Angelegenheit – Kulturabteilung. Frankfurt a.M. Februar 1952. S. 56.

*Whiting, John W.* (1961) The Cross-culture Method. In: *Moore, Frank, W. (Ed.)* (1961) Readings in Cross-Cultural Methodology. New Haven: HRAF Press 1961, S.287-309.

*Wilhelm, Theodor* (1943) Kulturpolitische Stützpunkte des USA-Imperialismus. In: Internationale Zeitschrift für Erziehungswissenschaft 12 (1943) 1, S. 1-23.

*Wilson, David N.* (1994) Technical-Vocational Education: Comparative and International Studies. In: *Husén, T. / Postlethwaite, T. N. (Ed.)* (1994) The International Encyclopedia of Education. Oxford u.a.: Pergamon, S. 6.261-6.266.



# International Handbook on VET

## Presentation of a draft modified monitoring grid<sup>15</sup>

*Prof. Dr. Ute Clement, Dr. Philipp Grollmann, Prof. Dr. Matthias Pilz*

**Abstract:** This is the presentation of considerations of the editorial board of the German language publication: “Internationales Handbuch der Berufsbildung” on updating the monitoring grid for country studies within the handbook. The aim of the handbook is being a practical, up-to-date source of information for experts in the field of technical and vocational education and training that is based on an up-to-date scientific approach to the analysis of Vet systems.

**Keywords:** Monitoring; VET systems

### Introduction

This is the presentation of considerations of the editorial board of the German language publication: “Internationales Handbuch der Berufsbildung (IHBB)” [International Handbook of Vocational Education] on updating the monitoring grid for country studies within the handbook. The IHBB is a research based resource that contains descriptive country studies of national VET systems.

The International Handbook of Vocational Education and Training provides information about the various avenues of initial and continuing vocational training in the majority of European states and the major industrialised countries and newly industrialised countries outside Europe. It analyses the structures and puts them in the context of the respective political and economic system in each country whilst it also brings in the relevant international aspects.

The handbook aims at being a practical, up-to-date source of information for experts in the field of education and vocational training, teachers, trainers, students, companies, chambers of trade and commerce, professional associations, trade unions, government agencies, international organisations, non-profit foundations and politics. The country studies are added, updated or renewed on a regular basis. The format of the book is a loose-leaf binder. In the last 15 years the book has developed towards a standard resource to the VET community and has increased its profile over the years.

The country-specific studies were developed as a part of a research project since 1992 by members of the academic staff of the German Institute for International Educational Re-

---

<sup>15</sup> This paper was developed as part of ongoing discussions in and on behalf of the editorial board of the „Internationales Handbuch für Berufsbildung“. Beside the authors the following individuals have contributed to this as editors: Dietmar Frommberger and Uwe Lauterbach.

search (DIPF) with the collaboration of over 80 academic researchers and specialists on vocational education from Germany and abroad. Initially, reports based on the direct experience of German training specialists who have participated in study trips abroad organised by the Carl Duisberg Gesellschaft were taken as a basis. The research project was funded since 1992 by the Federal Ministry of Research, Technology, Education and Science (BMBF). In the meantime it has become independent of this funding and from 2010 the BIBB will take over the responsibility for the management together with Otto-von-Guericke Universität Magdeburg (Philipp Grollmann and Dietmar Frommberger).

Apart from country studies there are also some introductory contributions of a more analytical nature, such as a chapter on the state of the art of comparative vocational education and a chapter that draws major lines of comparison between the different countries that the volume covers. In addition an overview is given on policies and practices of vocational education and training of International organisations such as the EU.

The following countries have been subject to country studies in the IHBB.

**Table 1: Country Studies in the International Handbook**

Chapters of the IHBB
Albania, Argentina, Austria, Australia, Belarus, Belgium, Brazil (revised version + supplement), Bulgaria, Canada (+ supplement 2003), Chile, China (People Republic) Czech Republic (revised version 1998), Cyprus, Denmark (+ supplement 2000), Finland, France, Greece (+ supplement 2000), Great Britain: England/Wales (revised version 2000), Great Britain/Scotland, Hungary (revised version), Ireland, Israel, Italy, Japan, Republic of Korea, Luxembourg, Netherlands, Norway, Poland, Portugal, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Ukraine, Uruguay, Vietnam.
In processing: Germany, Estonia, Finland (revised version), India, Lithuania, Malaysia, Poland (revised version), Romania, Slovenia (revised version), South Africa, Switzerland (revised version), Tunisia, and Turkey.

### Targeted Problem

In line with taking on new members to the editorial board of the handbook as well as with changes in its institutional setup, a discussion started in the editorial board about the conceptual make-up of the handbook and its country studies.

There were three major issues that led the discussion of this. First of all, the handbook should maintain its high reputation and scientific quality. Therefore, it should be looked at

recent developments and changes of approaches to monitoring VET systems. In addition the needs of major audiences should be reflected as well as the breadth of groups that make use of the handbook.

### *User Groups*

**Table 2: Target Groups of the International Handbook**

Users/Target Groups and use cases
University/college: assignments, presentations, university libraries
Associations: mobility consulting
Ministries: mobility consulting, guests (background)
Development cooperation: development aid worker, consulting
Research: EU-projects, “thought pattern”, theory formation
Companies (big): settlement in foreign countries, global players
Private education institutions: e.g. planification of EU-projects
Partner and researcher in foreign countries: structure that is compatible in other countries
Science, vocational education research

First of all, the range of user groups and “use cases” of the handbook were defined. Groups of users that were identified can be seen in Table 1.

### *Expectations*

In a second step, some exploratory interviews were carried out with industry representatives in order to find out about the specific user interests that they have when taking the handbook to consultation. A first finding is that the visibility of the handbook needs to be increased. Whilst many big enterprises in Germany have subscribed to the Handbook and receive the regular updates, those who are taking HRD and staff decisions for the companies in other seem not be aware of the handbook. The interview with users from the industry environment has led to the following list of expectations and questions on which answers are sought for:

- What is implied by the certificates/degrees, who can I employ?
- Which possibilities exist, in the respective context, to realize the dual or “German” training forms?
- Which contacts can be found in the respective country in the field of vocational education and training?

- Which persons and institutions exist concerning the questions of concrete implementation and realization of vocational training, for instance, the contacts to learning institutions?
- In addition, the handbook country studies were characterized as “too long”- so focusing would be appreciated.

### *Quality*

What are the signposts of quality of the handbook? The considerations of the handbook editors led to the following pillars that are partly realised and partly might need revision in the future. Each handbook study is supervised through the editors and written jointly between experts from the respective country and a usually a German language VET expert (balance between outside and inside perspectives, bilateral reflection). Chapters are unified through a common structure while highlighting country specific feature at the same time. Interrelationships between different items in the structure are explained on the basis of country-specific features. In terms of content the handbook covered a wide selection of countries including countries for which only few English or German language sources are available.

The handbook combines information on the following subjects with an analytical view based on the question on their conditions, causes and pathways of development:

- Curricula, structure of vocational education with regards to contents;
- Methods, pedagogy, teaching- and training-practice;
- Individuals' career development, how does the acquisition of vocational competencies function throughout the career development,

It seeks to provide information on the level of vocational education practice with contextual explanations. This is another feature of the handbook as opposed to other available informational sources.

### **Methodology**

It is the aim of the editors to further explore and analyse the interests of the different stakeholder groups of the IHBB. At the same a first draft reporting grid has been developed, taking into account the abovementioned findings. The most challenging issue will be to approach the national reporting not from the systems' level anymore but rather from a perspective that leans towards looking at typical ways of individual competence development, career paths and learning processes.

In Abstract the table shows how this would be manifested in a new structure of country studies.

**Table 3: Structure of country studies in the International Handbook**

Country study structure as is	Draft future structure
<ul style="list-style-type: none"> <li>• Basic data (Grunddaten)</li> <li>• Abbreviation (Abkürzung)</li> <li>• Preface/Introduction (Einleitung)</li> <li>• Introduction to the geographic, social, political and economic frameworks (Einführung in die geografischen, gesellschaftlichen, politischen und ökonomischen Rahmenbedingungen)</li> <li>• Responsibilities and agencies in education (including further education) – as well as training-systems (Zuständigkeiten und Träger im Bildungs-, Ausbildungs-, und Weiterbildungswesen)</li> <li>• Overview of the education system (Übersicht über das Bildungswesen)</li> <li>• Vocational education system (Berufliches Bildungswesen)</li> <li>• Further education and vocational education (Weiterbildung und berufliche Weiterbildung)</li> <li>• Personnel in vocational education system (Personal im beruflichen Bildungswesen)</li> <li>• Transnational mobility, international cooperation in vocational education (Länderübergreifende Mobilität, internationale Berufsbildungszusammenarbeit)</li> <li>• Abstract (Zusammenfassung)</li> </ul>	<ul style="list-style-type: none"> <li>• Executive Abstract (Zusammenfassung)</li> <li>• Introduction (histographic etc.) Einleitung (Histografisches etc.)</li> <li>• Individual (Individuum); Education progressions, influences on the education process, the beginning of the 1st threshold, transition data, biography, education participation, branches, transitions (Bildungsverläufe, Einflüsse auf Bildungsverläufe, Beginn an der 1. Schwelle; Übergangsdaten, Biographien, Bildungsbeteiligung, Einmündungen); Examples (Beispiel)</li> <li>• Learning in public and private vocational education institutions (Lernen in öffentlichen und privaten Berufsbildungsinstitutionen) Infrastructure/methods, pedagogy, curricula, exams, etc., personnel training (Rahmenbedingungen/Methodik, Didaktik, Curricula, Prüfungen etc.; Ausbildung des Personals)</li> <li>• Examples (Beispiel)</li> <li>• Learning inside the company (Lernen im Betrieb) Company recruitment strategy/methods, pedagogy, curricula, exams, etc. trainer training (Betriebliche Rekrutierungsstrategien/Rahmenbedingungen/Methodik, Didaktik, Curricula, Prüfen etc.; Ausbildung der Ausbilder) Examples</li> </ul>

<ul style="list-style-type: none"> <li>• References (Literatur)</li> <li>• Documents, legal foundations, addresses etc. (Dokumente, Rechtsgrundlagen, Anschriften u.Ä.)</li> </ul>	<ul style="list-style-type: none"> <li>• (Beispiel)</li> <li>• Framework (Rahmen) Laws (Gesetzliche Verortung) Standards for teacher education (Standards für Lehrerbildung) Modernization discourse (Modernisierungsdiskurse)</li> </ul>
--	---

## Results and Perspectives for further development

*For the described instrument/approach as such*

First of all, it we will have in mind how far certain items from the old structure can be transferred to the new concept. In a second step, the new approach is piloted with two countries. The countries, Japan and France; represent extreme or almost ideal types of the societal organisation of vocational education. It will be particularly interesting to see if a new structure is feasible or not with such paradigmatic cases of countries.

*Potentials and challenges for the use in/for international comparisons*

After this piloting, it is envisaged to discuss the new approach once again with representatives of VET research and from the different user groups in order to make final adaptations. As a next step, it would be worthwhile looking at the transferability to an internationalised structure and an English language version. But this is still some way to go.

## References

Lauterbach, U. u. a. (Hrsg.): Internationales Handbuch der Berufsbildung (IHBB). Bielefeld: Bertelsmann 2006.

## Indicators & Benchmarks

# Entry-Level-Competence Tests: A Means for Evaluation the Fulfilment of Occupational Standards

*Prof. Dr. Marc Schütte, Prof. Dr. Georg Spöttl*

**Abstract:** Companies recruiting new staff are relying on assessments rather than on tests. In Oman, the Ministry of Labour seeks to get to know the quality of employees who are starting their careers. This is the target group for entry-level-tests based on a work-process oriented approach. These tests allow determining whether the new employees are adequately qualified for their future tasks.

**Keywords:** Advanced occupational standard, entry-level competence, item generation, work-process, item field test, test use specifications

## Introduction

Within the framework of the development of “Occupational Standards and Skill Testing Centres (OSSTC)” in the Sultanate of Oman, the authors were entrusted with the development of entry-level-competence tests. These tests should first and foremost serve to determine which competencies the foreign workforce can bring into the country. A great number of Asian nations send workforce for technical occupations to the Sultanate of Oman. These employees then take over maintenance and repair tasks in different technical disciplines or are working in the field of tourism, banking and others.

As curricular oriented tests are not suitable for the determination of competencies in this target group, a work oriented test should be designed which should focus on the occupational challenges. The design of a test to be applied in this cultural environment has an additional advantage: specific cultural features need not to be considered.

Based on e.g. a reference system for the test design, the occupational standards so-far developed for around 50 occupations were selected. They consist of three elements, i.e.:

- The core work-process
- The core competencies and
- The detailing of the core work-process (Spöttl 2007).

Above all the core competencies are formulated as context oriented standards and therefore offer excellent prerequisites for the development of text items. The detailing yields further information which helps to include work process coherences in the text items.



## Entry-level competence tests

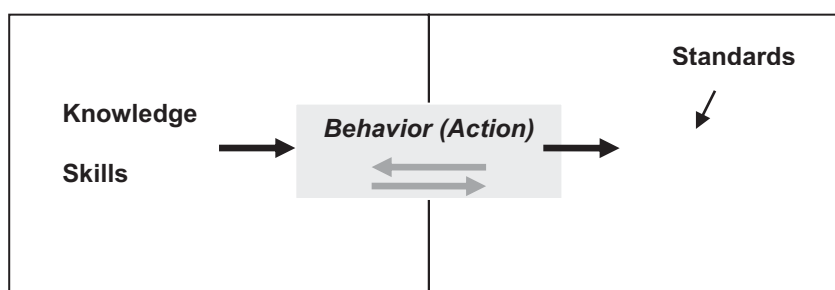
### Test purpose

A preparatory step of test development is to create an explicit statement of the test purpose (cf. Clauser & Case, 2006; Schmeiser & Welch, 2006). The test purpose statement specifies the object of measurement (quality or construct) and the usage of test scores (inferences).

In short, OSSTC tests are designed to measure whether a test taker has minimal competence that is necessary for skilled work at entry level of companies. The concept of “minimal competence” needs further explanation. Our explanation makes use of two differences.

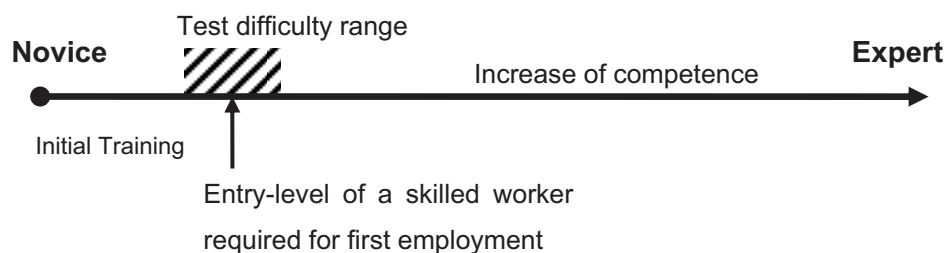
First, there is a difference between “competence” and “competencies”. The term “competence” is referring to the ability of a person to deliver outputs, i.e. consequences of action, which fulfil criteria for performance (given explicitly or implicitly). One of the areas where outputs are delivered are occupational fields where qualified staff has to master work-processes and must cover all relevant tasks. For example, “maintaining vehicle safety in terms of road-worthy operations and functions” on the one hand demonstrates a requirement in the field “car service and repair”. On the other hand it is the success criterion for providing “standard service” – a core work-process (CWP) in the occupation of car mechatronic. “Competence”, therefore, is clearly related to output evaluation. “Competencies” on the other hand refer to performance determining resources: sets of behaviour that are instrumental in the delivery of desired results or outcomes (Kurz & Bartram 2002). Usually, person-related performance inputs are classified as knowledge, skills, and abilities. Carrying out tasks of a standard service, for example, requires that the technician has learned to operate all tasks belonging to the algorithm of standard service properly. Figure 1 visualizes the relationship between the two terms which are two sides of the same medal. Competence tests, by definition, try to measure and predict the *evaluated outcome of person-environment (situation = work process) interaction*.

**Figure 1: Competencies and Competence**



Second, there is a difference between “novice” and “expert”. An expert is a highly efficient worker who has acquired competencies necessary for good performance. A novice is a beginner that has no or very limited experience and capabilities with work related tasks and conditions.

**Figure 2: Minimal Competence (entry level)**



By referring to minimum competence, test scores shall be used to classify advanced beginners, who have gained first work experience (e.g. after a minimal initial training) and are ready to enter an occupational career. Figure 2 shows the “target range” of entry-level competence tests.

Entry-level competence tests are designed to support two areas of decision-making (inferences):

- *Certification.* Certifying workers is a strategy to minimize risks associated with unskilled actions. Presented as evidence within application, a certificate will reduce uncertainty in connection with employment decisions.
- *Evaluation.* Entry-level tests can be used to reveal evidence about VET effectiveness in terms of graduate employability. In this context, test scores will be analyzed on group level. Repeated measurement is the method of choice for monitoring changes in the VET system (e.g. teacher training, curriculum changes).

### Test dimensions

Entry-level tests have a two-dimensional structure. The first dimension is given by the CWPs of an occupational standard. The second dimension pertains the type of cognitive performance that is prompted by test materials.

Content validity requires that each occupation based on CWP is represented by a group (set) of test items. Furthermore, CWPs have to be met and balanced with respect to the item number and item difficulty. Different average item numbers and also different average item difficulties between CWPs have a distortion potential on test scores. If it is not possible to reach a balance between CWPs, counter measures are necessary, e.g. definition of minimum correct answers per CWP in the test use specification (see below).

The second dimension is linked to the distinction between active and receptive cognition. Active cognition goes along with analyzing, evaluating, and creating, whereas receptive cog-

nition is associated with remembering, understanding, and applying. Active cognition is prompted by items that confront the test taker with problem solving. A receptive item is one that can be solved by applying an acquired rule. The distinction corresponds with the taxonomy provided by Bloom (1956) (see also Anderson & Krathwohl 2001). Companies do expect more abilities regarding receptive performances than active performances from entry-level workers; nevertheless it is desirable that entry-level workers have the potential to become active members of the work organization. This in mind, an active/passive ratio of 2:3 or even 1:3 seems appropriate for entry-level tests.

### Items

Four-option multiple-choice – one question or stem with four answering options, one correct and three incorrect – is the preferred item format in entry-level tests. The literature provides extensive construction rules (e.g. Haladyna & Downing 1989). For example, items should minimize reading time, focus on a single problem, and use vocabulary that is consistent with test taker level of understanding. In addition, questions should contain most positive phrasing. Finally, options should not overlap, be consistent in length, and be homogeneous in content.

Multiple-choice items have advantages and disadvantages. The advantages are:

- The format is well known and self-explaining. Thus, test takers – with different cultural or educational background – don't need much instruction how to deal with them. Familiarization with this item format also facilitates the participation of practitioners in the process of item generation.
- Item generation requires relatively little effort provided that sufficient items are available. Bearing in mind that a number of items have to be replaced or revised before test release, this makes multiple-choice items very economical. In addition, test delivery (paper & pencil or computer-based) and test administration can be achieved with moderate resources.
- It is possible to cover a broad range of knowledge in one test. Because relative completeness regarding CWP is a validity factor of entry-level tests, this potential is essential. [Typically, entry-level tests in Oman use 10-15 items per CWP, adding up to 90-120 test items].

There are also disadvantages. Because the disadvantages limit the diagnostic power of entry-level tests, it is necessary provide coping strategies:

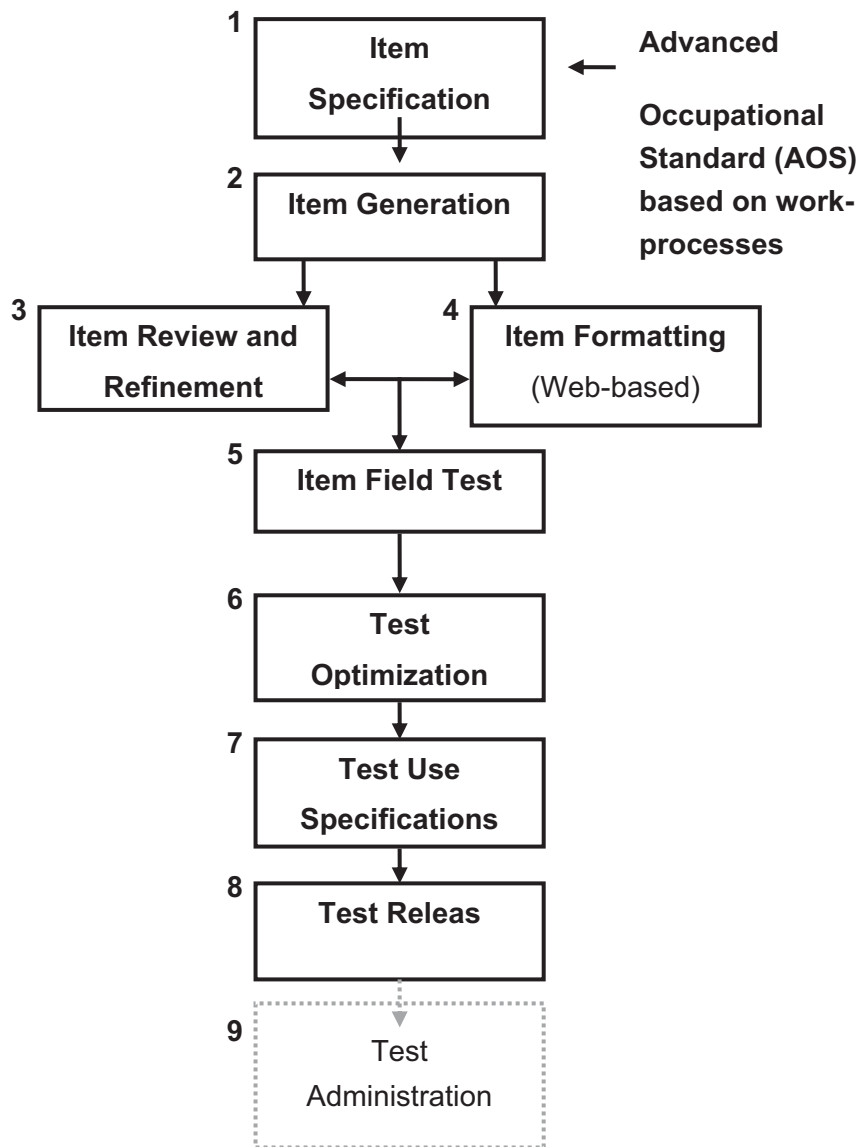
- Test takers have to deal with *symbolic representations* (text, pictures; multimedia) instead of interacting with real life content. There are different strategies to minimize the influence

of language skills on test results. Some are incorporated in item writing rules. A direct strategy is the development of the test in more than one language. [Entry-level tests in Oman are delivered in English because this is the mandatory language in the context of work].

- Tests lack sensitivity regarding sensomotor skills and negotiating dynamic tasks. Yet, only work samples can measure coping with dynamic tasks and conditions. Because work samples draw vastly from resources in development and are associated with many practical constraints, they are not suitable as a stand-alone solution for entry-level competence testing. Still it might be an optimum to use both test formats in combination.

### **Test development overview**

Figure 3 presents the steps of entry-level test development which are used in the Sultanate of Oman (Schütte et al, 2009). Test development, thus, is realized through a rather linear sequence. However, steps 2-4 usually have to be repeated two or more times – depending on the quality of new generated test content. Moreover, the step “test optimization” (step 6) includes activities that have been carried out in the steps before, e.g. replacing underachieving test content in the item field test.



### *Item Generation*

Item generation are realized in workshops over 2 or 3 days. Most important participants in the workshop are practitioners with substantial experiential knowledge from the shop-floor level. In addition, one external subject matter expert and a number of facilitators are necessary to run the workshop.

The role of the external expert is to do on the spot evaluation of items, i.e. checking technical correctness and plausibility. The expert is also a dialog partner, helping practitioners to explicate their experiential knowledge regarding CWP in a way that good items can be harvested. This role description needs a person with substantial technical expertise, excellent communication skills and a good general overview about CWPs.

On the first day, participants are introduced to the purpose and agenda of the workshop. The introduction includes presentations (with examples) of the concept of entry-level competence, item writing rules, and test dimensions.

The process of item generation is realized actively through group brainstorming and discussion. The practitioners are encouraged to verbally express ideas for items (including a short statement why they think it is a good item). When an item becomes clearer, it is written in front of the group via computer for mutually fine tuning. In the end, the participants have to judge the difficulty of the item for the target population. The progress of item generation is monitored through a visual chart.

The items need to be reviewed before the field test. Subject matter experts, testing experts, and practitioners from the workshop (follow-up interviews) are involved in this process.

## **Results**

In the following, we present field test data and experiences in item generation related to the development of an entry-level competence test for car technicians in Oman.

### **Item generation and field test**

Six practitioners from five companies followed the invitation to the item generation workshop. Work experience in the group varied from 6 to 26 years ( $M=17$ ).

In total, 120 items, evenly distributed over 8 CWPs (=15 items/CWP) which are specifying the occupational profile, have been prepared for a field test: 72 items (60%) were received from the workshop, 48 items (40%) were borrowed from sources outside of Oman. All 120 items of the field test version have been reviewed for validity aspects by practitioners.

The ability of the workshop group to generate items remarkably depended on CWPs. The group even declared a ninth CWP irrelevant for entry-level which, therefore, was excluded from the test.

The field test sample consisted of 58 workers (from six companies) which represent entry-level competence – according to supervisor judgement. A questionnaire was administered alongside with the test to obtain potential moderator variables, particularly skills strength in English (there was no native born speaker of English in the sample).

### **Item and test analysis**

Analysis of items considered classical item indices such as item difficulty and item discrimination because they are relatively simple to compute and are easy to understand (even for lay people). In addition, classical item statistics do not require sample sizes as large as required by statistics derived from Item Response Theory (IRT).

It is important to notice that the interpretation of item indices depends on the test purpose. For criterion-referenced tests, appropriate representation of domain content is more relevant than optimizing, for example, item discrimination and item difficulty (Schmeiser & Welch 2006, 338-339). Based on recommendations in the test literature, 20 items from the pool of 120 were discarded. The final test, thus, consists of 100 items.

Cronbach's coefficient alpha was used as an internal consistency reliability estimate. Alpha increases as a function of item interrelatedness and test length. In the actual case, the coefficient for the 100 items test was .91. A reliability coefficient of .70 or higher is considered acceptable for scales. Coefficients for CWP subscales (eventually useful to provide competence profiles) varied between .14 and .64.

Possible moderator variables of test results which have been analyzed based on self-assessment and self-report data (rating scales). A multiple linear regression was calculated to predict subject's test score based on English skills, work related skills, years in actual position, educational degree, and age. A significant regression equation was found ( $F(5, 48) = 4.882, p < .001$ ), with an  $R^2$  of .337 (stating that 34 percent of the test score variation is "explained" by the regression formula). Only language skills and age were significant predictors. The simple correlation between test score and English skills measured with a 5-point rating scale) was .42. Roughly, 16 percent of test score variation is thus explained by test taker's skills in English language.

### **Test use specifications**

Rules regarding test use have been specified in two directions: A cut-off score, i.e. a test score that a certification seeker must reach in order to pass the test (and be able to receive a certificate), and minimum correct answers, i.e. the minimum number of items related to every CWP a test taker must answer correctly; a person that reach a test score above the cut-off score threshold but fail to demonstrate sufficient competence in every CWP will not pass the test.

*Cut-off score.* The estimation of a cut-off score for the test was realized by considering empirical data (cut-off scores typically are estimated through expert judgement). A test calibration sample of 76 employees from 8 companies carried out the 100-item test in paper & pencil format. The sample was selected to represent entry-level competence in the domain by the companies. There was no time limit given; in average the individuals needed 3 hours to complete the test.

**Figure 4: Results of the test calibration sample**

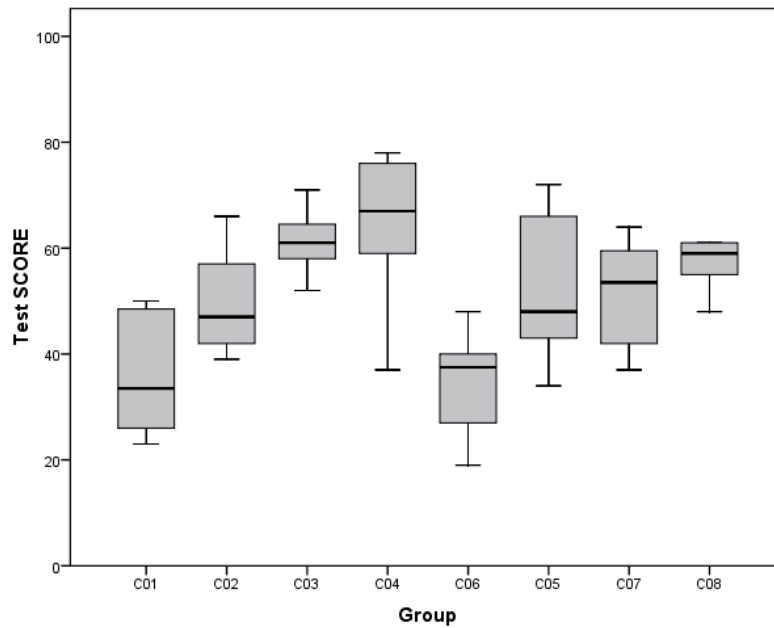


Figure 4 shows the results in the test “calibration” sample for the 8 companies (C01-C08) in form of box plots. Each box represents the lower and upper quartile distance (50 percent of the data); the line dividing the box is the distribution median. In average, 51 out of 100 items were answered correctly (SD=14). The maximum score was 78 and the minimum score was 19. Thus, entry-level competence in work organizations today seems to be relatively low in the Sultanate. This might be an effect of a lack of skilled workers. However, companies show substantial variation in test scores. The difference between the highest (C04) and lowest (C01) median was 33 points.

For specification of a cut-off score, it is assumed that the average test score is a good approximation of entry-level competence. Therefore, either the arithmetic mean (51 points) or the median (52 points) should be used. Alternatively, one could use the easier to achieve 25-percent percentile of the total distribution which is located at 40 points.

For comparison, a group of 57 technicians who freshly graduated in 2009 from different Vocational Training Centers in Oman (after a one-year, sometimes two-years’ training course) received an average test score of 32 (SD=8). This result, though exploratory till repeated measures are available, clearly indicates a need for improvement in the direction of work orientation of VET.

*Minimum correct answers. Specification of minimum correct answers is a strategy to avoid that a test taker with selective competence accomplishes the test by solving a subset of items regarding CWPs. Minimum correct answers have been estimated by using a simple-formula: item difficulty minus one standard deviation multiplied with the number of items belonging to the CWP.*



## Perspectives

An important result of the described testing approach is the fact that the development of a test design oriented towards work-processes is feasible. However, this can only be achieved by simultaneously developing items in cooperation with experts and/or practitioners. They know the requirements that occur during the work-process and can then include this knowledge into the items. This step is not self-evident for test designers who have to master the testing theories. Practitioners are always presumed to lack the necessary Abstraction level for the design of tests. The project proved that as soon as practitioners were adequately involved, they quickly became the most important actors for the item design. They are indispensable for the implementation of work-process requirements into the tests.

## References

- Anderson, L.W. & Krathwohl, D. (Eds.) 2001. A taxonomy of learning, teaching, and assessing: A revision of Bloom/s taxonomy of educational objectives. New York: Longman.
- Bloom, B.S. (Ed.), 1956. Taxonomy of educational objectives: Book 1, Cognitive domain. New York: Longman.
- Clouser, B.E. & Case, S.M. 2006, Testing for Licensure and Certification in the Professions. In: R.L. Brennan (Ed.), Educational Measurement. Westport: American Council on Education, Praeger Publishers, 701-731.
- Haladyna, T.M. & Downing, S.M. 1989. A taxonomy of multiple-choice item writing rules, Applied Measurement in Education, 2, 37-50.
- Kurz, R. & Bartram, D. (2002). Competency and individual performance: Modelling the world of work. In: I.T. Robertson, M. Callinan & D. Bartram (Eds.), Organizational effectiveness: The role of psychology. New York: John Wiley, 227-255.
- Lave, J. & Wenger, E. 1991, Situated Learning: Legitimate Peripheral Participation. New York: Cambridge University Press.
- Schmeiser, C.B. & Welch, C.J. 2006, Test Development. In: R.L. Brennan (Ed.), Educational Measurement. Westport: American Council on Education, Praeger Publishers, 307-353.
- Schütte, M., Klaffke, H., Hitz, H. & Al Khatib, A.R. (2009). Automotive Mechatronic Occupational Entry-Level Test. Test Development Documentation and First Test Application. Muscat, Sultanate Oman: OSSTC, Ministry of Manpower.
- Spöttl, G. (2007): Work-process analyses. Muscat, Sultanate Oman: OSSTC, Ministry of Manpower.

# Monitoring of Qualification and Employment in Austria – an empirical approach based on the labour force survey (LFS)

Dr. Lorenz Lassnigg, Dr. Stefan Vogtenhuber

**Abstract:** The purpose of the study is to make use of the information about ISCED-study-fields available with the LFS since 2004 for a monitoring of the relationship of initial VET to employment in Austria. First a meaningful classification of VET supply (upper and post-secondary ET, higher education) has been constructed; second a set of indicators has been developed and explored using available data. The indicators include three sections (a. demography and gender; b. employment, unemployment and income; c. competences, occupations and trades). The monitoring should produce signals about risks and opportunities across the ET-supply spectrum, comparing the empirical results for the various programmes. The results are not meant to produce „hard evidence“, but rather empirical indications which should draw attention to certain areas of risk, in order to decide about further inquiries. The theoretical approach is “constructivist”: the results should provide an input into a social system of knowledge production among the various stakeholders to create common reflection about the empirical relationship between supply and demand.

**Keywords:** monitoring, indicators, VET, knowledge production

## Introduction

An attempt for the development of a monitoring system of the relationship between initial VET and employment in Austria is reported, using new Labour Force Survey-(LFS)-data. The background is a strong initial vocational education and training-(VET) system, which has to – in one way or another – govern its relationship to employment (e.g., by setting decisions about the structure of qualifications, or about the financial inputs and the regional distribution of ET institutions, etc.) The rationale of the monitoring system is not a narrow technical concept of skills demand or matching, but an open conception of the VET-employment relationship which is influenced by several factors from the demand and the supply side.

Currently the very complex structures of governance of the Austrian VET system are based mainly on informal flows of information that are held by the various stakeholders, but are not formally documented. Some formal exercises of forecasting labour demand are mostly not related to the supply side. The data about the labour market development do not include information about the VET specialisations, which has been available in the past only every ten years with the population census (Lassnigg 2006, 2004a).

More recently, a new variable about the ISCED fields of study has been included in the European LFS. These data are explored in our study as a potential basis for the develop-

ment of a regular monitoring system that can provide information about the empirical relationship between initial VET and employment. We contend that this empirical information, if it is provided on a regular basis, can contribute to a more evidence-based decision making in the governance system of VET and higher education (HE). Currently the system is to some degree driven by informal and localised knowledge held by the various stakeholders and groups or networks of stakeholders in the system, and ultimately the market is correcting the decisions based on this informal knowledge about the ET supply. If formal information about the use of qualifications is lacking, several kinds of myopia or distorted perception might arise: E.g., strong and powerful suppliers might be interested to “market” their supply; on the demand side there might be an interest to create an oversupply that can beat down the wages; new demands might not be visible because of a lack of related powerful stakeholders, etc.

### **Targeted Problem**

The problem targeted by the project has been threefold: (1) the lack of regular and timely information about the relation of initial ET and employment; (2) the lack of a usable and comprehensive classification of the specialisations of the initial VET system; (3) the creation of regular information flows that might infuse the provided information into the governance system of initial ET.

### **Lack of information**

Initial VET is a strong and specialised system in Austria, with its main domain at the upper secondary level. About 85% of the 20-24-years age cohort, and about 80% of the 25-64-years population have attained an upper secondary qualification (NBB I, D5, D8), most of them a VET qualification. The supply of initial ET is still strongly differentiated and specialised, with more than 500 specialisations in two sectors (full time schools with upper and lower level, and apprenticeship) at upper secondary level<sup>16</sup> and even more programmes in higher education.<sup>17</sup> Of course, many of these programmes are attended by few students only; nevertheless, infrastructure and personnel must be provided and planned including decision about scarce resources.

Employment prospects should be at least be taken into account in these decision, what, however, is not possible if even simple information about the employment of graduates or completers of the programmes is not available.

---

<sup>16</sup> We find about 50 school types with about 300 specialisations in VET full-time schools and about 250 apprenticeship occupations (see <http://www.abc.berufsbildendeschulen.at/de/news.asp>).

<sup>17</sup> Almost 500 programmes are documented at the polytechnics or universities of applied science (<http://www.fachhochschulen.ac.at/de/studienangebot>), and the programmes at universities also exceed 200 by much (<http://www.studieren.at/categories/Studienrichtungen/>).

### *Lack of a classification of ET supply*

A first step to develop regular information about employment is the construction of a classification that is able to describe the differentiated supply in a meaningful way. Such a kind of classification does not exist so far; rather there are lists of programmes in different orders, mostly based on institutional categories. The lack of a shared classification means that no language exists among the stakeholders to even communicate the overall structure of the system.

Moreover, taking into account the high number of specialisations, it is not possible to create an overview at this level of differentiation. As a first task, a classification has been created which should be usable by the stakeholders. Therefore a classification has been developed stepwise through a series of workshops including experts from the supply and demand side, from the public employment service, from statistics, and from research and development (R&D). The result of this procedure was a classification of 44 programme categories spanning the overall spectrum of ET supply from compulsory education to higher education.

The ISCED study fields were used as a reference, and the classification strategy was to identify major programmes<sup>18</sup> in terms of size by specific categories, and to merge the smaller programmes to meaningful categories due to level and field (Lassnigg & Vogtenhuber 2007, 46-47).

### **Creation of regular information flows to initiate knowledge management/production among stakeholders**

Based on the experience, that R&D-based forecasts and analyses have not been recognized – not to mention used – in the past a basic idea of the study has been to create a non-technical descriptive tool for a regular monitoring of the relation of ET-supply and employment. The stakeholders involved in the development of the classification should also be regularly informed about the new monitoring results.

Because of the institutional fragmentation of the system to different sectors (full-time schools and colleges, apprenticeship, polytechnics, and universities) and to different groupings of specialisations within the different sectors (engineering, business, social work, health, agriculture, production, services, etc.) the stakeholders are considering only their sector/specialisation, without looking at the overall development. To overcome this fragmenta-

---

<sup>18</sup> Almost 500 programmes are documented at the polytechnics or universities of applied science (<http://www.fachhochschulen.ac.at/de/studienangebot>), and the programmes at universities also exceed 200 by much (<http://www.studieren.at/categories/Studienrichtungen/>).

<sup>18</sup> About 50 programmes were attended by at least 1.000 students in the reference year (2002/03); this shows the strongly skewed distribution of the supply; the distribution also shows that the specialization is much higher in traditional production fields and in engineering than in service oriented programmes..

tion the project has taken a comparative view, showing at set of indicators for the different specialisations.

Moreover, in order to initiate knowledge management and knowledge production among stakeholders, the system was not designed to produce ready truths but rather indications that should be combined with the informal knowledge in the various sectors/specialisations, and produce questions for further consideration and analysis. Thus a kind of language and common ground should be developed, to create awareness about what goes on empirically, as basis for reflection about the goals and objectives of programmes and their practical implementation (Lassnigg 2004b)

## **Methodology**

The methodology used is rather simple, by creating a multi-faceted set of detailed indicators, and Abstract indexes to provide a comparative overview across the classification of initial ET supply. The agreed classification of ET supply includes 44 categories at the different levels according to the structure of the Austrian system (17 apprenticeship categories including „masters“; 7 lower level secondary VET schools; 6 upper level secondary VET colleges; 2 post secondary categories; 10 higher education categories combining polytechnics and universities; 2 categories without specialisation: compulsory school and upper level academic school).

The development of indicators was constrained by severe data problems, as the Austrian LFS does not allow the disaggregation of yearly cohorts by age-years and sex because of too small sample size. We are confronted with the problem that the main European and Austrian data source about employment and the labour market still does not allow a simple descriptive follow-up of the transition from school to work of young people for statistical reasons. Therefore we created a set of more indirect indicators that include three sections based on the pooled LSF 2004-06:

- Demography: age, migration, gender; 11 indicators
- Employment: employment rates, unemployment, income; 9 indicators
- Utilisation: competences, occupations, trades; 5 (condensed) indicators

**Figure 1: Indicators**

<b>Indicators demography (11)</b>	<b>Weighting (in relation to average)* and interpretation</b>
(1) % female among the employed	<i>[descriptive: +1 female, -1 male]</i>
(2) % female among those who completed programme (completers)	<i>[descriptive: +1 female, -1 male]</i>
(3) % female among completers / % female among “young”** employed	+/- 0,5; above (below) average: high (low) out of LF among females
(4, 5) % young, % old employed among total employed	+1 young / -1 old; young: high supply dynamic; old: high replacement
(6) % young - % old employed	+1; high replacement
(7) completers / employed	-1; high supply dynamics
(8) average cohort of older employees / completers	Low +1 expansion; high -1 replacement
(9) % non-nationals among employed	+/- 0,5; low (high) demand
(10, 11) % traditional migration countries; % EU new member states	+/- 0,5; 0,5; low (high) demand
<b>Indicators employment (9)</b>	
(1) Employment rate	-1/+1; high/low demand
(2, 3) Markedly lower female employment rate than male; young, total	+0,5; +0,5; low demand
(4, 5) Low employment rate among young as compared to total m; f	+0,5; + 0,5; low demand
(6) unemployed / completers	+1+2; high unemployment
(7) unemployed / completers f // unemployed / completers m	+0,5; high female unemployment
(8, 9) Income as compared to average of education level above +10% or below -10%	total; corrected for gender -/+ 0,5; 0,5; low (high) risk

<b>Indicators utilisation (6)</b>	
Competence level	
(1, 2) % higher, lower competence level of employed relative to education level	-/+ 1; low (high) risk
Occupation, trade	
(3) Index forecast of main occupations	high, low demand -/+ 1
(4) Index forecast of main trades	high, low demand -/+ 1
(5, 6) Concentration of trades, occupations per VET programme (GINI-Index)	high, low concentration -/+ 0,5; 0,5

\* Interpretation of signs: + always “high risk”; - “low risk”; 1 > 0,5 higher/lower risk

\*\* “young” were defined differently according to upper secondary and higher education levels

The *demography* indicators reflect firstly the descriptive aspect of segregation of programmes by gender, which is quite high in the Austrian system. Second the relationship of young and old employees indicates potential replacement demand or a high current additional supply of young completers. Third, an increased employment of non-nationals as compared to the average is seen as an indicator for high demand, as the employment of non-nationals is handled rather restrictive.

The *employment* indicators include firstly overall and gender specific employment and unemployment figures, and unemployment related to completers. The relative income as compared to the average of the ET level is also taken as an indicator for demand.

Finally some more complex indicators should reflect current and future *utilisation* of qualifications. First, the relative competence level of employees should reflect higher/lower demand, second forecasting figures are used, and thirdly the concentration of occupations/trades per ET programme should also reflect demand.

The procedure includes the following steps

(1) the calculation of the 25 indicators, and sorting of ET-programmes for each from maximum to minimum;

(2) graphical representation of each indicator: upward right increasing risk, downward left increasing opportunities;<sup>19</sup>

(3) inspection of each figure and definition of the section indicators by setting the cutting points: assignment of (+) and (-) to categories above/below the cutting point with 0,5, 1 or sometimes 2, according to the weight for risks/opportunities; inspecting the empirical distributions, at the point where discontinuities are occurring, the cutting points are set.

(4) summing up per section, and totally for the composite indicator.

## **Results and Perspectives for further development**

*For the described instrument/approach as such*

The results allow for a comparison of the 44 categories of VET programmes first by the composite indicator. The composite indicator varies between +5 and -7, with two thirds of the categories between +/-2. Nine categories are indicated with increased risk, and six categories with increased opportunities (figure 2, 3).

Increased risk results to some degree from demography, in particular in the expanding upper level programmes. In the more traditional programmes rather utilisation gives the strongest contribution. On the side of reduced risks rather utilisation indicators play the strongest role, followed by demography.

Concerning the levels of programmes there is no clear picture of upper or lower levels, both increased and reduced risk programmes are mixed. There are three big programmes in each of the risk categories, female programmes are over represented among the increased risk programmes and male programmes are even more over represented in the reduced risk programmes.

A second step allows for an inspection of the different sections from where the risks/opportunities are resulting (Figure 4-6). Figure 4 shows the demographic indicators. There are five programmes, which show a very strong demographic dynamic. All of them are situated at the upper level of higher education or VET colleges. Only few programmes show high indications of replacement, mainly in the apprenticeship sector including the masters (Meister). The demographic structure also shows that almost one third of the programme categories are strongly biased by gender (6 male, 6 female).

---

<sup>19</sup> The results of this procedure can be found in the internet: [<http://www.equi.at/material/indicators-monitoring.pdf>]



**Figure 2: Results composite indicator and contribution of sections**

<b>Indication of increased risks</b>	<b>Weighting (in relation to average)* and interpretation</b>
VET college services (5 pts.) medium, female	Demography 5,5; <i>employment –0,5</i>
Apprenticeship food, textiles (5 pts.) small	Utilisation 4; employment 2,5; <i>demography –1,5</i>
VET school services (4,5 pts.) big, female	Utilisation 2; employment 1,5; demography 1
VET college business (4 pts.) medium, female	Demography 3,5; utilisation 1; <i>employment –0,5</i>
Apprenticeship hairdresser (4 pts.) big, female	Utilisation 2; employment 2
University business (3,5 pts.) big, male	Demography 4,5; <i>employment –1</i>
Apprenticeship health (3,5 pts.) medium, female	Demography 2; utilisation 1; employment 0,5
Apprenticeship retail (3,5 pts.) small	Employment 2; utilisation 1, demography 0,5
Apprenticeship agriculture (2,5 pts.) medium	Demography 1; utilisation 1; employment 0,5
<b>Indication of reduced risks</b>	
VET school health (-7 pts.) medium, female	<i>Utilisation -5; employment -2</i>
HE production and construction (-4,5 pts.) small, m	<i>Utilisation -2; employment -1,5; demography -1</i>
Apprenticeship construction (-3,5 pts.) big, male	<i>Demography –1,5; utilisation -1; employment -1</i>
HE medicine (-3,5 pts.) small, male	<i>Utilisation -4; demography 0,5</i>
VET college pedagogy (-3 pts.) big, female	<i>Utilisation -2; employment -1</i>
Apprenticeship electricity, electronics (-2,5 pts.) big, male	<i>Demography –1; utilisation -1; employment –0,5</i>

The employment related indicators (Figure 5) show positive indications (high employment rates, low unemployment related to the completers) in some of the higher education programmes, and negative indications are present rather in the apprenticeship sector (high unemployment and low employment rates).

Figure 6 shows the utilisation indicators. Few programmes have strong positive indications from the forecasts and also from the utilisation of qualifications. The latter is distributed in a rather mixed manner across the spectrum of programmes.

Summarizing the results for the individual programmes we can firstly look at the programmes with increased risk. Six to eight of the 44 programmes show indications for increased risks, originating in different dimensions:

- high supply dynamics, without visible demand: *VET colleges in services and business*, and *higher education business* programmes;
- bad employment and bad forecasts: *apprenticeship food and textiles* and *apprenticeship hairdresser*;
- each of the three sections not favourable: *VET school services*, *apprenticeship health and retail*.

Those programmes are rather low level qualifications and rather big programmes.

For four to five of 44 VET programmes good opportunities are indicated, coming from the competences and the forecasts, and partly from employment – the supply dynamic is rather low:

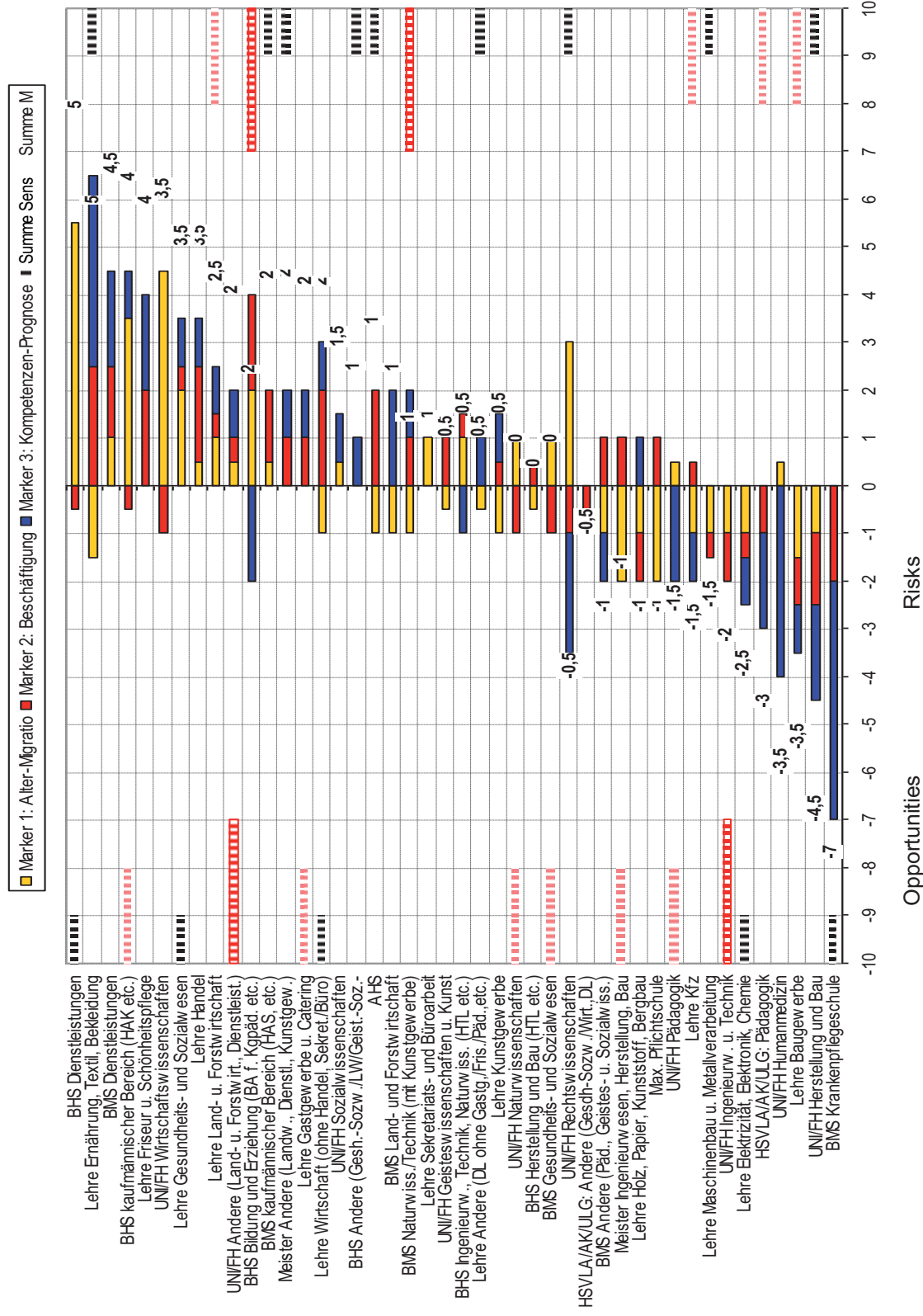
- two programmes with good competence and forecast indications are in the health services (*higher education medicine* and *VET colleges health*)
- mixed positive indications show two programmes in *construction* (higher education and apprenticeship) and the *postsecondary pedagogy* programmes.

Those programmes are at higher or medium level and small or medium size

The remaining 31 to 34 programmes show moderate values, some are slightly biased to risks (ca. seven programmes) or to opportunities (ca. nine programmes). *Law studies* show a polarised structure of high supply dynamics and good values at employment and utilisation. Thirteen programmes don't show a deviation from average, supply dynamic is often low, employment and utilisation is rather unfavourable.

**Figure 3: Composite indicator**

Summenmarker





**Figure 5: Section indicator employment**

Marker 2: Beschäftigung

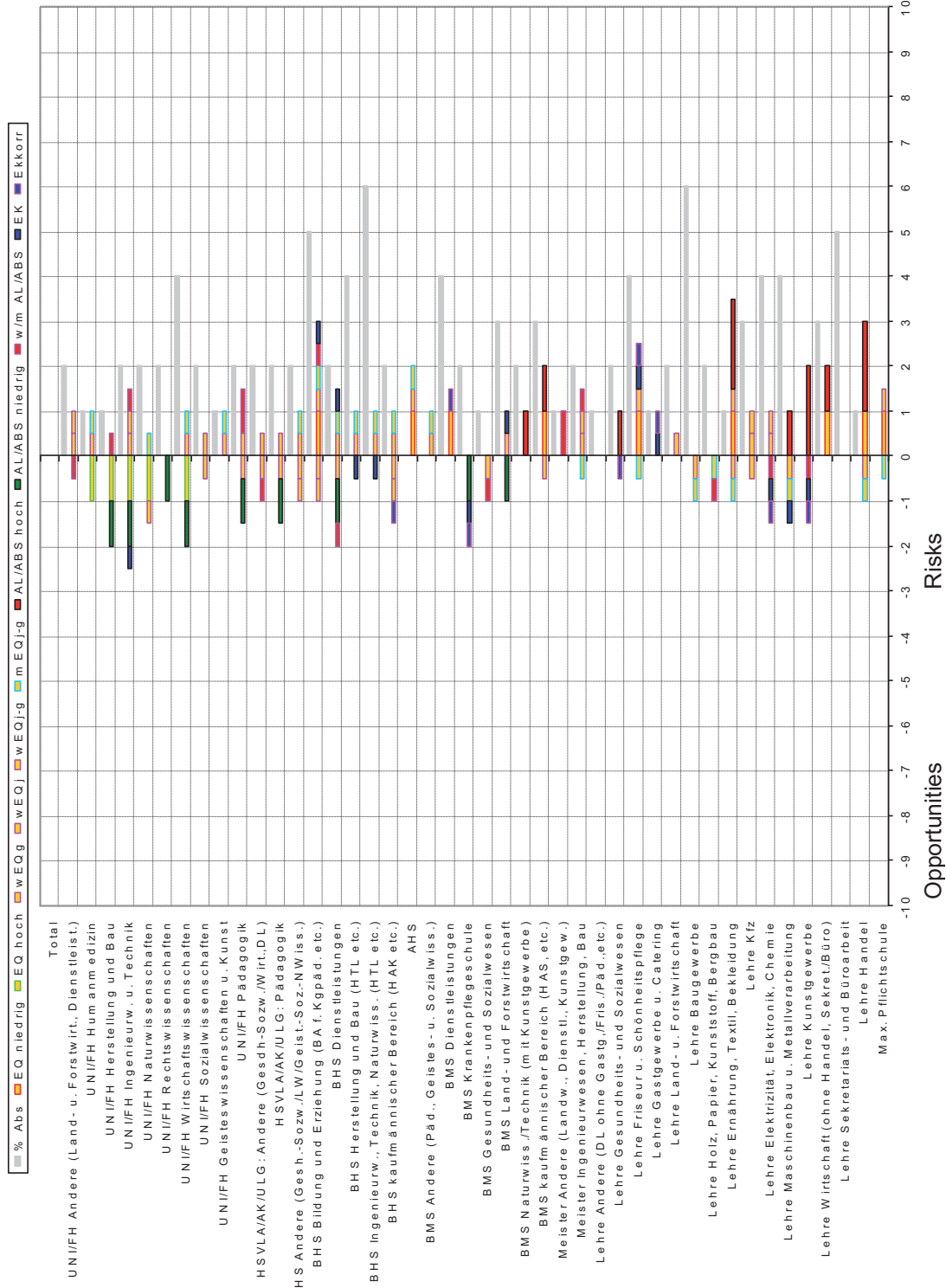
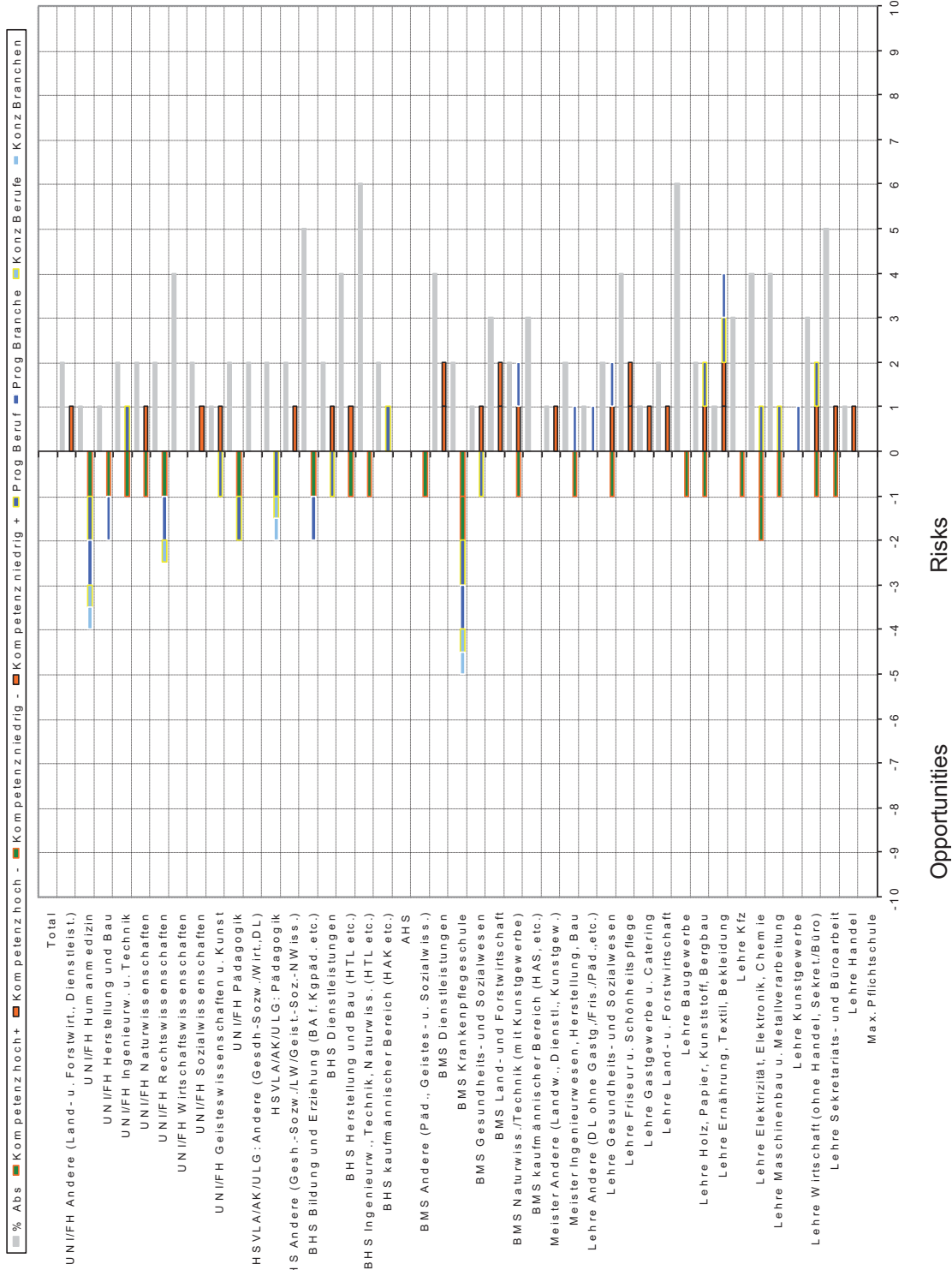


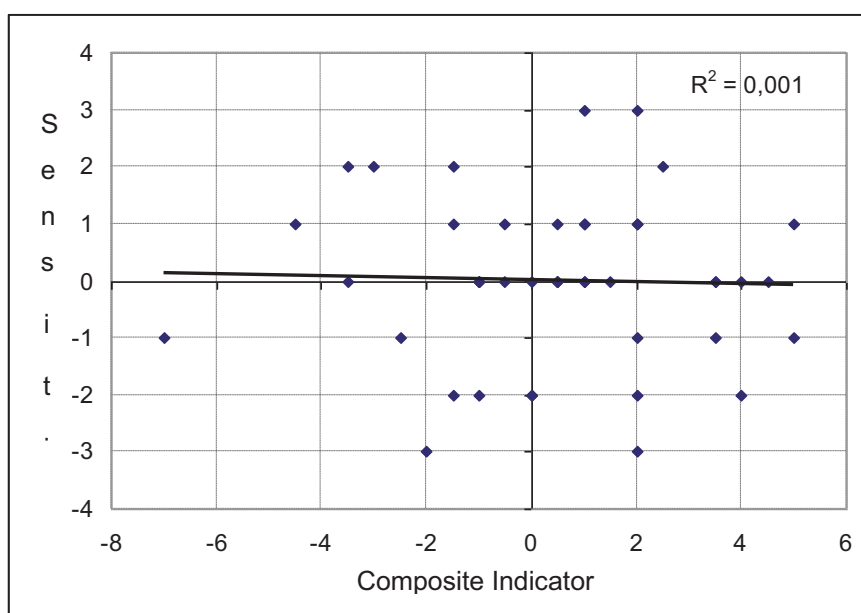
Figure 6: Section indicator utilisation

Marker 3: Kompetenzen, Berufe, Wirtschaftsklassen



Potential weaknesses of the methodology applied might be seen in its simplicity. One is the “handmade” setting of the cutting points. We can check for this by looking whether some programmes are often neighbouring the cutting point, so that they would shift into the risk/opportunity category if the cutting point were slightly different. If we check this we can see that the maximum at which some categories are neighbouring the cutting point are three times out of 25 indicators (4 programmes, none of them in the risk or opportunity group), 10 programmes are neighbouring the cutting point for two times and 14 programmes one time. Altogether 28 out of 44 programmes fail this sensitivity test; that is more than half of the programmes (see figure 3 where the bars at the left and right edges indicate the number of neighbouring indicators at the risk-right or the opportunity-left side). Figure 7 shows that the sensitivity marker is uncorrelated with the composite indicator.

**Figure 7: Scatter of the rank of programmes on the composite indicator with sensitivity**



More sensitivity analyses should be done for the control of the composition of the indicators, e.g., by taking a technical cutting point (the x highest/lowest values), and comparing the results.

Another step of validation, which could be particularly interesting if comparisons by country are made, might be modelling the composition of the composite indicator by its components.

As for this first run of the study only the LFS 2004-06 was available, we can also follow-up the development.<sup>20</sup>

<sup>20</sup> This will be done for the final paper, and maybe already for the presentation at the Bonn-workshop.

### *Potentials and challenges for the use in/for international comparisons*

In principle the methodology can be used for international comparisons, as the data of the LFS are available for the EU member states. A main challenge will be the differing structures of VET systems between countries. Two versions could be envisaged: one that takes the differing structures as point of departure, and analyses the components of the risks and opportunities; another that constructs a kind of standard structure of ET programmes at a similar level of specialisation.

### **References**

- Lassnigg, L. & Vogtenhuber, S. 2007. Klassifikation für Ausbildung und Beruf: Darstellung der Klassifikation. IHS-research report. Vienna [Internet: <http://www.equi.at/dateien/IHS-Ausbildungsklassifikatio.pdf>; 2009-11-23].
- Lassnigg, L. 2004a. To match or mismatch? The Austrian VET system on struggle with diverse and changing demand. Berufs- und Wirtschaftspädagogik - online Nr.7 [Internet: [http://www.bwpat.de/7eu/lassnigg\\_at\\_bwpat7.pdf](http://www.bwpat.de/7eu/lassnigg_at_bwpat7.pdf); 2009-11-23].
- Lassnigg, L. 2004b. Indicators for Quality in VET. European Experience. IHS Sociological series Nr.63 [Internet: <http://www.ihs.ac.at/publications/soc/rs63.pdf>; 2009-11-23].
- Lassnigg, L. 2006. Approaches for the anticipation of skill needs in the "Transitional Labour Market" perspective - the Austrian experience. WZB Discussion paper SP I 2006-105 [Internet: <http://skylia.wz-berlin.de/pdf/2006/i06-105.pdf>].
- NBB-Nationaler Bildungsbericht Österreich 2009. Vol.I: Das Schulsystem im Spiegel von Daten und Indikatoren. Graz: Leykam [Internet: <http://www.bifie.at/buch/657/1/d>; 2009-11-23].



# Using Survey Instruments in Longitudinal Evaluations of Vocationally Orientated Higher Education: Problems and Pitfalls

*Prof. Jim Stewart, Victoria Harte*

**Abstract:** The EU emphasises vocational relevance and employability in policy directed at provision of higher education. This has perhaps been particularly the case in the UK with a raft of government sponsored initiatives intended to improve the impact and contribution of public funds in universities on national and firm economic performance. The authors designed and implemented a longitudinal evaluation of the work of the Institute for Enterprise, a UK government funded centre at Leeds Metropolitan University. The study utilised an established survey instrument and intended to assess the impact of enterprise education on undergraduates over a three year period. Experience of 18 months has suggested a number of limitations in the survey instrument and indeed in the use of surveys as a method of evaluating vocationally orientated education. The paper describes, explains and analyses experience of this project as a basis for exploring and discussing how best to design evaluation studies.

**Keywords:** employability, higher education, longitudinal evaluation

## Introduction

This paper describes a planned three year, longitudinal study of the impact of enterprise focused modules, on higher education students, delivered in numerous subject disciplines and the problems and pitfalls that the authors encountered within the first 18 months of the project. The study is part of the work of The Institute for Enterprise, a Centre for Excellence in Teaching & Learning (CETL) within Leeds Metropolitan University and funded by the Higher Education Funding Council for England (HEFCE).

The project has now ended in its original form. The decision to end the project as originally designed was taken at its 2 year point in May 2009. The main focus of this paper is to discuss the project and the reasons for ending within the context of methodology for evaluating Enterprise Education (EE). The paper will describe the project but will not cover the detail of the actual results.

The objectives of the paper are to:

- outline and discuss the project;
- outline and discuss the chosen research design and methods;
- discuss in detail the problems encountered;
- discuss the reasons for ending the project as originally designed;
- offer some insight into the opportunities and limitations of researching EE and engaging higher education students in institutional research.

## Targeted Problem

It is clear that support for enterprise curriculum in the UK is very strong (Matlay 2007; Moreland 2006). From a UK perspective there are two main factors behind the requirement for higher education to take enterprise and entrepreneurship more seriously. First, there is the overwhelming importance of single person enterprises to the UK economy. As one of the major outcomes of higher education is to prepare students for the world of work, it needs to recognise the significance and incidence of self-employment for the employability aspects of the HE curriculum. Secondly, the UK Government sees small businesses as a driving force for the UK economy and also as a contribution to making Britain a more entrepreneurial society (Moreland, 2006). The European Commission also endorses the need to increase the number of start-ups and business successions as well as improve the support for young entrepreneurs (European Commission, 2004a, b).

While there has been a boom in the development and take up of EE there appears to be a deficit of evaluation of its impact (Kailer, 2005; Charney & Libecap, 2000). Investment of resources in delivering undergraduate modules to encourage students to consider the creation of enterprise and appreciate entrepreneurial behaviour has been vast. EE is a growth industry and has been widely intensified at universities. The number of entrepreneurship chairs increased by 120% within five years in the US alone with over 270 endowed positions in 2000 (Kailer, 2006; Charney & Libecap, 2000). Most evaluations have been short term (Kailer, 2005; Harte & Stewart, 2009) with negligible investment in long term evaluations. Investment in evaluating the subject does not reflect its development. Despite that, many evaluations have taken place but findings of evaluation studies rarely have impact. There are obvious difficulties with evaluating EE and despite calls for a unified framework (Fayolle, 2006; NCGE, 2008) the replication of evaluation research will not ease the task; rather it will be a hindrance and add to the many inadequate evaluations that have taken place to date.

An EU project named 'Entreva.net' analysed methods for evaluating EE and most studies were taxonomised as 'mere monitoring'. Only one quarter of all the studies analysed could be called evaluation (Stampfl & Hytti, 2002; Hytti & Kuoposjaervi, 2004). The types of EE that are offered across the UK by higher education institution's (HEI's) are so varied that one methodology of evaluation would not be suitable because of the varied ways curriculum is designed and delivered. Some evaluations have measured impact but on different variables; others measure pure inclination or intent whereas some measure only outcomes or the number of successful start-ups and level of sales volume and turnover. Other evaluations have been based on the written assignments and exam performance of students (Kailer, 2005). Furthermore, the number of evaluation findings that have had an impact on the curriculum where improvements have been made is also low. The influential evaluation of the Berger Entrepreneurship Program reports no changes or improvements to existing EE. Rather the

evaluation focussed on the impact of EE on the high number of entrepreneurship graduates it produced and the increase in funding that the university attracted as a consequence (Charney & Libecap, 2000).

The project reported here had a direct focus on and purpose of conducting an evaluation in one university using a longitudinal design. This was our main focus and of our research question:

“To explore the impact of enterprise modules in higher education on students’ own perceptions of their self-efficacy and their motivations towards particular career intentions.”

The rest of the paper deals with the problems experienced attempting to address that question through a longitudinal survey based research design. We begin in the next section with a brief overview of that design.

## **Methodology**

The original design began in October 2007 with a robust survey instrument (questionnaire), developed and tested by Cambridge Management Institute (CMI), which we were using with their permission. It is worth noting that when CMI undertook their project, the Enterprisers Programme, they didn’t apply the questionnaire longitudinally. The programmes they evaluated were week-long residential programmes designed to develop entrepreneurial skills and build confidence, and their total sample size of their first programme in 2003 was 55. They did however confirm that the instrument could be used in designs such as ours.

### *Survey Method – the questionnaire*

The questionnaire is underpinned by Bandura’s (1997) theory of self-efficacy (SE). It is intended to capture changes in student perceptions of their skills and abilities (SE) and their attitudes to different career options whilst studying enterprise related modules (Boyd & Vozikis, 1994). Changes in the perceptions of the student are captured by the use of one pre-test and two post-test completions of the questionnaire which respondents complete on the first day of studying an enterprise module, a second time at the end of the module and a third at a point in time following the module, usually six months later. The pre- and post-test design illustrates how the respondents’ ratings of their self-efficacy either increase or decrease over time. In the case of the CMI programme the last day was day 5, whereas our students’ last day was at the end of the semester, so approximately 12 weeks between completing the pre-test and the first post-test questionnaire. The changes are captured using a series of questions separated into four sections: section 1 - ‘general entrepreneurial skills and characteristics’; section 2 - ‘direct relation to business contexts i.e. profit or non-profit venture’; section 3 – ‘specific entrepreneurial intentions’; and section 4 – ‘motivations to

choose a particular career choice'. Other data collected on the questionnaire requests general baseline data such as age, gender, level of course, name of course etc. and further questions enquire about antecedents, role models and prior exposure to enterprise.

#### *Population, Samples and Data collection*

The total population was defined as 'all those students enrolled on enterprise or enterprise related modules within Leeds Metropolitan University'. It proved impossible to establish an accurate figure to meet this definition. The sample was identified using data from the Student Management Information System by identifying appropriate modules and then permissions from lecturers to attend lectures to distribute the questionnaires. The intended structure was a sample in semester 1 of 07/08 and a separate new sample in semester 2 of 07/08, following the process of three questionnaires being completed (i.e. 1 pre-test and 2 post-test) then follow the same students through levels 2 and 3, repeating the process. However, this approach did not work out in practice.

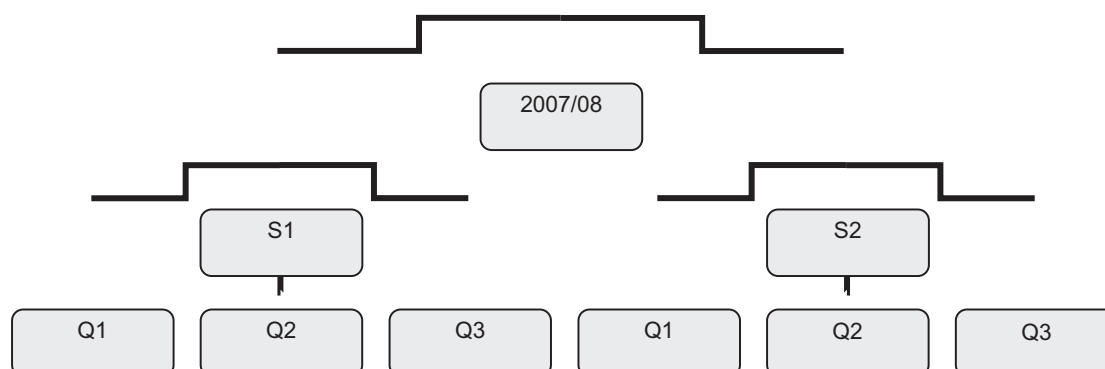
#### **Results and Perspectives for further development**

It is now well established and accepted that there is a downwards trend in response rates to social surveys (Thompson and Surface, 2007) and that the trend applies to surveys of higher education student populations (Sax et al, 2003). The trend clearly adds to the problem of attrition generally experienced in longitudinal studies (Burton et al, 2006). Nevertheless, in principle, the intended structure was a valid approach to embarking on a longitudinal design. However, numerous problems occurred within the first year (2007/08).

#### *Attrition & Non-response*

The actual data collected was very different from that proposed data and expected. The diagram below is a pictorial illustration of actual data collected for the period 2007/08.

**Figure 1: Actual Data Collected**



As can be seen, the follow-up data in semester 1 was very disappointing and resulted in only  $n=13$  responses. Despite numerous reminders the number of responses did not increase and plans had to be made as to what approach to adopt in the second semester. The very low response rate to Q2 in semester 1 resulted in the decision not to attempt to collect Q2 data from the second sample in semester 2. Instead, it was decided to scale down the sample size and focus on a much smaller sample. The authors believed that this approach would build-up a better relationship between the researcher and participant and possibly enhance the participants' desire to engage more and increase their saliency and interests in the research itself (Laurie & Lynn 2008).

Collection of data in year 1 was via distribution at enterprise module lectures and achieved a response rate of nearly 200 in semester 1 and approx 300 in semester 2 (pre-test questionnaire). The rationale for this approach was that we expected a high attrition rate at the second wave (first post-test questionnaire). This expectation was in line with general experience (Burton et al, 2006). We were correct about the high rate of attrition but did not expect to have such a low response at the second wave. This was in part due to a problem related to budgetary constraints communicated to us by senior management along with a proposition/argument from the same source that students would prefer 'online' completion. We were asked to change the data collection method from paper-based to online completion. The reason was the costs of inputting data. We accepted this but still had reservations about the notion that students are 'technologically driven'. However, neither party (us, the researchers, or senior management) had evidence to suggest students were technologically driven or not. Changes to the data collection method were communicated to the students by letter asking them to go to a web address and complete the questionnaire online. Despite the costs involved in posting the letters it was made clear that a copy of the questionnaire was not to be included because there would be no funding available for inputting responses. This decision is one of a number of contributing factors that we have concluded may explain the low response rate. A study by Sax et al. (2003) found that offering an option of paper or online completion produced higher response rates than either as a single option. We acknowledge the potential dissonance between paper rather than e-mail communication in relation to the 'new' online completion method, which may also help to explain the low response rate. However, we were aware that numerous e-mail addresses for the sample were incorrect as this was confirmed by the high number of 'failed delivery' messages we received from previous communication to the sample. This, therefore, left us with only one method of communication via post.

Another factor may include distribution at lectures. We have reflected that students may have completed the questionnaire 'under duress' even though 'choice to participate' was explained prior to the questionnaires being handed out. Furthermore, research ethics guidelines were

adhered to and informed consent was sought in the form of a disclaimer with a full explanation on the first page of the questionnaire. So, while 'duress' may have been a factor we do not believe it was significant.

A further possible cause of attrition is the structure of the questionnaire and perceived interest (Lynn et al 2005, Burton et al, 2006). The questionnaire may be viewed by some as quite arduous to complete and includes some relatively complex questions. Lynn et al (2005) offer the suggestion that at second and subsequent waves respondents have information about the survey which they did not encounter at the beginning. This is because they have had experience of the survey at wave one and, potentially, they conclude that they have no interest. This could be said to be the case with our study. With respect to perceived benefits, the majority of modules included were compulsory. Therefore, the perceived interest and benefits by the students may have been low and thus impacted upon respondents' interest in and willingness to contribute to the research (Lynn et al 2005, Burton et al, 2006). In addition, though we have no substantive research data to justify this, we still believe that had students had a paper copy of the questionnaire at the second wave, more would have completed and returned them than did online. There is some evidence in the literature to support this belief and in particular the studies by Sax and her colleagues (2003) and by Thompson and Surface (2007).

The data we expected to capture at the 2 year point in the project was not achieved. We believe this was primarily due to changes in the method of distributing the questionnaire as described above. Almost 190 students responded to the pre-test application of the questionnaire in Semester 1 but of the same sample only n=13 responded to the first post-test application of the questionnaire. These figures raise some significant questions for research design. One lesson is simply that the method of questionnaire distribution, completion and return is a significant decision in research design. Our lesson here confirms that of previous research (Sax et al, 2003, Burton, et al, 2006, Thompson and Surface, 2007). A further lesson is that changing those methods during a project is probably inadvisable. Other questions include the most effective means of encouraging responses among higher education students. It seems from our project that a personal appeal by the researcher has some effect. It also seems that handling paper copies has a positive effect. This raises the second and related question on the efficacy of online completion. There are a number of issues here. First, whether the current generation i.e. millennial's (Oblinger 2003) prefers online interaction as much as is sometimes claimed. Second, whether the deliberate effort and action of logging on to a specific website led to a poor response rate. Related to this is the question of whether the easier option of completing a questionnaire already in the hand and putting it in an envelope also already in the hand would have led to a higher response rate. A third and final general question is whether it is advisable when researching higher education students to use a

research design that relies on multiple completions from respondents (see also Sax et al, 2003 and Burton et al, 2006). Our experience does not answer these questions but it gives pause for thought. It seems that this is particularly significant and important in longitudinal studies that usually rely on large samples and data sets. An additional lesson seems to be that it is unwise to assume that the familiarity with and preference for technology of the current HE generation necessarily translates into their participation in research projects.

A final point is the suitability of the CMI questionnaire in higher education. The Enterprisers Programme (EP) data was obtained in a controlled environment not comparable with a typical semester in higher education and the curriculum content on the EP reflected the design and content of the questionnaire, making the two very compatible. This relationship is not immediately obvious when applying the questionnaire to EE curriculum not based on the EP.

### **Conclusions and Further Research**

We can draw a number of conclusions from the experience described in this paper. First, the general principles described in the literature on research design have sound validity. This is especially true in relation to longitudinal studies. Second, that distribution of questionnaires, which can be seen as a fairly straightforward and low level decision, can in fact be the most critical decision of all and can make or break a project that relies on large samples and data sets. Our third conclusion is that design of research projects with higher education students as the subject population needs careful thought. In particular, the received wisdom to do with the technology preferences of HE students needs to be questioned and in particular when it comes to completing research instruments online (see also Thompson and Surface, 2007).

Improving the evaluation of EE seems to us a neglected topic. Improving methodologies for evaluation does require attention and is mentioned by some authors. The NCGE (2008) and Fayolle (2006) both call for common or unified frameworks for evaluation of impact but with the many various forms of EE curriculum at many HEI's this will be difficult to achieve. The 'one-size-fits-all' concept certainly doesn't apply in this instance, unless all EE curricula is unified and based on the 'theory of planned behaviour' (Fayolle, 2006). However, what also has to be taken into account is the impact of optional experiences such as extra-curricular activities that are not part of mainstream routes.

Our future attempts to evaluate the impact of EE at Leeds Met will take on a different form as we have learned that the CMI instrument does not contribute to a 'unified framework'. Evaluating EE is very much like the subject itself: generally outside traditional boundaries; ambiguous; opportunistic, and very difficult to define.



## References

- Boyd, N.G. & Vozikis, G.S. (1994) The Influence of Self-Efficacy on the Development of Entrepreneurial Intentions and Actions. *Entrepreneurship, Theory & Practice*, Spring, 1994.
- Burton, J., Laurie, H. and Lynn, P. (2006) The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *Journal of the Royal Statistical Society*, 169. Part 3.
- Charney, A. & Libecap, G. (2000), The Impact of Entrepreneurship Education: An Evaluation of the Berger Entrepreneurship Program at the University of Arizona 1995 – 1999, Revised Final Report to The Kauffman Center for Entrepreneurial Leadership, Eller College of Business and Public Administration, University of Arizona.
- Church, A. H. (1993) Estimating the effect of incentives on mail survey response rates; a meta-analysis. *Public Opinion Quarterly*, 57, pp62-79. in Laurie, H. & Lynn, P. *The Use of Respondent Incentives on Longitudinal Surveys in Peter Lynn Methodology of Longitudinal Surveys*, John Wiley, chapter 12, 205-233, 2009.
- European Commission – Enterprise Directorate-General (2004b), Final Report of the Expert Group “Education for Entrepreneurship”, October, Bruxelles.
- Fayolle, A., Gailly, B. & Lassas-Clerc, N. (2006) Assessing the impact of entrepreneurship education programmes: a new methodology. *Journal of European Industrial Training*, 30, 9, pp-701-720, 2006.
- Harte, V. Stewart, J. (2009) A longitudinal approach to evaluating undergraduate students studying ‘enterprise education’ modules; pitfalls problems and struggles, the 8th European Conference on Research Methodology for Business and Management Studies, University of Malta, Valletta, Malta, 22-23 June 2009.
- Hytti, U. & Kuopusjaervi, P. (2004), *entreva.NET - Evaluating and Measuring Entrepreneurship and Enterprise Education: Methods, Tools And Practices*, Small Business Institute, Business Research and Development Centre, Turku School of Economics and Business Administration, Turku.
- Krueger, N.F. & Brazeal, D. (1994) *Entrepreneurial Potential and Potential Entrepreneurs*. *Entrepreneurship, Theory & Practice*, Summer, 1994.
- Dearing, R. (1997) *Higher Education in the Learning Society*. London: DfES.
- HM Treasury (2008) *Enterprise: unlocking the UK's talent*. Available at: <http://www.berr.gov.uk/files/file44992.pdf>
- Lambert, R. (2003) *Lambert Review of Business-University Collaboration*. London H.M. Treasury. Available at <http://www.lambertreview.org.uk> and [www.hm-treasury.gov.uk/media/lambertemergingissues\\_173.pdf](http://www.hm-treasury.gov.uk/media/lambertemergingissues_173.pdf)



Laurie, H. & Lynn, P. The Use of Respondent Incentives on Longitudinal Surveys in Peter Lynn Methodology of Longitudinal Surveys , John Wiley, chapter 12, 205-233, 2009.

Lynn, P. et al (December 2005) 'A Review of Methodological Research Pertinent to Longitudinal Survey Design and Data Collection', Working Papers of the Institute for Social and Economic Research, paper 2005-29. Colchester: University of Essex.

Lynn, P. (2005) Longitudinal surveys methodology. International Statistical Seminar no.45. Basque Statistical Institute (EUSTAT). Spain.

Matlay, H. & Carey, C. (2007) Entrepreneurship education in the UK: a longitudinal perspective.

Journal of Small Business and Enterprise Development, Vol 14, 2, pp 252 – 263.

Moreland, N. (2006) Entrepreneurship and higher education: an employability perspective. York, Higher Education Academy and ESECT Learning and Employability Series.

National Council for Graduate Entrepreneurship (2008) Developing Entrepreneurial Graduates: putting entrepreneurship at the centre of higher education. London: NESTA.

Oblinger, D. (2003) Boomers, Gen-Xers, and Millennials: Understanding the "New Students." EDUCAUSE Review, Vol 38, 4 p36-40,42,44-45.

Peterman, N.E. & Kennedy, J. (2003) Enterprise Education: Influencing Students' Perceptions of Entrepreneurship. Entrepreneurship, Theory & Practice, Winter, 2003.

Porter, S. & Umbach, P. (2006) Student survey response rates across institutions: why do they vary? Research in Higher Education, Vol 47, 2, pp 229-247.

Sax. L. J., Gilmartin, S.K, and Bryant, A.N. (2003) Assessing response rates and non-response bias in web and paper surveys. Research in Higher Education, Vol. 44, No. 4.

Sherman, P. (2005) A Proposed Look at the Role of Self-Efficacy as a Mediator between Experiential Entrepreneurship Education and Performance Outcomes. Allied Academies International Conference, University of Evansville.

Stampfl, C. & Hytti, U. (2002), Entrepreneurship als Herausforderung an das Bildungswesen – Ansätze in Österreich und europäischer Vergleich – Ergebnisse des Projektes ENTREDU, Institut für Bildungsforschung der Wirtschaft, Vienna.

Stewart, J. & Knowles, V. (2003) Mentoring in undergraduate business management programmes. Journal of European Industrial Training, Vol 1, 27, no.3.

Stewart, J. (2007) 'The ethics of HRD', in Rigg, C., Stewart, J. and Trehan, K. (Eds.) Critical Human Resource Development: Beyond Orthodoxy, London: FT Prentice Hall.

Thompson, L.F. and Surface, E.A. (2007), Employee Surveys Administered Online: Attitudes Toward the Medium, Non-response and Data Representativeness. *Organizational Research Methods*, Vol.10, No.2.

Heinonen, J. et al (2006) Breeding Entrepreneurship in Life Sciences: study on a training programme for the Graduate Schools of the Academy of Finland. Turku School of Economics, Business Research and Development Centre <http://www.tukkk.fi/ytkk>

Wylie, C. (2001) Competent children: findings and issues from the first 7 years. Ministry of Social Policy seminar, The Long Road to Knowledge. New Zealand Council for Educational Research.

# **The use of large-scale administrative data sets to monitor progression from VET into Higher Education in the UK: possibilities and methodological challenges.**

*Dr. Geoff Hayward*

**Abstract:** A variety of measures exist that purport to measure the effectiveness of Vocational Education and Training (VET) Systems. However, there is a lack of longitudinal data. This paper focuses on the methodological challenges of using administrative data sets. Such data are typically computerized records that are gathered for some administrative purpose, but contain information that can be used for other purposes as well. A classic example is birth records. These data are maintained as a matter of public record and have long been computerized to facilitate their use. The challenges of working with administrative data to answer substantive research questions of interest should not be under-estimated. Upon reflection about the use of such data, three crucial questions will be formulated that need to be considered when working with such data.

**Keywords:** administrative Data, longitudinal data

## **Introduction**

A variety of measures exist that purport to measure the effectiveness of Vocational Education and Training (VET) Systems. However, typically these are measures of stocks - the number or proportion of people enrolled on particular courses at particular levels; the number or proportion holding qualifications of a particular at a particular level – or various measures of return to individuals who hold a particular level of qualification, deemed to be a proxy for the value of a qualification in the labour market. Such measures are useful but they tell us little about the progression of individuals, and the factors that affect such progression (including the contribution of VET programmes), between different phases in their learning career, for example from secondary to tertiary education or from VET programmes into the labour market, relative to other individuals who have followed other learning pathways. Without detailed information about such transitions and factors affecting the probability of success it is difficult to provide individuals with sound career advice and guidance about which pathway to follow to realise their ambitions. Yet arguably this is how we should judge the value of VET programmes: the extent to which they provide access to valuable outcomes, such as good jobs or progression to higher level education and training programmes that in turn lead to successful insertion into the labour market, as opposed to, say, warehousing people in education and training programmes through times of economic downturn.

The challenge with providing evidence about the success, or otherwise, of such transitions, is that it requires longitudinal data of sufficient quality to exert the statistical control needed to make, *ceteris paribus*, valid claims about the effectiveness (or otherwise) of particular VET

interventions to support transitions to successful outcomes. Such data is very expensive to generate and is, accordingly, rare. It does exist, for example in England the various birth cohort studies, but the gaps between each cohort precludes drawing conclusions about issues of interest associated with changes in the labour market and the Higher Education (HE) widening participation agenda in the UK, for instance, since 1990.

Other panel studies, such as the English Longitudinal Survey of Young People (LSYPE), will provide useful information for one group of young people progressing to age 25. The main role of the study is to identify, and enable analysis and understanding of, the key factors affecting young people's progress in transition from the later years of compulsory education, through any subsequent education or training, to entry into the labour market or other outcomes. Data from the study will be used, among other things, to monitor the progress of the cohort group, evaluate the success or otherwise of policy aimed at this group and provide an evidence base for further policy development. Sample boosts took place for deprivation factors and for ethnicity. However, the data is for one cohort starting in 2003/04 so is of limited use at the moment in terms of examining labour market progression.

Yet other panel data sets are either too restricted in their age coverage (for example the Youth Cohort Study) or too limited in their sample size (eg the British Household Panel Survey) to be of much use in answering the questions of interest in measuring the effectiveness of VET systems.

Another alternative is to link administrative data sets and so follow individuals through different phases of their learning career. This was the approach adopted by the ESRC funded 'Degrees of Success Project: the Transition between VET and HE'. In this paper, I reflect on the challenges of using the administrative data sets employed by this project and report on some of the main findings to illustrate the potential utility of such administrative data sets in exploring the effectiveness of VET programmes as preparation for HE progression.

### **Background to the Project**

The project "Degrees of Success: The Transition between VET and HE" analysed the transitional landscape of students coming to Higher Education (HE) from different educational backgrounds, with a focus on those holding vocational qualifications. The project described the distributional patterns and seeks to explain them by looking at different factors and characteristics of students and institutions.<sup>21</sup> The empirical basis was threefold: a) an analysis of administrative large-scale databases, b) data from two surveys administered by the project to

---

<sup>21</sup> See Dunbar-Goddet and Ertl (2007) for more details about the project design and research questions.

students in five Higher Education Institutions (HEIs) in three subject areas and c) interview data with admissions staff, lecturers and students.<sup>22</sup>

This paper focuses on the methodological challenges of using administrative data sets. Such data are typically computerized records that are gathered for some administrative purpose, but contain information that can be used for other purposes as well. A classic example is birth records. These data are maintained as a matter of public record and have long been computerized to facilitate their use. In the case of education administrative data are typically collected for funding and accountability purposes.

The main strength of administrative data is their supposed availability. No new data collection is required by the research team as another body has gathered and entered the data, and it is apparently therefore lying there waiting to be analyzed. Available databases are often relatively large – in the case of the Degrees of Success project in excess of half a million records. That usually means it is possible to analyze population subgroups separately, and administrative data can often be useful in identifying problems specific to a particular subgroup or geographic area.

However, administrative data are not without serious problems and limitations. First, the data can be "dirty" and careful scrutiny of all fields is needed to identify aberrant entries. Second, administrative data are often incomplete and there is a consequential issue about how to deal with missing data. Third it is critical to remember that the data were gathered for another purpose, for example in the case of the data sets employed in the Degrees of Success study to collect information about course attendance for funding purposes or for managing entry to Higher Education. The fields of greatest interest, for example the qualifications held by applicants to HE, may or may not be central to the primary record keeping or payment purpose.

### **The data sets**

No single dataset yet exists within the UK which makes the tracking of students over different stages of a learning career from completion of compulsory secondary education to completion of an undergraduate degree across all qualification types possible. Therefore one has to look for different datasets and try to combine them in a meaningful way. The Degrees of Success project aimed to combine three different data sources: a) the post-compulsory, but pre-HE, data included in the Individual Learner Record (ILR) from the Learning and Skills Council (LSC), and the Pupil Level Annual School Census (PLASC) from the DfES; b) the applicant/admission data from UCAS; and c) the Higher Education Statistical Agency data on students in Higher Education. All three administrative datasets hold the full population of the

---

<sup>22</sup> See Dunbar-Goddet and Ertl (2008) for an introduction and results of the first survey and Hoelscher et al. (2008) for some results from the quantitative analyses.

specific groups, which should permit drawing quite reliable transition landscapes between different educational stages and pathways.

### **School and college data sets**

However, the PLASC data set holds limited data on students enrolled in post-16 VET provision in schools. The restricted fields available at the time of the project were partly a function of the need to minimise the administrative burden on school managers when completing a return for each student on each field. New approaches to managing this data set in the future may yield better functionality for answering research questions about transitions from VET provision. Further, at the time of the project plans were only being formulated to produce a unique pupil identification code that would enable an individual to be tracked across all data sets. Without such an identifier, matching of cases in the PLASC data set to other data sets needs to be done by fuzzy matching processes. However, the reliability of this process is not open to public scrutiny as it is undertaken by a private foundation. Consequently no further use was made of this data set. The Individual learner Record is two databases. One consists of the qualifications a learner is taking, the second a record of learners. The purpose of the data is to release funding to Further Education, Sixth Form, and Tertiary Colleges and a range of private training providers for the programmes they deliver. Potentially this is a useful data set but it only covers a limited portion of the post-16 population – it excludes school students. Consequently using this data to track the progress of students confounds two factors: the qualification taken and the centre where learning took place with no means of controlling for the absence of school based students. Further, the data sets require considerable re-working to make them suitable for the type of analysis envisaged for the Degrees of Success Project. Finally, matching of data from the ILR to UCAS and HESA data sets again would require fuzzy matching processes. Subsequent to the project the PLASC and ILR have been merged with HESA data.

### **UCAS data**

UCAS organises the applications of all applicants for full-time undergraduate study at Higher Education Institutions (HEIs) in the United Kingdom.<sup>23</sup> These huge administrative datasets provide information about the gender, age, socioeconomic status, ethnic background, and potential disabilities of applicants, as well as a restricted postcode of the applicants' home address, their UCAS tariff score (a numerical measure of attainment but which is only available for some qualification pathways), their choices (for which subject they applied and at which type of institution) and, most importantly for this project, with the applicants' prior qualifications and whether they are accepted or not by an HEI. We compared data from a ten-year time-span (1995-2004), so we are able to show trends over time.

---

<sup>23</sup> There are some minor additional routes for direct entry, but the numbers are very small.

The UCAS data is collected for administrative rather than for research purposes. Therefore, although the data is a rich source of information, it is limited in different ways.<sup>24</sup> First, it only holds information about the application process. Although this includes some information about the applicants' prior qualifications, the data is not very detailed. Furthermore, although one knows whether they are accepted by a University, it is not clear from this data alone a) why they are (not) accepted,<sup>25</sup> and b) how successful they are later in their studies. Thus, anything like individual educational pathways or learning careers over a longer period cannot be drawn on the basis of this data alone. Second, the data are, in the main, only for full-time HE students.<sup>26</sup> Although full-time students make up the majority of all students in UK HE, this might not be true for the special subgroup of students with a Vocational Education and Training (VET) background. An additional problem with time-comparisons is the stability of particular items. Most of the variables in the dataset stayed the same over the entire ten years; however, changes occurred in the coding of socio-economic status, a crucial variable, although UCAS provides information about how to transform the old codes into the new ones.<sup>27</sup> Changes were also made in the coding of prior qualifications, due to changes in the qualifications pupils could achieve. As far as one can see, these changes produce only minor changes over the time period. Another problem is the classification of qualifications. The dataset we used is not perfect for providing full information about prior achievements.<sup>28</sup> Further limitations of the data are discussed when each of the variables are described below.

These limitations notwithstanding, the results produced below have a high reliability for the largest group of students. One reason for this is that we have data for the whole population of applicants, not merely a sample.<sup>29</sup> The following analyses take only UK (domiciled) applicants into account, although some cases have been dropped, as they had no information about choices (less than 1%). Table 1 gives an overview of the datasets for the three years 1995, 2003/04 and 2004/05.

<sup>24</sup> See Gorard et al. (2006: Appendix A) for a more detailed analysis of problems with current analyses in HE.

<sup>25</sup> See Wilde & Hoelscher (2007) for a more detailed analysis.

<sup>26</sup> Even a small minority of full-time students can bypass UCAS through direct applications for HE courses or through internal progression from FE level to HE level at FE colleges. Additionally not all HEIs are members of UCAS. HESA data for 2004 show that 59 % of all first year students are full-time students, another 6 % on sandwich courses and 36 % are on part-time study. For students studying for a degree, the figure for part-time study reduces to 13 %, full-time raises to 78 % (National Statistics reports a figure of 90%, see [www.statistics.gov.uk/cci/nugget\\_print.asp?ID=9](http://www.statistics.gov.uk/cci/nugget_print.asp?ID=9)). It appears that the figures have been quite stable since the 1980s. A comparison over time by the DfES shows that in 1970/71 26.1% of the students in HE were part-time, 34.3% in 1980/1 and 33.9% in 1990/1. The figures are not directly comparable to current data, as they changed from headcounts to enrolments (see <http://www.statistics.gov.uk/STATBASE/xsdataset.asp?vlnk=189>, accessed on 12/04/07, for further information).

<sup>27</sup> The provision of socio-economic status (SES) information by the applicant is voluntary and therefore not that reliable for comparisons.

<sup>28</sup> Other problems with qualifications in surveys are discussed in Jenkins & Sabates (2007).

<sup>29</sup> As we use the whole population, traditional significance tests are not applicable, and the respective figures are not given below. There is, however, an ongoing debate if one can nevertheless view a whole population as a sample, e.g. the population of one year of applicants as a sample for applicants in different years (see Behnke, 2005). This depends on the research questions one poses. We will address this problem in another methodological paper.



**Table 1: Characteristics of applicants: Overview**

YEAR	N of cases	Women	Age (MEAN, Standard deviation)	SES <sup>1)</sup> (Mean, Standard deviation)	Accepted
1995	369,701	51.4 %	21.0 (5.88)	3.0 (1.89)	71.8 %
2003/04	406,165	54.3 %	20.6 (5.71)	3.1 (1.9)	82.2 %
2004/05	409,526	55.2 %	20.6 (5.62)	3.1 (1.9)	81.6 %

<sup>1)</sup> Socio-economic status is classified in seven categories (1 = High, 7 = Low) and is not a metric variable. Nevertheless, mean and standard deviation (in brackets) are shown to give a first impression of the distribution (see appendix B).

## HESA data

HESA holds data on all individuals who have enrolled at a Higher Education Institution (HEI), the type of HEI they are attending and also details about completion/non-completion rates. There are fields for qualifications held but these are often only partially completed and some socio-economic and demographic data. Importantly the HESA data has a student identifier that can be mapped back to the identifier given to the student when they applied through UCAS, enabling the matching of the datasets. Thus, the more detailed background information provided by the UCAS data can be matched to data about progression into and success within Higher Education.

I now present some initial results produced from analysing this data to illustrate its utility before turning to some concluding remarks.

## 2. Applicants and their characteristics by educational pathways

### *Defining the educational pathways*

The UCAS data holds detailed information about prior qualifications. On the basis of this information, five educational backgrounds could be differentiated<sup>30</sup>:

- General academic Vocational

Foundation and Access courses (FaA)

- Other
- No qualifications

<sup>30</sup> The full list of qualifications, and the coding of the five educational backgrounds, is given in appendix B.



Table 2 shows the changes in proportions applying for Higher Education from the different qualification groups. There are distinct changes over time. The number of applicants holding a general academic qualification has increased by around 5% over the last ten years, resulting in an overwhelming majority of applicants coming through the traditional general academic route. An even larger increase can be found for applicants holding vocational qualifications. They increased their share from around 18% to 25% over the same time span. However, the biggest increase can be found for “other” qualifications.<sup>31</sup> Interestingly, the share of people coming from a Foundation/Access (FaA) background is stable over time, despite the introduction of the new Foundation qualification during the period analysed.<sup>32</sup> The number of applicants with no reported qualification has decreased by half.

**Table 2: Qualifications held by applicants (%)**

	1995	2003	2004
<b>General academic</b>	70.7	75.3	75.7
<b>Vocational</b>	17.8	25.8	24.9
<b>Foundation/Access</b>	7.8	8.6	8.9
<b>Other</b>	5.1	13.5	14.1
<b>No qualification</b>	6.0	3.5	3.2
<b>TOTAL</b>	107.3	126.7	126.8

(more than 100%, as applicants can hold multiple qualifications)

However, the increase in the shares of applicants holding a certain qualification does not tell the whole story. The figures for each year add up to more than 100 %, because people can hold qualifications from more than one strand. One main message in Table 2 is the overall increase of qualifications held. While in 1995 only a small minority (7%) held qualifications from different pathways, this number has increased to 27% in 2003/04. The increase of around 12% in the overall figures of applicants (see Table 1) was accompanied by an increase in the diversity of the qualifications held from different pathways.

The different combinations of qualifications were therefore recoded into distinct qualification patterns. Out of the four qualification strands, seven “educational pathways” an applicant could be on were created in line with the project’s purpose to compare those with vocational qualifications with those coming through other routes. These pathways are:

<sup>31</sup> The UK and Overseas Degrees account for just over 20 % of all “Other” qualifications in 2003 and 2004.

<sup>32</sup> Around 44 % of these applicants have completed the newly introduced Foundation courses (not available in 1995). Foundation courses are one-year university preparation courses suitable for mature students who may not have formal qualifications, and also for students of any age without the entry qualifications for specific degree programmes, especially overseas students who have studied a non-British curriculum.

- only general academic ("Academic"),
- only vocational qualifications ("Vocational"),
- only Foundation and Access-courses ("FaA"),
- only other ("Other"),
- academic and vocational qualifications ("Ac + Voc"),
- academic and FaA ("Ac + FaA"), and
- any other combination ("Other Combination").

Table 3 shows the proportions according to these pathways, now adding up to 100 %.

<b>Table 3: Qualification pathways (in %)</b>			
	<b>1995</b>	<b>2003</b>	<b>2004</b>
<b>Academic</b>	63.4	50.8	51.3
<b>Vocational</b>	13.6	10.1	9.5
<b>FaA</b>	6.2	3.9	4.2
<b>Other</b>	3.5	5.9	6.0
<b>Ac + Voc</b>	4.2	14.1	13.7
<b>Ac + FaA</b>	1.5	3.2	3.1
<b>Other combination</b>	1.6	8.5	8.9
<b>No qualification</b>	6.0	3.5	3.2
<b>Total</b>	<i>100</i>	<i>100</i>	<i>100</i>

The growing diversity of qualification types applicants hold is again clearly visible. The first three groups decreased over time, while "Other" and combinations of different pathways increased. Nevertheless, the "Academic" group is still the biggest, accounting for more than half of all applicants.

The increase in vocational qualifications (Table 2) is mostly a result of an increase in their combination with academic qualifications. While the "Vocational" group was much larger than the combined "Ac + Voc" group in 1995, it decreased over time, and in 2003/04 the combined group is the larger group, having more than tripled its share of all applicants.<sup>33</sup>

<sup>33</sup> However, the changes might be, at least in part, a result of changes in the coding of data by UCAS for the different years.

In Abstract, the majority of those applying to Higher Education (70.7% in 1995 and 75.7% in 2004) held general academic qualifications but the proportion applying with vocational qualifications increased from 18 per cent (1995) to 25 per cent (2004). However, this growth was due to an increase in those combining vocational and general academic qualifications, up from 4 per cent to 14 per cent over this time period, while the proportion with vocational qualifications only decreased from 14 per cent to 10 per cent.

### **3. Do VET pathways widen Higher Education participation?**

This section presents results derived from the combined HESA UCAS data sets. Analysis of this data set indicates that vocational routes do open access to HE for non-traditional students. Applicants with a vocational background come from lower socio-economic groups, are more often male, older, are more likely to come from a non-white ethnic background, and are more often disabled, than those from the traditional academic route.

Unsurprisingly, those with “Academic” qualifications mainly come from State school sixth forms, Sixth Form colleges, Independent and Grammar schools. Most applicants from the “Vocational” pathways obtained their most recent qualification within the Further Education sector. A high proportion of those with combined academic and vocational qualifications gained their qualifications in state-school sixth forms and they are more similar to the traditional general academic applicants from such schools. Consequently, the increasing proportion of students applying to HE with a combination of academic and vocational qualifications may be leading to some increase in widening participation but this effect is not substantial.

Those with vocational backgrounds are much more likely to apply for courses in post-92 HEIs, and FE colleges offering HE provision than those with purely academic backgrounds. Those who combine academic and vocational qualifications are less likely to apply to at least one pre-92 institutions than those with just general academic qualifications, but more likely than those with just vocational qualifications. However, the application rate to post-92 HEIs is similar for those with just vocational and those combining vocational and academic qualifications.

Students with a vocational background are over represented in ‘less prestigious’ HEIs.

VET students are over-represented in applied fields such as creative arts and design and computer science and under-represented in areas such as medicine, dentistry, law, languages, history and philosophy. It could be that the more applied subjects chosen by VET students are over-represented in Post-1992 universities and other HEIs.

Controlling for the distribution of subjects across institutions it becomes clear, however, that students with an academic background were much more likely to study at pre-1992 universities than their counterparts with vocational qualifications. Their overrepresentation was especially visible in those subjects in which they were normally underrepresented. *If* someone

with an academic qualification was studying “Agriculture”, “Computer Science” or “Creative Arts” at all, *then* they were studying at a Pre-1992 institution.

Controlling for socio-economic background, other demographic factors and the type of educational institution attended prior to entry to HE, compared to traditional A level students those with VET qualifications have a much higher risk of not obtaining a place in HE, and of dropping out after their first year (Figure 1). But the picture is much more favourable for those combining the two pathways, who are nearly as successful at entering and completing the first year of HE than those with only general academic qualifications. However, multi-level models suggest that the situation is more complicated with the influence of VET background on dropping out varying significantly between institutions: those institutions with high drop out rates in general are where vocational students are most at risk of dropping out irrespective of the subject being studied. Furthermore, the risk of dropping out is less in HEIs which have more VET students.

**Figure 1: Attrition in different pathways**

Of 100 English students applying for a full-time degree programme in HE from one of the following educational pathways ...



#### 4. Concluding remarks

This paper has introduced two of the data sources of the “Degrees of Success: The transition between VET and HE” project: the UCAS applicant data and HESA data. The purpose of the paper was threefold. The first part made general remarks about these administrative data sets, their coverage and resulting limitations. Some problems were discussed, and the point was made that as the combined data contains nearly all full-time applicants for a place in HE, it is, nevertheless, a highly valuable data source.

The main aim of the project was to investigate the transition landscapes and the success of students in HE who have a vocational educational background. The second part of the paper therefore gave a detailed description of key characteristics of the applicants along their different educational pathways. The coding of eight distinct educational pathways: only general academic (“Academic”), only vocational (“Vocational”), only Foundation/Access (“FaA”), only other (“Other”), academic and vocational (“Ac + Voc”), academic and Foundation/Access

("Ac + FaA"), all other combinations ("Other combination") and no qualification ("No qualification"), was explained. The description took into account three domains of variables: socio-demographic characteristics, information about prior educational experiences and some data about the application process. Looking at these characteristics, it becomes obvious that vocational routes open access to HE for non-traditional students. Applicants with a vocational background are from lower socio-economic classes, are more often male and from a non-white ethnic background and are more often disabled, than those from the traditional general academic route. Additionally, improving access for students with vocational qualifications would track more students into some areas of study that are seen as crucial for the future competitiveness of the British economy (i.e. "Mathematical and Computer Sciences", "Engineering", "Technologies").

The third part of the paper provided some results from an analysis of the combined UCAS-HESA data set. In particular, it demonstrated that students aiming to enter full-time UK HE from a purely vocational background were at a significant disadvantage compared to their counterparts following an academic pathway. Furthermore, they tended to attend less prestigious HEIs and were more at risk of dropping out during the first year of Higher Education study.

While students from a VET background contribute to a widening in access, it remains a question whether their distribution across HEIs represents fair access for them. Of particular concern is the extent to which students from VET backgrounds find themselves primarily in less well-resourced HEIs. In addition, studying at a pre-1992 HEI brings extra economic benefits: students with an academic background have an advantage on the labour market. Clearly, as boundary objects Vocational Qualifications in the UK signify different attributes to academic qualifications, confirm status hierarchies and are interpreted, rightly or wrongly, as preparing students less well for HE study. HE remains a highly segmented market from the perspective of both VET applicants and from the perspective of HE admissions staff and tutors.

The challenges of working with these administrative data to answer substantive research questions of interest should not be under-estimated. Upon reflection three crucial questions need to be considered when working with such data.

### **1. Who owns data collected using taxpayers' money and who can legitimately use this information to answer research questions?**

The data collected by all the agencies involved is done so using public funding, yet the data are not readily publicly available. They can be requested for research purposes but this involves paying not inconsiderable sums of money for the data to be extracted. Further data has to be requested which means having an exhaustive knowledge of the variables and fields held in different data sets in order to specify accurately the data needed to answer a

particular research question. Ongoing, iterative negotiations were needed to ensure best quality data was obtained. There are good reasons for maintaining confidentiality around such data, for example identifying individuals and institutions which could lead to personal or organizational damage if the data are improperly used. However, other data sets are made available to researchers at no cost with appropriate checks and balances to maintain confidentiality. If wider use is to be made of these potentially valuable administrative data sets for independently assessing the performance of the VET system then, arguably, these data sets need to be made publicly available at minimum if any cost.

## **2. Who is missing from such administrative data sets and why?**

It is clear that whole groups of applicants and HE students were missing from one or more of the data sets and that different data was being collected about the same age learners because of the different use that data was being put to administratively. Of particular importance for judging the effectiveness of VET provision for supporting progression to HE is the absence of part-time students from the UCAS data set. This reflects a historical tradition of direct entry for such students to an HEI without passing through the UCAS system, which is geared to support, in the main, full-time students. This means that important progression routes, for example from apprenticeship programmes to part-time Foundation Degrees, are all but invisible in current administrative data even though such progression is currently a major policy concern. Clearly steps will need to be taken to develop the administrative data sets to facilitate the tracking of learners through a highly heterogeneous set of pre-HE pathways.

## **3. What are the incentives and disincentives for organizations to provide the full returns needed to maintain comprehensive administrative databases that can be used for research purposes?**

An abiding source of frustration in working with these administrative data sets is the preponderance of missing data. This suggests that certain fields, notably qualifications held and background information about learners, are routinely not completed or only partially completed. In part, this is likely to be the result of the time needed to collect such information from students spread across a university campus. If the information is not essential to trigger the release of monies from a funding body then the temptation must be to not complete that field but focus attention on the commercially more crucial ones. This is entirely rational for an administrator but completely frustrating for the researcher. Given the urgent policy need to evaluate the effectiveness and efficiency of VET provision, and given the need to utilize administrative data sets to do so, additional thought needs to be given to how administrators can be incentivized to provide full returns.

A second reason for poor data around qualifications held is that it is difficult for bodies such as HESA and UCAS to keep their data collection systems up to date with a rapidly evolving

qualification system. This problem will be exacerbated by the fast-track procedures for registering new qualifications in the various UK Qualification and Credit Frameworks (QCF), the intention to make such qualifications highly flexible in terms of their structure and composition in order to meet 'employer needs', and the potential incorporation of company training programmes as qualifications to generate the release of government funding for such training within the English QCF. Such activity potentially further reduces the utility of vocational qualifications for supporting progression to HE not least because little will be known about them and no data will be collected about their use in relation to HE progression. They will, in effect, be invisible and so reduce the potential for those studying on VET programmes to progress into Higher Education. The use of appropriately constructed administrative data sets would enable the detailed examination of such issues on the UK's HE widening participation agenda.

## References

Behnke, J. (2005). Lassen sich Signifikanztests auf Vollerhebungen anwenden? Einige essayistische Anmerkungen. *Politische Vierteljahresschrift*, 46(1), O-1-O-15.

Department for Education and Skills (DfES). (2007). Participation Rates in Higher Education: Academic Years 1999/2000 - 2005/2006 (provisional). Retrieved March 30, 2007, from <http://www.dfes.gov.uk/rsgateway/index.shtml>

Dunbar-Goddet, H., & Ertl, H. (2008). Investigating the learning experience of HE students with a vocational pathway: Findings from a questionnaire survey: Department of Education, University of Oxford.

Dunbar-Goddet, H., & Ertl, H. (2007). Degrees of Success: Research on the Transition from Vocational Education and Training to Higher Education: Department of Education, University of Oxford.

Gorard, S., Smith, E., May, H., Thomas, L., Adnett, N., & Slack, K. (2006). Review of widening participation research: addressing the barriers to participation in higher education. A report to HEFCE by the University of York, Higher Education Academy and Institute for Access Studies (Review). York: University of York, Higher Education Academy and Institute for Access Studies.

HEFCE. (2006). *Pathways to higher education. Access courses* (No. May 2006/16): Higher Education Funding Council for England.

Hoelscher, M., Hayward, G., Ertl, H., & Dunbar-Goddet, H. (2008). The Transition from Vocational Education and Training to Higher Education: A successful pathway? *Research Papers in Education*, 23(2).

Jenkins, A., & Sabates, R. (2007). *The Classification of Qualifications in Social Surveys* (No. Working Paper 2007/2). London: Institut of Education.

UCAS. (2000). Statistical Bulletin on Widening Participation. Retrieved 15/02, 2008

Wilde, S., & Hoelscher, M. (2007). Missed Opportunities? Non-placed applicants (NPA) in the UCAS data. Project completed for UCAS by Oxford University Department of Education. Oxford/Cheltenham: UCAS

Younger, M., & Warrington, M. (2005). *Raising Boys' Achievement* (No. Research Report 636). London: (University of Cambridge, Faculty of Education).

#### *Used datasets*

UCAS Applicants Data Sets 1995/06, 2003/04 and 2004/05. Copyright UCAS 2007.

UCAS cannot accept responsibility for any inferences or conclusions derived from the data by third parties. All results are presented in accordance with the UCAS/HESA rounding methodology.



# Methods and instruments for quality assurance of the initial and vocational education and training system in Romania

*Dana Stroie*

**Abstract:** In Romanian initial VET system (TVET)<sup>34</sup>, a National Quality Assurance (QA) Framework, consisting of methodology and instruments through which quality is ensured at system and provider level, was developed and implemented, starting with 2003. The pillar of QA in TVET is the annual schools' self-assessment, finalised with the Self assessment Report. Each Self assessment Report is subject of a validation process, conducted by county inspectorates. On basis of the validated Self Assessment Reports, an annual Report regarding the quality assurance of TVET at county level is elaborated by county inspectorates. All county reports are integrated in an annual Report regarding the quality assurance of TVET at national level, part of the Report on the national educational system, submitted on an annual basis to the Parliament by the Minister of Education. This mechanism ensures a coherent approach at national, regional, county and provider level for the development of the TVET system.

**Keywords:** quality assurance, national framework, self assessment, validation

## Introduction

The processes of transforming society in the last years, as well as the transition towards post-industrialism and a knowledge-based society represent challenges for the TVET system. Altogether, with ensuring access to education and achieving a high level of qualification, the relevance and the quality of training become significant. In this context, quality assurance (QA) in the TVET system gains new meaning and implies new institutional roles.

The measures taken at European level in this field have inspired the developments in Romania, determining in 2003 the development of the National QA Framework for TVET on the basis of the European Common Quality Assurance Framework in VET (CQAF). The National Quality Assurance Framework in TVET consists of the totality of principles, methodologies and instruments through which quality in TVET is assured, at system and provider level. Implemented on a pilot basis for 2 school years, NQAF was fully adopted at national level through Ministry of Education decision starting with 2006 – 2007 school year.

## Targeted problem

In the general context of the pre-university education system, Romanian technical vocational education and training is specific because it provides both academic education and initial vocational training. TVET main goals mirror its double role of education and training – social and economic:

---

<sup>34</sup> In Romania, initial VET is part of the formal system of education.

- ❑ assurance of students personal and professional development so that they are able to become active citizens of their community;
- ❑ assurance of equal chances/equitable access to technical and vocational education, as well as professional development of each student, depending on her/his individual aspirations and learning potential;
- ❑ assurance of professional development chances for each student for achieving qualification and competences in line with existent employment opportunities, and in a life-long learning perspective;
- ❑ assurance of the quality conditions in organization and development of education and training processes in every school providing TVET.

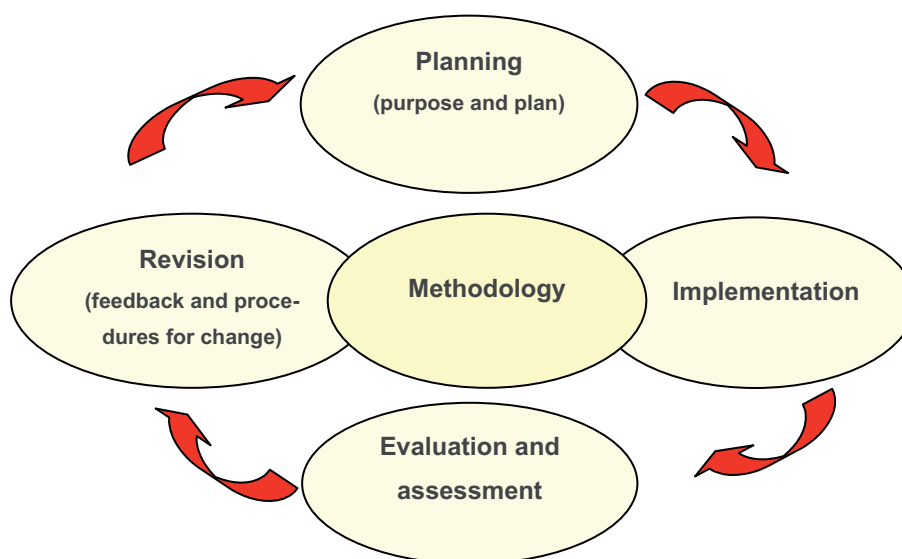
This implies specific quality assurance requirements, to ensure confidence in the quality of academic education and in the validity of the professional certificates. In this context, a multi faceted mechanism for monitoring and evaluation of IVET system was set up, including self assessment of TVET providers, accreditation of training programmes and external monitoring inspection.

## **Methodology**

### **1. General structure**

The aspects focusing on the performance of the TVET providers and of their programmes are structured in 7 areas, called quality principles. The quality principles are based on the European Common Quality Assurance Framework for VET model (Technical Working Group, 2004):

**Figure 1: Quality diagram of CQAF model**



**Table 5: Correlation between the CQAF elements and the NQAF quality principles**

Elements of the European Common Quality Assurance Framework for VET	Quality Principles of the National Quality Assurance Framework for TVET
<i>Methodology</i>	1. Quality management
<i>Planning</i>	2. Management responsibilities 3. Resource management 4. Design, development and revision of learning programme
<i>Implementation</i>	5. Teaching, training and learning 6. Assessment and certification of learning
<i>Evaluation</i>	7. Evaluation and improvement of quality
<i>Revision</i>	1. Quality management 7. Evaluation and improvement of quality

**The Quality principles are:**

**Quality management** – the management develops the mission, vision and values of the organization following consultative processes. The quality manual is developed, including policies and procedures, strategic and operational plans and documentation regarding quality assurance.

**Management responsibilities** – the management gets actively involved in quality assurance of learning programmes. Partnerships with stakeholders are developed and maintained. Relevant information is collected, stored and analyzed regularly and communicated to stakeholders. There is an effective financial management.

**Resource management (physical and human)** – the organization provides students with a safe and supportive environment. The learning spaces are properly equipped and meet the collective and individual needs of students. The resources, teaching methods and the related premises allow the access and active participation of all students. Staff is employed according to clear criteria (minimum standards regarding qualifications and experience) of recruitment and selection, the organization defines job descriptions, evaluation of staff performance is transparent; there is a staff policy, that includes induction procedure and continuous training programmes.

**Design, development and revision of learning programmes** – the organization is permanently concerned with the improvement of learning programmes, to meet the needs of students and staff (internal stakeholders) and of employers and the community (external stakeholders). The learning programmes are centred on the student.

**Teaching, training and learning** – the organization provides equal access to learning programmes and supports all students, preventing any form of discrimination. Students receive complete information about the training provision and benefit from effective counselling and career guidance. Students' rights and responsibilities are clearly defined. Student-centred teaching and training methods are mainly used. Students are encouraged to assume responsibility for their own learning process.

**Assessment and certification of learning** – the organization develops and uses effective processes of assessment and monitoring of learning, in order to support students' progress. Teachers participate regularly in activities of standardization of assessment. Assessment and certification meet national standards and legislation requirements.

**Evaluation and improvement of quality** – the performance of the organization is evaluated and monitored. Following the identification of weaknesses through the evaluation process, improvement measures are developed. These are implemented and monitored, as part of a new cycle of quality assurance.

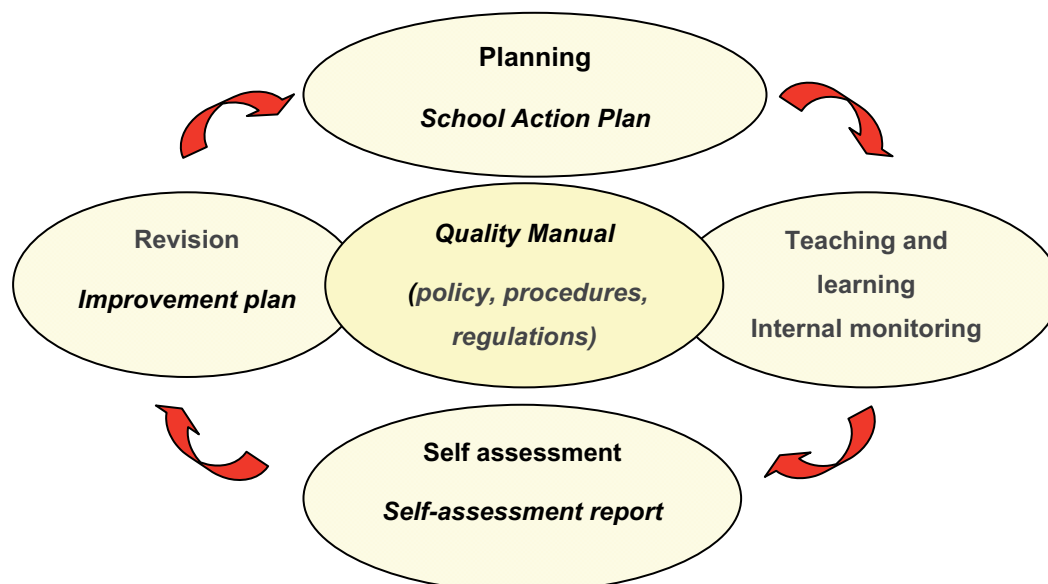
Each quality principle is made up of a number of performance descriptors, based on quality criteria within the European Common Quality Assurance Framework. These describe the activities that lead to quality assurance and improvement of vocational education and training.

## 1. Internal and external QA processes

### 2.1. Internal processes

All TVET providers are required by QA in education Law<sup>35</sup> to have a QA system in place to demonstrate their capacity to plan, internal monitor, evaluate and improve the quality of programmes and the services offered to students. Internal processes to assure quality at TVET provider level are presented in fig.2:

**Figure 2: Quality diagram for the internal processes at TVET provider level**



**Internal processes at TVET provider level include:**

**Planning** – all activities to establish the main objectives in quality of TVET offer, as well as the necessary resources and means, through:

- ❑ *strategic planning*: setting up general objectives regarding QA;
- ❑ *operational planning*: setting up of activities, of necessary resources and means, allocation of terms and responsibilities

<sup>35</sup> QA in education Law (2006) endorses the principles of the Romanian policy for Quality Assurance in Education for the entire national system

In Romanian TVET, quality planning results are included in the School Action Plan (SAP). SAP is a strategic planning document, for 2007 – 2013, developed at TVET provider level as a response in terms of institutional development to the regional and local priorities;

**Internal monitoring** – verifies whether the planned activities are carried out according to the agreed terms and responsibilities and evaluates the teaching and learning process through lessons observations. Internal monitoring identifies in time the school's difficulties in attaining the established objectives and allows identification of appropriate solutions;

**Self-assessment:** the main process in quality assurance, through which TVET providers evaluate their performance based on evidence and produce the Self-assessment report;

**Revision and development of the improvement plan;**

**Development of the Quality Manual**, that includes the quality policy, the procedures for quality assurance, decisions to allocate roles and responsibilities etc.

## 2.2. External processes

**Accreditation:** by QA in education Law it is mandatory for each vocational qualification programme offered by TVET schools. The accreditation requires the covering of two successive stages:

- ❑ **provisional authorization**, that grants the right to carry out the education and training process and to organize the admission to education and training programmes, as the case may be; provisional authorization guarantees that a set of compulsory national authorization standards, mainly input and process requirements, is met;
- ❑ **accreditation**, that additionally grants the right to issue diplomas and certificates recognized by the Ministry of Education and Research and to organize graduation / certification exams; this second stage guarantees that a set of compulsory national accreditation standards, mainly output and outcomes requirements, is met.

**External monitoring:** all TVET programmes are externally monitored on a regular basis by the county inspectorates. Monitoring involves guidance and support to TVET providers in the quality assurance process, quality control to verify the fulfilment of the quality requirements, validation of Self Assessment Reports and approval of the Improvement plans.

## 3. QA Instruments

Quality must be owned by TVET schools; they are responsible for the quality of their own services, and should therefore endeavour to establish a culture of quality at their organisations. Self-assessment against performance descriptors is a key tool in quality assurance

and improvement. As part of the NQAF in TVET a detailed Self assessment manual for schools and an Inspection manual for external monitoring have been elaborated to support these processes.

### **3.1. Self-Assessment Manual**

Achieving quality in the delivery of VET programmes is a long-term process, which requires careful planning and organisation. One of the most important steps towards establishing a quality system is the practice of self-assessment of TVET providers' performance against defined quality standards. The guidance contained in the Self assessment manual serves to assist TVET schools in undertaking their self-assessment process on all aspects of their work as part of a national quality assurance system.

The Self assessment manual outlines the seven quality principles and associated performance descriptors, which form the basis for self-assessment, external monitoring/inspection, and quality improvement. The manual also gives some general guidance on the self-assessment process and improvement planning. It details the requirements for setting up the self-assessment process, gathering and generating evidence, using evidence to make evaluative decisions on the level of performance, and for developing the improvement plan. It includes a Abstract of the self-assessment process, and examples of sources of evidence. The Self assessment manual also explains the internal monitoring procedure.

### **3.2 Inspection Manual for External Monitoring**

The Inspection Manual for external monitoring explains the external monitoring procedure. Full regular inspections are only carried out every 4 to 5 years but TVET schools need guidance and support throughout the self-assessment process. Therefore, they might receive the external monitoring visit at any time during the annual self-assessment cycle, on their own demand or according to the Inspectorates external monitoring programme. The process for the annual validation of VET providers' Self assessment Reports and approval of the Improvement Plans is also explained in the manual. The validation normally does not require a visit; although, it may coincide with an external monitoring visit. Inspectors will read the Self Assessment Report comparing it to the Self Assessment Report and Improvement Plan from the previous year, as well as external monitoring reports (where available). Finally inspectors will decide whether the current Self Assessment Report and Improvement Plan are valid or need to be revised in order to reflect TVET providers' real performance.

The Inspection Manual also includes the seven quality principles, associated performance descriptors as well as the general requirements regarding evidence to make evaluative decisions on the level of performance included in the Self assessment Manual.

#### 4. Stakeholders Roles

One of the most challenging questions is how to best bridge internal and external monitoring and evaluation mechanisms, bringing together top down and bottom up approaches, in an effective and efficient way. This involves clear stakeholders' roles and responsibilities at different TVET level:

- at school level, establishment of the Commission for **Evaluation and Quality Assurance** (CEAC), including staff, students and employers representatives, is mandatory. CEAC coordinates the self-assessment process, elaborates the Self-assessment report, assuring the transparency of the process through an active publication and communication strategy. On basis of the Self assessment report, CEAC develops the Improvement plan, defining areas that need to be changed, as well as the objectives, activities, resources (human and physical) responsible for the Improvement plan.
- at the local/ county level, **School Inspectorates** are responsible for the external monitoring, validation of schools' Self-assessment reports and approval of their Improvement plans. In case of non compliance with quality standards, schools are subject of a special monitoring procedure, designed to offer guidance in the improvement process. External monitoring has led to a major change in the role of inspections, which is currently less about control and more about support and dialogue.

On basis of the validated Self Assessment Reports, an annual *Report regarding the quality assurance of TVET at county level* is elaborated by county inspectorates.

- at central level:
- the **Romanian Agency for Quality Assurance in Pre-university education (ARACIP)** is assigned with the key role in promoting and applying the policies regarding quality of preuniversity education (including TVET). The Agency is responsible for accreditation of all preuniversity providers, for a periodic renewal of accreditation (every five years) and for the development of authorization and accreditation standards.
- the **National Centre for TVET Development** (CNDIPT) supports TVET providers and inspectorates on specific issues regarding the development, implementation and assessment of programmes leading to qualifications, ensuring the TVET specificity and is organising support programmes for actors empowerment and institutional capacity building. CNDIPT also integrates all county reports in an annual *Report regarding the quality assurance of TVET at national level*.



- the **Ministry of Education** (MoE) is responsible for the development of national policies and strategies for the development of education (TVET sector included) and coordinates their implementation at national level. MoE submits to the Parliament an annual *Report on the national educational system*;
- the Romanian **Government** is the statutory body authorized with the power to approve or reject further development QA methodologies for VET. It plays a key role in the general QA process and implementation of the legal and regulatory basis for quality assurance.

## Results and Perspectives for further development

For the described instrument/approach as such

- **Application of NQAF to CVT providers**, in order to assure a unitary QA framework for the whole VET system. The self-assessment and external monitoring process should become part of the CVT management system. This requires the development of quality objectives for continuous vocational education and training, improvement plans and improvement strategies having a structured cyclical character within the CVT providers in the way this is already done for the TVET sub-system.

- **External quality assurance monitoring and evaluation agencies** and bodies (school inspectorates, ARACIP, CNDIPT) should themselves set up their own QA management system and be a subject of external monitoring / regular peer review in order to achieve transparency and visibility of their work, to ensure public control on their activities and also to convince stakeholders in the credibility of their evaluations.

- **Networking between TVET schools** to foster cooperation and ease inter institutional support in the domain. Networking can become a strong driver to a shared quality agenda in times of limited resources and growing demands for investments into good quality of education and training;

- **Collection and dissemination of information on** good practice in quality assurance to fulfil TVET institutions' public roles and responsibilities to provide information on the quality of training opportunities available to their students; alongside marketing opportunity for institutions, such information would enable individual TVET institutions to compare themselves with others and verify that it meets stakeholders expectations in respect of the quality of its provision.

## Potentials and challenges for the use in/for international comparisons

- One particularly promising instrument of quality assurance and development is Peer Review, the external evaluation of TVET schools conducted by peers - experts of equal standing with those being externally evaluated. Current European developments in quality assurance in VET have emphasized the multiple benefits of peer reviews:
- "higher acceptance from the VET providers than other forms of external control like inspection, testing, meta-evaluation etc;
- critical feedback from the peers is accompanied by elements of 'peer coaching/consulting'. The expertise of persons from different institutions and - in the case of transnational reviews - different countries and systems is tapped and a mutual learning process stimulated. Areas of improvement are identified and discussed during the review and thus continuing quality development is enhanced;
- mutual Peer Review often leads to the emergence of quality networks of education providers. Exchange is fostered and transparency increased. Transnational Peer Review contributes to transfer of know-how and experiences and a common understanding in Europe and improves co-operation and comparability in the 'jungle' of quality assurance and development instruments currently in use on the national and the European levels."(<http://www.peer-review-education.net>)

Within the TVET QA Mechanisms, Peer Review can be implemented as a new methodology for ensuring and improving quality. It can be used to complement external monitoring of the quality of VET provision. The current Romanian external monitoring procedure of the TVET QA mechanisms already meets the requirements of the European Peer Review procedure, and is in line with the peer review quality areas, criteria, and indicators.

## References

Technical Working Group, 2004, Fundamentals of a common Quality Assurance Framework 'CQAF' for VET in Europe, p 5 <http://www.peer-review-education.net>

# Gender Sensitivity in Instrument Designs and Evaluation of Vocational Education and Training (VET) Systems

*Theodorah. U.Ogali*

## **Abstract**

Gender issues have continued to occupy the centre burner in vocational education and training (VET) program designs and implementation. It is obvious that there are physiological, biological and sociological differences between the sexes. These remarkable differences have not been regularly accounted for in data collection instruments and interpretations. This paper is designed to proffer practical suggestions on how VET evaluation instruments and processes could accommodate gender sensitivity factors so as to maintain the delicate dynamic equilibrium in combinatorial analysis of all system's elements. In addressing these concerns, the paper examined linguistic, cognitive, psychological and social perspectives of both sexes in wording as well as structuring instruments for qualitative and quantitative data collection. Another aspect of the paper discussed gender sensitivity in interpretation of collated data for VET systems evaluation. In the final analysis, all dimensions of gender sensitivity will prove catalytic in effective VET systems evaluation.

**Keywords:** Gender sensitivity, Instrument design, Evaluation, VET

## **Introduction**

Boys and girls are born with in built differences that go beyond genetalia. Meaning that the girl-child or boy-child has innate predispositions in reacting to certain happenings more strongly and to learn some things, like words, more easily (Keeton & Baskin 1985). The empirical study of sex differences in many facets of human endowments are quite few compared to those of gender differences because society has prevailed on scientists with overwhelming stereotypes. It was only recently that Burman, Bitan & Booth (2008) empirically established that the activation in the left inferior frontal and fusiform regions of girls was also correlated with linguistic accuracy regardless of stimulus modality, whereas correlation with performance accuracy in boys depended on the modality of word presentation (either in visual or auditory association cortex). Many other sex related differences like strength, psychic, intelligence, the brain (Harasty, 2000) and character have been beclouded by gender stereotyping and accounts for fewer empirical studies. The woman's biology has been used to explain how traditional sex roles have developed and why they are similar in so many cultures. Her vulnerability during pregnancy and lactation sets limits to her achievement. Sometimes, these real and supposed differences have been eulogized simultaneously and without conscious contradiction, however they have served to justify and maintain the woman's inferior status, dictating motherhood role as peculiarly appropriate to the woman (Bolt, Wilson & Larsen, 1979). According to Haralambos and Herald (1980) society is made up of human bi-

ogrammar. Biogrammar being genetically based programme that predisposes mankind to behave in certain ways. For example, the man is more aggressive and dominant than the woman. This is derived from differences between male and female hormones.

Distinct from sex differences, gender refers to the historical and sociological differences between women and men as ascribed by any given society. Gender refers to the socially constructed, rather than the biologically defined sex roles and attributes of females and males. Men and women differ socially and biologically but also share some similarities such as being gregarious and wanting to be valued as individuals. These chequered socio-historical reasons have resulted in women's subordination, inequality in opportunity to a better life and poor implementation of policies meant for empowering women to full partnership at work, home, community and school (Schlafly, 1977, Maduewesi, 2005). In some African families, it has been erroneously argued that the opportunity cost is the girl-child education because parents reap the dividend of the boy-child and the husband's family gain in that of the girl-child (Eboh, 1995). The scenario of gender imbalance is not uniform across the world because the more developed a country becomes the more she creates opportunities to correct the imbalance.

The overt and covert differences between the sexes reflect in their world views. Women are more deeply rooted in the species type (issues relating generally to humanity) than men who are more greatly differentiated and more specifically individualized (Simmel, 1978). At regional, country, school and community levels, women leaders have shown significant differences in leadership styles. It would be misleading to make any generalization implying that women make better leaders than men or vice versa. But the spice engendered in any country, political or school structure that meritoriously rotates leadership positions between men and women cannot be fully described. At the school level, various countries have provided different forms of gender oriented education like single sex schools, gender oriented counselling and vocations. As accommodating as this might seem there is still a yawning gap in many policy development and implementation processes.

Specifically, similar reasons of gender differences have played varying roles in influencing policies on vocational education and training (VET). Certain vocations have been labelled women like nursing, teaching, cosmetology etc while others like carpentry, masonry, and auto mechanics have been labelled for men only. In evaluating VET systems, it is not enough to have women researchers or state hypotheses comparing men and women but the design stage of evaluation process should address gender sensitivity issues. By taking these into account, the evaluation process as well as instrument design would be deemed to be gender complaint. Critical VET systems policy indicators provide detailed data which can point to and guide the nature of interventions which need to be made in order to effect meaningful change. They assess gender differentials in provision, access allocation of resources, par-

ticipation and achievement/impact within the educational system. Gender sensitivity entail asking questions, in relation to men and women, about who is doing what, who owns what, who makes decisions about what and how, who gains and losses by a planned intervention. It also involves examining what is happening within the household and makes linkages with the different levels of the wider society. Gender sensitivity reaches a better understanding on how communities work from the perspective of relationships between men and women. Data collection processes, instrument designs and interpretation of findings would have to account for these forms of gender sensitivity (O'Rourke, & O'Rourke, 2001, Gwartney, 2009, Mickelson, & Harrington, 2009).

### **Targeted Problem**

Vocational Education and Training (VET) system evaluation like many other fields have not always elicited the accurate information from respondents, because their instruments do not take cognizance the psychological, biological and sociological differences between the sexes (Gwartney, 2009, Mickelson, & Harrington, 2009). Vocational education training programmes have traditionally been sex-segregated, channelling males and females into different courses sometimes without reference to personal interest, aptitude and achievement. This situation has resulted to failures of many intervention programmes that were supposedly well articulated. Gender insensitivity in the VET systems' perceptions, attitudes and data collection activities have been the bane of many evaluation exercises.

### **Methodology**

The paper is based on intensive internet-based literature, books and journals searches. Various views of different authors were reconciled with what is practically feasible under evaluation processes. Presentation is made within two major sub headings.

### **Results and Perspectives for further development**

#### *Instrument Designs for Data collection in VET Evaluation and their Gender dimensions*

There are several instruments of data collection in VET systems evaluation. However, the discussion of these instruments is limited to the most commonly used namely questionnaire, focus group discussion and interview guides. The basic focus in gender sensitivity in instrument design for VET systems evaluation is to provide guide to carrying out a gender impact analysis of government policies, plans, programmes, and projects in the formal and non-formal educational sector.

These guidelines identify certain key indicators pertaining to gender analysis of the education sector, and they constitute an attempt to capture and summarize data on these indicators of gender equity. Responses to the questions posed in the instruments will yield the data neces-

sary for analysis of the gender impact of policies, plans programmes and projects. The analysis must both diagnose and prescribe. It should:

1. Identify gender gaps where these exist.
2. Proffer reasons for, as well as identify the implications of these gaps.
3. Recommend key actions that could be taken to ensure progress, and in some cases effect immediate closure of the gaps.

The guidelines and instruments can be used as one-time assessment tools, or they can be used to provide baseline information prior to planning an intervention project or programme; using the tools at the end of the programme or project will allow for a pre/post intervention comparison, and an evaluation of its impact.

The questionnaire is a commonly used instrument for data collection in many quantitative evaluation designs. A well-designed questionnaire is essential to improving response rate and increasing the quality of the data collected (O'Rourke, & O'Rourke, 2001). O'Rourke (2001) described several guidelines of physical appearance of the questionnaire for improving receptivity. Between both sexes, it is obvious that the females are more inclined to beauty and order; such that colourful and carefully arranged questionnaires would have more females' favourable disposition. Gender sensitivity dictates thoughtfulness in arranging the order of questions. Especially in Africa and such other cultures where women are hesitant to discussing personal issues; sensitive questions like age and marital status should be at the end of the items. It is also important to avoid sexist language in wording the questionnaire items. It is no longer fashionable to use 'he' to depict both males and female. Sexist bias can occur when pronouns are used carelessly (American Psychological Association, 2005) and this can ruin a well structured questionnaire.

The focus group discussion (FGD) is a major qualitative data collection method. Focus group discussion is used to explore a topic in depth through group discussion, e.g., about reactions to an experience or suggestion, understanding common complaints (McNamara, 2008). In using this method, the FGD guide is essential if the discussants are to be properly focused. The issues mentioned about wording and sexist language for the questionnaire is also pertinent in developing the FGD guide. However, there are more socio-cultural dimensions of the FGD. Every culture has its observances on how; when and where both sexes can meet and some are quite sensitive in their stipulations on discussing intimate issues; though scope of intimate issues varies widely as you have different cultures. In the specific context of VET systems evaluation, FGDs are used for exploratory discussions, to determine those opinions, attitudes, and knowledge held by the target group that regulate how people behave and the effectiveness of potential intervention strategies and programmes. From practical purposes, the FGD session should be held in an atmosphere that is considered natural and relaxed for the discussants. In some cultures such as India, Africa, and Thailand, indigenously-initiated

focus groups develop without prior planning, as neighbours and friends join in household and reference group interviews (Scrimshaw & Hurtado, 1987).

Interview is another popular qualitative data collection method. Just like the FGD, there is need to develop a good guide for the interviewer. Since interviews could be conducted in close confines between the interviewer and the interviewee, there are many psychological dimensions in such social intercourse. In this regard, the nature of data collected could subtly be influenced by the choice of words, dressing and personality of the interviewee. When decently dressed females go for data collection, the men tend to supply more than the needed information. In very religious communities, stricter observances are demanded in choosing field researchers in order to ensure valid data collection.

### **Potentials and challenges of Data interpretations in VET: socio-cultural and international comparisons**

As vital as data collection is to the evaluation process, data interpretation is another critical issue in gender sensitivity. Social (Presser & Stinson, 1998), religious and psychological biases have been implicated in wrong interpretation of data. It is not fair to either sex to infer that any would be more objective than the other. Suffice it to state that some scholars like Geary, Sauls, Liu, & Hoard (2000) have identified sex difference in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning. This follows that the data interpretation component should be mindful of these differences especially if the intended interventions are to succeed. In addition, educationally and economically empowered may not fit strictly into any mould of depicting women as emotionally fragile.

The serious challenge adopting gender sensitive VET systems evaluation is the cost of employing the experts and taking the extra steps in providing for gender equity. Furthermore, it is common for researchers to presume that issues on gender equity imply placating females. Indeed, gender sensitivity is not restricted to female issues but providing sensible options for both males and females. Often times, the males seldom complain aloud due to male egoism. Any society is deemed developed proportionately to how it sustains the dynamic balance between the sexes. The cost implication makes it difficult for gender sensitivity to be strictly observed by countries with little budgetary provision for VET systems evaluation. It therefore follows that for the sake of near even global development, there should be templates (using a set of gender complaint principles) developed for international uses on gender sensitive instruments.

Conclusively, in order to ensure enduring gender sensitivity in every society, gender considerations should be included in all educational and training curricula. This means providing a curriculum which in content language and methodology meets the educational needs and entitlements of girls/boys and which recognises the contributions of women/men in society and values female/men knowledge and experience.



## References

- American Psychological Association (2005). Concise rules of APA style. Washington DC: Author
- Bolt M., Wilson, A. & Larsen W. (1979). Women's biology and Mankind's Destiny: The population explosion and women's changing roles. In I. Freeman (Ed.) *Women: A Feminist Perspective* (2<sup>nd</sup> edition) California: Mayfield Publishing Company.
- Burman, D. D., Bitan, T. & Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia* 46, 1349–1362
- Eboh, B.C. (1995). Socio-cultural Barriers to Effective Participation of Women in Nigeria's Technological Development: *Journal of Technical Education Review* 4(4) 35-50.
- Geary, D. C., Saults, S. J., Liu, F. & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology* 77, 337–353
- Gwartney, P. A. (2009). A cognitive tool to predict how well respondents perform survey-related tasks. Paper presented at the annual meeting of the American Association For Public Opinion Association, Fontainebleau Resort, Miami Beach, FL Retrieved on November 15 from [http://www.allacademic.com/meta/p16952\\_index.html](http://www.allacademic.com/meta/p16952_index.html)
- Haralambos M. & Herald R.M. (1980). *Sociology: Themes and perspectives*. London: Oxford Press.
- Harasty, J. (2000). Language processing in both sexes: Evidence for brain studies, *Brain* 123, 404 – 406
- Keeton K & Baskin Y. (1985). *Women of Tomorrow*: New York: St. Martins / Marek.
- Maduewesi, E.Y (2005). *Benchmark and Global Trends in Education*. Benin City Dasyilva Influence Enterprises.
- McNamara, C. (2008). Some Major Types of Program Evaluation. Retrieved on November 16, 2009 from [http://www.managementhelp.org/evaluatn/fnl\\_eval.htm](http://www.managementhelp.org/evaluatn/fnl_eval.htm)
- Mickelson, W. & Harrington, J. (2009). Fostering storytelling: A qualitative approach to enhance questionnaire development. Paper presented at the annual meeting of the American Association For Public Opinion Association, Fontainebleau Resort, Miami Beach, FL, <Not Available>. 2009-11-14 [http://www.allacademic.com/meta/p17157\\_index.html](http://www.allacademic.com/meta/p17157_index.html)
- O'Rourke, T. & O'Rourke, D. (2001). The ordering and wording of questionnaire items: Part I *American Journal of Health Studies*. Retrieved on November14, 2009 from [http://findarticles.com/p/articles/mi\\_m0CTG/is\\_3\\_17/ai\\_90301873/](http://findarticles.com/p/articles/mi_m0CTG/is_3_17/ai_90301873/)



O'Rourke, T. (2001). Techniques to improve questionnaire format. *American Journal of Health Studies*. 17(1). 36-38.

Presser, S. & Stinson, L. (Feb. 1998). Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review*, 63 (1) 137 – 145

Schlaflly P. (1977). *The Power of the Positive Woman*. New York: Jove Publications Inc.

Scrimshaw, S. & Hurtado, E (1987). *Rapid assessment procedures for nutrition and primary health care* tokyo. Tokyo: The United Nations University.

Simmel, G. (1978). *The Philosophy of Money*. Boston: Routledge & Kegan Paul.

# The “Learning Transfer System Inventory” (LTSI) as a diagnostic tool of training effectiveness in Spain

*Prof. Pilar Pineda, Carla Quesada and Victoria Moreno*

**Abstract:** Providing an effective training for workers is a key to the progress in the context of the current crisis. There is a validated tool that can evaluate training efficacy, in terms of transfer, the “Learning Transfer System Inventory” -LTSI- from Holton & Bates (1998), which diagnoses the factors that are acting as barriers or facilitators for effective training. Our study measures the factors affecting transfer and training effectiveness in Spain, through the administration of the LTSI to 1.042 trainees. The results of this study allow us to validate the survey in Spanish (SLTSI) and contribute to a better understanding of the transfer process in Spanish firms. They also note the importance of developing the role of the supervisor/boss in the transfer process, and provide concrete strategies that can increase the effectiveness of continuing training in Spain and other Hispanic countries.

**Keywords:** training effectiveness, transfer of learning, LTSI in Spain, evaluation.

## Introduction

Nowadays, organizations recognize the value of training for human resource development and for improving their competitiveness in the market. This is why the volume of investment that is made in the training is high. In the US, organizations invested 134 billion dollars in 2008 to fund the direct costs of continuing training (ASTD 2008). In Europe, the last Cranfield Study Executive Report says that in 2005 the average investment in training was 2.99% of payroll (ESADE, 2005). However, the results of training, in terms of transfer of learning, do not correspond with that investment: training only generates between 10% and 30% of changes in the workplace (Holton, Bates & Ruona 2000; Saks & Belcourt 2006), implying that between 90% and 70% of the investments are lost. This figures are alarming, especially in a context like the current crisis in which it is urgent to maximize investments. It is therefore necessary to examine in depth the results of training, to determine its effectiveness and identify how training can be improved to generate greater changes in the performance of workers and increased profits for businesses.

Since Kirkpatrick's created in the 60s his model of the *4 levels of evaluation*, many researchers and practitioners have attempted to measure the results of the training according to the following levels: *reaction*, *learning*, *behavior*, and *results*. We find that it is very important to evaluate the level of behavior or also called transfer of training (Swanson 1996; Pineda 2002; Burke & Hutchins 2008), because this level refers to the extent of which the training learned is applied to the workplace, and generates improvements in the workers' performance; put it in another perspective, measures training efficacy.

In the literature on evaluation of transfer, the contributions of theoretical models explain the transfer process and the mechanisms that facilitate or hinder transfer, such as the models of Baldwin & Ford (1988), Rouiller & Goldstein (1993), Thayer & Teachout (1995), Noe (1996), Holton (1996, 2005), Lim (2002), Nijman (2006), and Burke & Hutchins (2008).

However, the only global validated questionnaire that exists for measuring transfer factors is the “Learning Transfer System Inventory” (LTSI) from Holton & Bates (1998, 2000). The LTSI has diagnosed the system of factors affecting training efficacy in 17 countries, in which it has been administered to approximately 7,000 people.

The LTSI is based on the idea that performance improvement is affected by 16 factors classified into three types: factors related to the ability of the trainee, factors related to motivation, and factors related to the work environment.

**Table 1: Classification of the LTSI factors according to the type of factor**

LTSI	Factors
<b>ABILITY</b>	1. Content validity 2. Transfer Design 3. Personal Capacity 4. Opportunity to Use
<b>MOTIVATION</b>	5. Motivation to Transfer 6. Learner Readiness 7. Performance Self-Efficacy 8. Transfer Effort-Performance Expectations 9. Performance-Outcome Expectations
<b>WORK ENVIRONMENT</b>	10. Supervisor Support 11. Supervisor Sanctions 12. Peer Support 13. Personal Outcomes – Positive 14. Personal Outcomes – Negative 15. Performance Coaching 16. Resistance to Change

Factors are grouped into two groups: those that refer to training in general in the organization - factors 7,8,9,15 y 16 - and those that refer to the specific training program that it is being evaluated - the remaining factors.

In the latest researches, authors have included a new factor, "intent to transfer" (Ajzen & Driver 1992; Combs & Luthans 2007; Griffeth, Hom & Gaertner 2000; Kirschenbaum & Weisberg 2002; Bogdan 2009) inspired by the Theory of Planned Behavior (Ajzen 1991) that indicates that the intent of a person shows their willingness to act.

### **Targeted Problem**

In this paper we demonstrate the results of the LTSI validation in Spain (SLTSI). The aim of the study, in addition to validating the tool, is to diagnose the efficacy of continuing training in Spain in terms of the training outcomes and factors that may cause a barrier, or can be a facilitator for training efficacy. The results of our study provide a better understanding of the transfer process in Spanish firms and allow us to identify strategies to increase continuing training effectiveness in this context.

### **Methodology**

A total of 1,042 surveys were collected in 11 Spanish companies from various sectors: health care (26.49%); insurance sector (6.72%); messaging (15.26%); training institutions (25.53%); and transport (26.01%). We evaluated a total of 103 training programs on: management and administration (8.93%), personal and relational skills (20.35%), technical skills (42.15%), office automation and computer applications (7.97%), and safety and workplace hazards (20.63%). Most of the trainees occupied a position of technical staff (57.05%), followed by middle managers (13.77%) and unqualified personnel (13.33%). The remaining positions represented in the sample are divided by management and administrative personnel (5.90%), managers (5.46%) and commercial salesmen (4.48%).

The *Learning Transfer System Inventory* (LTSI) survey used in the study is the adaptation of the American survey created by Holton, Bates and Ruona in 1998, to the Spanish language and context. The LTSI uses 55 items to measure the 16 factors in it. LTSI v.3 version also considers another factor: "intent to transfer". Trainees must respond based on a 5-point Likert scale (1 = strongly disagree, 5 = totally agree). We conducted a rigorous cross type reverse translation Spanish-English and English-Spanish (Chen, Holton & Bates, 2005:61).

Construct validation of the SLTSI is made by an exploratory factor analysis, taking as a criterion for retention of factors one Eigen value greater than or equal to unity, and factor loadings equal to or greater than 0.40. Constructs analyzed emerge as expected and the analysis reveals an almost identical structure of factors to the original LTSI, except for two items that require revision.

We conducted a descriptive analysis of the specific factors and general factors of the SLTSI, and multivariate analysis of variance of the results by training program and by the position of the trainee (MANOVA). Finally, we conducted a multiple regression analysis to assess the extent to which factors of the SLTSI affect transfer, using as criterion variable “intent to transfer”. We also measured the correlations between the variables involved in the analysis.

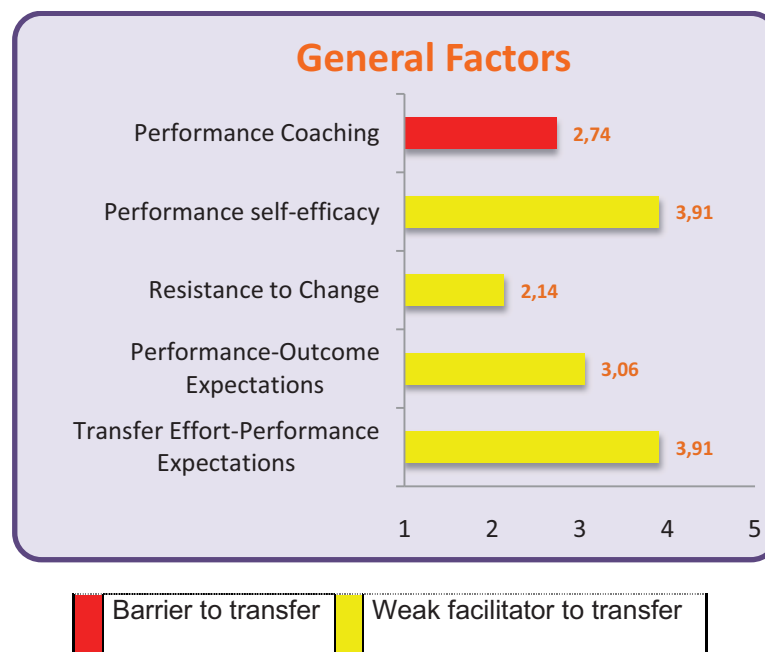
## Results and Perspectives for further development

### *The Learning Transfer System Inventory in Spain*

In all organizations studied, it appears that the **general factors** which determinate transfer, namely those that occur throughout all the training, obtain results between 3 and 4 points on a 1 to 5 scale (Graph 1). Thus, “performance self-efficacy” (3.91 points), “resistance to change” (2.14 points)<sup>36</sup>, and “transfer effort-performance expectations” (3.91), are present in the organizations analyzed but at a level that is not too high, thus acting as facilitators but with little force.

There is a limit to the “performance coaching”. Given the low score obtained by this factor, this is 2.74 points on a 1 to 5 scale, it is a barrier to transfer.

**Graph 1: General Factors of the Learning Transfer System Inventory in Spain**



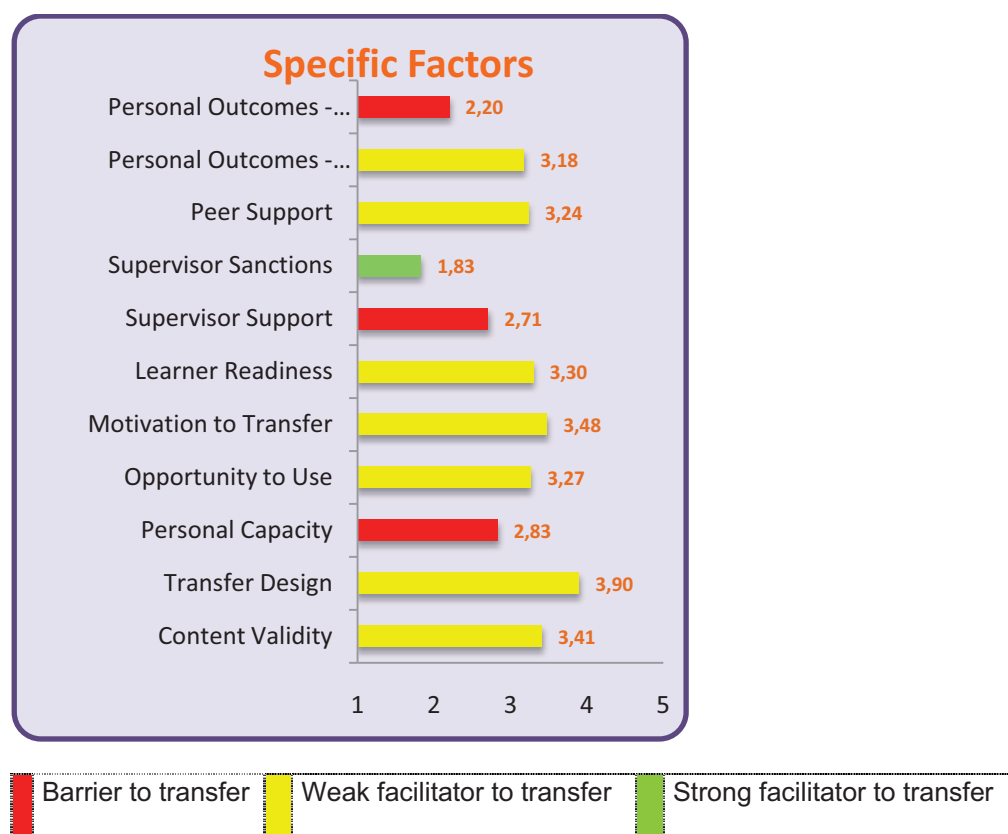
Among the **specific factors**, there is one that stands out due to its positive score: “supervisor sanctions to transfer” (1.83). In the organizations analyzed, the boss does not object to learning transfer. Contrary to this result, the boss shows little “transfer support” (2.71 points),

<sup>36</sup> The factors “resistance to change” and “supervisor sanction” are written in a negative direction, that is to say, they are barriers to transfer. It is expected that they have low scores so they can act as transfer facilitators.

trainees feel they have little “ability, time and energy to implement learning” (2.83), and the fact of not to transfer the training does not have “negative consequences in their work” (2.20). It is positive that there are not supervisor sanctions to transfer, since this would limit any possible implementation of training, especially in a context of instability as the current job market. But the lack of support from the boss and the trainees’ small personal capacity to transfer are aspects to improve, as they can limit the effectiveness of the training investment in the Spanish context.

Other specific factors obtain medium scores and, neither complicate, nor enhance the training efficacy. These factors include the “transfer design” (3.90), “motivation to transfer” (3.48) and “content validity” (3.41), which score slightly above the rest. These results are encouraging because the motivation is the key driver of the training efficacy. Transfer design and content validity depend directly on the training department. Despite the results are being positive, it is for that reason that it is relatively easy to improve them, as its improvement is under the direct control of the training management department.

**Graph 2: Specific Factors of the Learning Transfer System Inventory in Spain**



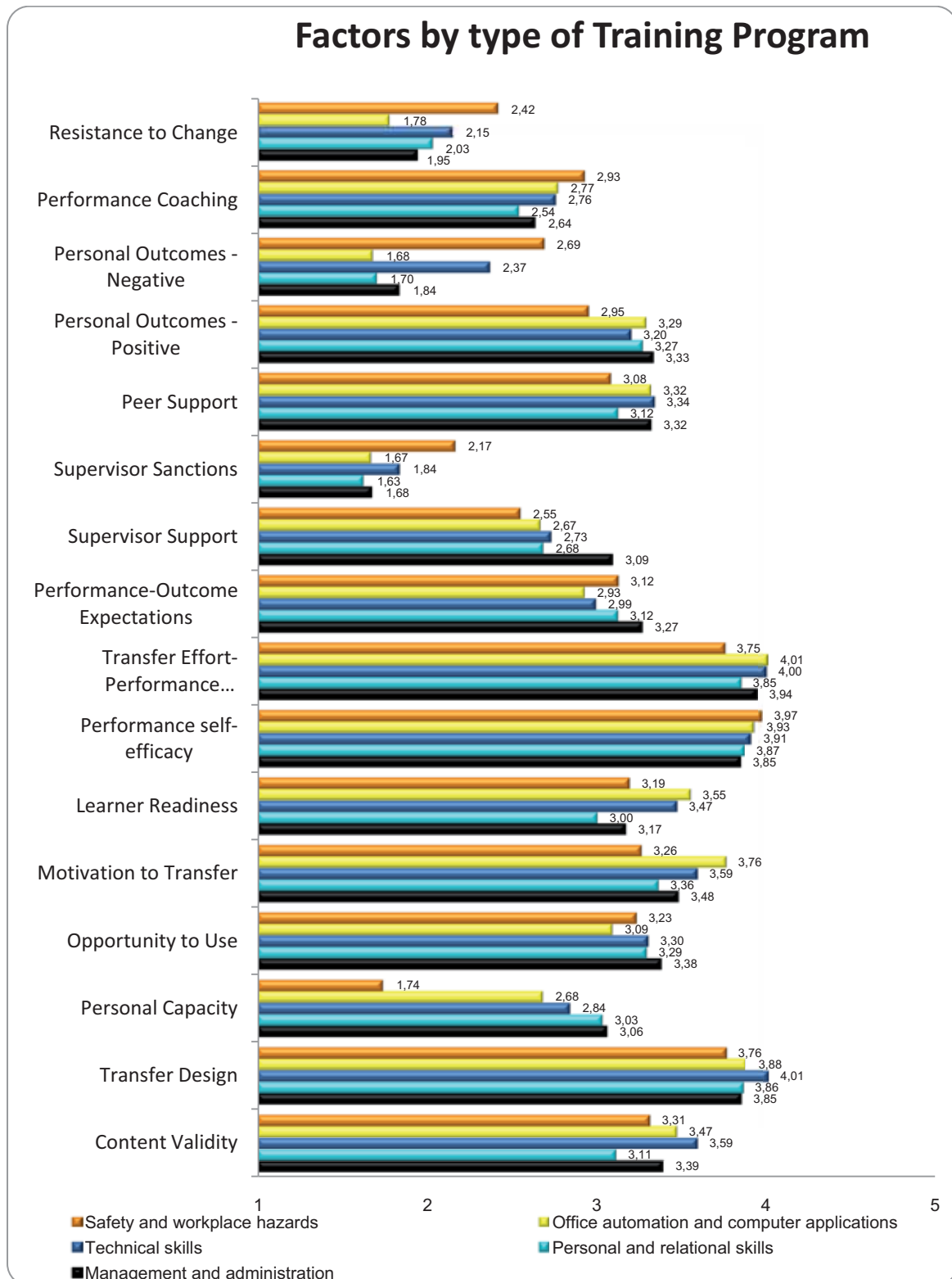
We conducted a multivariate analysis of variance (MANOVA) to determine whether these results vary with the **type of training program**. The results show that there are significant differences among the five types of training in all factors of ability, Wilk’s  $\Lambda = .90$ ,  $F(16, 3159) = 6.8$ ,  $p < .01$ -, in the factors of motivation, Wilk’s  $\Lambda = .85$ ,  $F(20, 3427) = 8.68$ ,  $p < .01$  and in

the work environment factors, Wilk's  $\Lambda = .78$ ,  $F(28, 3718) = 9.62$ ,  $p < .01$ . It is necessary to note that in all three types of factors the multivariate  $\eta^2$  based on Wilk's  $\Lambda$  is very weak, between .03 and .06, in other words, the variance of the factors is not primarily explained by the type of training program. Variances tests were conducted (ANOVA) on the dependent variables using the Bonferroni method, at the level of 0.01. Post hoc analysis were conducted of the significant univariate ANOVAs based on the pairwise comparisons of variables that identify what type of training program has a better transfer system that facilitates training effectiveness.

It is noted that the training programs on safety and workplace hazards present a situation not conducive for the transfer: "resistance to change" is very high, acting as a barrier, the "supervisor sanctions towards transfer" is also high, and so it is the perception about possible "negative personal outcomes", and "motivation to transfer" are lower than in other types of training programs. These results raise serious doubts about the usefulness of training programs on safety and workplace hazards, especially if one considers the large amount of training that is done, being a mandatory training program and because it is important in reducing workplace accidents.

Training programs on technical skills, on most factors, obtain medium scores. The exception would be the "transfer design" and the "content validity" where scores are significantly higher, indicating that this training program is well planned and meets the needs of the workers. Training programs on management and administration enhance "support from the supervisor", "positive personal outcomes" and "opportunity to use learning". Training programs on computer applications has less "resistance to change" and more "learner readiness" and "motivation to transfer", which indicates that the chances of a real transfer are high. In training programs related to personal and relational skills, "performance coaching" is significantly lower, while the other factors scores are situated in the average.

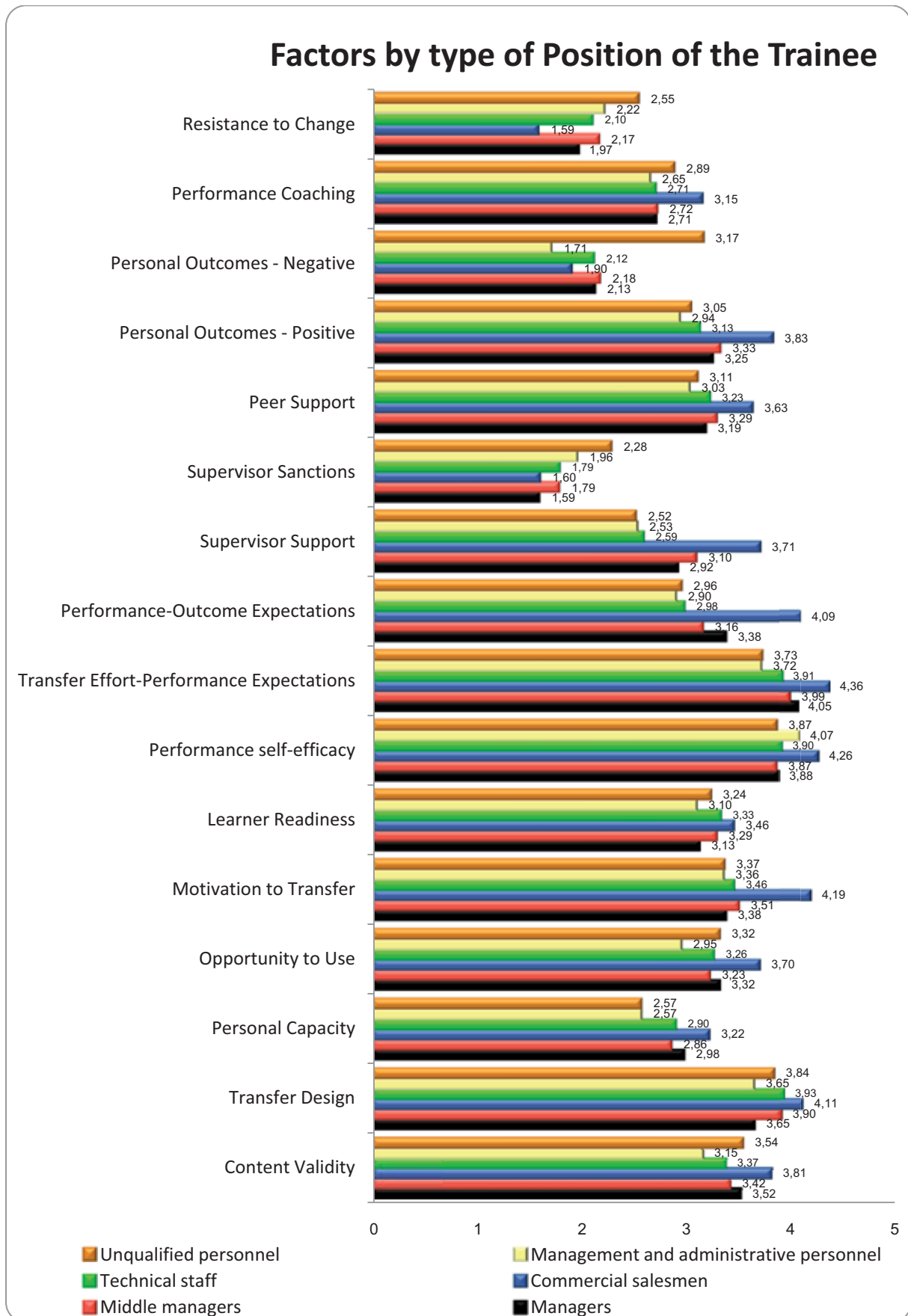
**Graph 3: Global Diagnosis of the factors by Type of Training program**



The multivariate analysis of variance (MANOVA) of the **factors by position** of the trainees shows that there are significant differences in the factors of ability Wilk's  $\Lambda = .92$ ,  $F(20, 3005) = 3.75$ ,  $p < .01$ , in the factors of motivation, Wilk's  $\Lambda = .87$ ,  $F(25, 3363) = 5.36$ ,  $p < .01$ , and in the factors of work environment Wilk's  $\Lambda = .93$ ,  $F(35, 3801) = 7.82$ ,  $p < .01$ , although with a weak  $\eta^2$  multivariate, .02, .03. and .06 respectively.



Graph 4: Global Diagnostic of the factors by Position of the Trainee



In general, the scores of the factors are higher in commercial positions than in other positions, indicating that this group has better conditions to transfer the training. In general, unqualified personnel have good conditions to transfer the training, except in supervisor support, performance self-efficacy, which are below average, and they find a high resistance to change.

The technical staff and the middle managers do not stand out on any factor in relation to other groups, except at their best transfer design and a better learner readiness. Managers also get average scores, which are relevant because of their greater personal capacity to transfer and a high content validity.

It is relevant that personal capacity is significantly lower on the case of unqualified personnel and the administration staff, than in the other groups. From this, we can deduce the hypothesis that the higher the position qualification, the better personal capacity. But this should be verified by other studies in depth.

### **Intent to Transfer in Spain**

“Intent to transfer” in all the organizations analyzed is positive: gets an average score of 3.85 points on a 1 to 5 scale. We think that this factor should have had a much higher score if the training had been highly effective.

To further explore the results, we conducted multiple regression analysis to assess the potential of the 16 factors of the SLTSI as predictors of the “intent to transfer”. Results show that the model of 16 factors correlates significantly with the intent to transfer,  $F(16, 1025) = 77,47$ ,  $p < .01$ , with a high correlation ( $R=0,74$ ), and explaining the 55% of the dependent variance. The model has predictive validity, but it is necessary to refine the tool to enhance it.

All correlations between the factors of the SLTSI and the intent to transfer are significant, except for “personal negative outcomes” (table 2). However, only the partial correlation between nine of the factors of the SLTSI and the intent to transfer is statistically significant ( $p < .05$ ), these factors are mainly motivational factors and ability factors, which suggests that these two blocks of factors are the ones that act as drivers in the prediction of intent to transfer. Still, it is necessary to take into account that the correlation is low, where none reaches the 0.3, and that the various factors of the SLTSI are correlated. Consequently, we can see the need to consider the factors of the SLTSI globally in order to be able to count on all its predictive potential.

**Table 2: Partial and bivariate correlations with the intent to transfer**

Predictors	Correlation between each predictor and the intent to transfer	Correlation between each predictor and the intent to transfer controlling the rest of predictor
Content Validity	,405**	,006
Transfer Design	,517**	,166**
Personal Capacity	,290**	,090**
Opportunity to Use	,522**	,201**
Motivation to Transfer	,516**	,136**
Learner Readiness	,272**	-,062*
Performance Self-Efficacy	,464**	,168**
Transfer Effort-Performance Expectations	,589**	,254**
Performance-Outcome Expectations	,297**	-,117**
Supervisor Support	,296**	,060
Supervisor Sanctions	-,164**	-,051
Peer Support	,484**	,140**
Personal outcomes – Positive	,426**	,060
Personal outcomes – Negative	,004	,029
Performance Coaching	,168**	-,012
Resistance to Change	-,227**	-,040

\*  $p < 0,05$  \*\*  $p < 0,01$

### **Potential and prospects for future development**

The act of having a validated tool to identify the possibilities of transferring training, hence, training programs effectiveness in terms of performance, is enormously useful in a context like the present, where the investment in training and life long learning is empowered (Eurostat 2009). If this tool can also predict transfer, and becomes an alternative measure to the direct evaluation of the results of training in the workplace, the advance for science-related human resources development would be spectacular.

The application of the LTSl in Spain provides a progress in this line of work. The validation of the tool in Spanish opens up numerous possibilities for implementing the SLTSl in other Hispanic territories, thereby improving its cross-cultural dimension and allowing to carry on analysis of the transfer of training to a much more globally international level (Holton 2005).

Results obtained in Spain indicate that, in the organizations of the sample, the factors that determine transfer of training can clearly be improved upon. It appears only one facilitator: the absence of sanctions by the supervisor, along with three other factors that, without being powerful facilitators, are very close to: the performance self-efficacy of the trainee, transfer effort-performance expectations and transfer design. There are significant barriers related to the role of the supervisor, because of that, it is necessary to introduce strategies to improve supervisor coaching towards a trainees' performance and the support offered by the trainees' supervisor to apply what they learn. The role of the supervisor as a mentor, as a key agent of the professional development of their employees, is not seen a lot in the Spanish context, which may explain the result. While in sense, it would be smart to review what is being done in training programs on management and administration and training programs addressed to commercial salesmen, because these factors do tend to get good ratings. These cases can be used as examples to improve supervisor support throughout training.

The predictive capacity of the SLTSl and the intent to transfer in the Spanish context is limited, making it required to further refine the tool and conduct further analysis on the factors that may act as predictors of transfer of training and that can be converted in substitute measures of the actual transfer. This is the goal that guides our future research.

## References

- Ajzen, I. & Driver, B.L. 1992, Application of the theory of planned behavior in leisure choice. *Journal of leisure research*, 24, 207-224
- ASTD. 2008, State of the Industry Report 2008. American Society for Training and Development. Alexandria, USA.
- Baldwin, T.T. & Ford, J.K. 1988, Transfer of training: a review and directions for future research. *Personnel Psychology*, 41, 63-105.
- Bogdan, Y. 2009, Towards a theoretical model of dispositional influences on transfer of learning. 10th International Conference on Human Resource Development: Research & Practice across Europe (Newcastle Business School, June 10.12 2009).
- Burke, L.A. & Hutchins; H.M. 2008, A Study of Best Practices in Training Transfer and Proposed Model of Transfer. *Human Resource Development Quarterly*, 19 (2), 107-125.
- Choi, M. & Ruona, W.E.A. 2008, An Update on the Learning Transfer System. Online Submission, Paper presented at the Academy of Human Resource Development International Research Conference in the Americas (Panama City, FL, Feb 20-24, 2008).
- Combs, G.M. & Luthans, F. 2007, Diversity training: Analysis of the impact of self-efficacy. *Human Resource Development Quarterly*, 18, 91-120.
- Early, D.M., Bryant, D.M., Pianta, R.C., Clifford, R.M., Egan, T.M., Yang, B. & Holton, E.F. III. 1996, The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7, 5-21.
- ESADE. 2005, Informe Cranfield ESADE. Gestión estratégica de Recursos Humanos. Esade, Barcelona.
- EUROSTAT. 2009, Continuing Vocational Training Survey-3  
[http://epp.eurostat.ec.europa.eu/cache/ITY\\_SDDS/EN/trng\\_cvts2\\_esms.htm](http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/trng_cvts2_esms.htm)
- EUROSTAT. 2009, Lifelong learning policy. [http://ec.europa.eu/education/lifelong-learning-policy/doc28\\_en.htm](http://ec.europa.eu/education/lifelong-learning-policy/doc28_en.htm)
- Griffeth, R.W., Hom, P.S. & Gaertner, S. 2000, A meta-analysis of antecedents and correlates of employee turnover. Update, moderator tests, and research implications for the next millennium, *Journal of Management*, 26, 463-488.
- Holton, E.F. III. 2005, Holton's evaluation model: new evidence and construct elaborations. *Advances in Developing Human Resources*, 7 (37), 37-54.
- Holton, E.F., Bates, R.A., & Ruona, W.E.A. 2000, Development of a generalized learning transfer system inventory. *Human Resource Development Quarterly*, 11(4), 333-360.

Kirkpatrick, J. 2005, Transferring learning to behavior. *T + D*, 59 (4), 19.

Kirschenbaum A. & J. Weisberg. 2001, Employee's Turnover Intentions and Job Destination Choices. *Journal of Organizational Behavior*, 22(1), 109-125.

Lim, D.H & Johnson, S.D. 2002, Trainee perceptions of factors that influence learning transfer. *International Journal of Training and Development*, 1, 36-48.

Nijman, D.J.M., Nijhof, W.J., Wognum, A.A.M., & Veldkamp, B.P. 2006, Exploring differential effects of supervisor support on transfer of training. *Journal of European Industrial Training*, 7, 529-549.

Noe, R.A. 1996, Trainees' attributes and attitudes: neglected influences on training effectiveness. *Academy of Management Review*, 11 (4) 736-749.

Pineda, P. 2002, *Gestión de la formación en las organizaciones*. Barcelona, Ariel.

Rouiller, J.Z., & Goldstein, I.L. 1993, The relationship between organizational transfer climate and positive transfer of training. *Human Resource Development Quarterly*, 4(4), 377-390.

Ruona, W.E.A., Leimbach, M., Holton, E.F., & Bates, R.A. 2002, The relationship between utility reactions and predicted learning transfer among trainees. *International Journal of Training and Development*, 6(4), 218-228.

Saks, A & Belcourt, M. 2006, An investigation of training activities and transfer of training in organizations. *Human Resource Management*, 45 (4), 629-648.

Swanson, R.A. 1996, Evaluation systems in HRD. PLS evaluation system: Sales communication case study. *Symposium Evaluation Systems in HRD*, Minneapolis.

Thayer, P.W. & Teachout, M.S. 1995, A climate for transfer model. *Armstrong Laboratory/Human Resources*, Brooks Air Force Base, TX.

## **Peer Learning & Qualitative Methodology**





# Peer Evaluation: Facilitating Changes on the System-Level

Sandra Speer

**Abstract:** Many forms of peer evaluation exist. They can vary from peer learning to peer review and are influenced by various forms of organization, rules and structures. This article focuses on peer evaluation on the system-level with its specific incentives, which facilitate policy learning and transfer. The first section provides a general overview of peer evaluation related to evaluation concepts. The following section addresses three main international peer evaluation mechanisms (OECD peer review systems, the peer learning within the Open Method of Coordination and the African Peer Review Mechanism). Learning from good practices can be facilitated, horizontal and vertical peer pressure fostered, and self-commitment strengthened.

**Keywords:** Peer Evaluation, Peer Review, Peer Learning

## 1. Introduction

Peer evaluation as a mechanism for quality of scholarly journals might be one of the oldest forms of formal evaluation. Also other forms of peer evaluation are widespread in research governance and in other policy fields. Especially in education, e.g. groups of schools or teachers carry out peer evaluations. The term „peer evaluation“ has not been consistently defined and there is no single method but rather a family of methods. The terms „peer review“ and „peer evaluation“ are often used synonymously, as in this text, but the former is more widespread. The unifying concept is a structured feedback given between colleagues or individuals with an equal standing and often based on prior self-evaluations. Peer evaluation is an expertise-oriented approach carried out by individuals, who are usually not trained in evaluation. Their qualifications stem from the field of practice they work in. Many standard textbooks in evaluation do not introduce peer evaluation as a separate evaluation approach. However, there are many different forms, which have to be distinguished according to their purposes, the underlying evaluative positions, the organization, their reciprocity and links to other form of evaluations.

The quality of peer reviews in academia has often been contested (e.g. Cichetti 1991; Gans und Shepherd 1994; Campanario 1998a; Campanario 1998b; Starbuck 2005). Research shows low conformity between the reviews, reviewer biases and dysfunctional incentives. However, several international organizations steer peer evaluation mechanisms. System-level peer evaluations shall lead to better governance and institutions. Who is learning from whom? This contribution cannot give an overview on all the variants of peer evaluation, but shall shed light on different forms of *system-level* peer evaluation with its specific structures and incentives, which differ from peer evaluations on other levels of analysis.

## 2. Theoretical Reflections on Peer Evaluation

Peer evaluation can be located closed to external or to internal evaluation, depending on the evaluative position of the peers and the evaluation criteria. In the one extreme, peers are direct colleagues and the evaluative criteria stem from within the same organization. In the other extreme, peers are from the same field of practice but have a very different institutional background, and additionally might stem from another culture/country. Then the evaluative criteria will be more influenced from outside and at the same time the influence on ensuring that evaluation results have an impact on future behavior will be smaller. This is usually the case in international system-level peer evaluations.

As always, also peer evaluation can have a formative or summative orientation. But two main concepts for the direction of the learning process can be distinguished. In the traditional peer evaluation, the evaluated receives feedback for accountability or development purposes. In another form, also called “peer learning”, the peers receive information on the evaluandum and can learn from that.

Peer evaluations on the micro- and the meso-level are often characterized by direct reciprocity. A teacher might be a peer evaluator to another teacher and vice-versa. The fact that exactly that person will evaluate the peer at another time might induce incentives for a friendly judgment and create conflicts of interest. But reciprocity is also taking place on the systems-level in the framework of country peer evaluations. Leeuw (2002) distinguishes two kinds of reciprocity: the first dimension is the “give-and-take-dimension” on the level of giving information and learning, and the second dimension is the “you-too-me-too-dimension” characterizing the reciprocity on the level of transparency and evaluation that the one institution asks from the other versus, the evaluation and transparency of their own work.

Balancing independence and reciprocity may be also a prerequisite for social capital building and trust between the parties involved in system-level peer evaluations.

Evaluation plays an evermore important role on all various levels within one policy field. Evaluation systems are set up and streams of information are flowing in (Rist & Stame 2008). In recent years standardized testing and achievement measurement is introduced to various educational fields. At the same time self-evaluations on the organizational level as well as on the classroom level are widespread in education, but newer to the VET sector in Europe (Di Battista et al. 2009). In many countries they can be seen as interrelated to the European Quality Assurance Reference Framework for VET (EQARF) (EC 2008). Peer evaluation might extend a self-evaluation or an external evaluation perspective. Sometimes peer evaluation has links to external evaluations on the system-level or to benchmarking systems. Peer evaluation can also connect self-evaluation with an external view by incorporating a kind of auditing function for other self-evaluations, especially when the validity of self-

evaluation is questioned (Kyriakides/Campbell 2004: 29). In that case, peer evaluation can be partly interpreted as meta-evaluation, but is usually not limited to that. The combination of different evaluation approaches is especially useful, when they have complementary goals and perspectives. An external macro evaluation might be accountability-driven and can deliver comparative data, which can also be used for benchmarking (Speer 2001:64).

### **3. A Comparison of Three System-level Peer Evaluation Mechanisms**

Peer evaluations, which are organized in a network of peers and institutions, often have specific rules, which are laid down in guidance documents and clearly defined tasks for the peers. The documents usually give also guide on methods and areas for the peer reviews. In most cases a coordinating body manages the peer reviews and is influencing the quality of the process. Hereunder, three different international peer evaluation mechanisms will be discussed.

#### **3.1 The OECD Peer Review Systems**

The Organisation for Economic Co-operation and Development (OECD) peer review is defined as “the systematic examination and assessment of the performance of a state by other states, with the ultimate goal of helping the reviewed state improve its policy making, adopt best practices, and comply with established standards and principles” (Pagani 2002: 4). Since the 1960’s, the OECD has been carrying out peer reviews in various policy fields and manages secretariats for their organization. The peer-evaluated country provides the information in a form of background analysis or some form of self-evaluation. Usually a peer review team, which always includes three to four experts from other member countries as well as members from other OECD directorates, undertakes a one-week evaluation mission in the country under review. They conduct interviews with national authorities, but also with NGOs, business representatives and researchers. Guidelines for preparing the peer review report have been agreed upon by the relevant OECD committees. The final report is generally approved by consensus of the OECD and the reviewed country. The key concept is mutual accountability and peer reviews are expected to exert their influence by using “peer pressure,” which is a mechanism for soft persuasion or coercion. Dialogue with peer countries, comparisons, public scrutiny and sometimes rankings shall exert pressure on public opinion, national administrations and policy makers (Pagani 2002: 5). Since most countries rarely want to be blamed publicly, peer pressure might be a powerful tool in promoting compliance. So here peer pressure is overtly built in this peer review system. The peer review, as carried out by the OECD on country level, has no direct implications for the evaluated country and no budgetary decisions are connected to it. But the country may be blamed, and the indirect effects may be strong in some exceptional cases. Politicians might then have to justify the policies they pursued or adapt them according to the results of the peer review.

On the OECD level, a large amount of knowledge is accumulated, like for example information about national aid agencies and their respective evaluation systems (Liverani & Lundgren 2007). However, peer evaluation may only lead to learning through public debate in cases where the peer review initiates a policy debate on the national level. There are some reasons why the pressure is often rather weak. The written reports are cumbersome, the recommendations relatively soft, and they are not well known outside the administration. Additionally, the peer review frequency might be too low (Lehtonen 2005: 182). The OECD peer reviews are not well known outside the community; they do not attract as much public interest as e.g. the OECD's "Programme for International Student Assessment" (PISA). The scandal of the weak German PISA results led to a large policy reform including evaluation having an important role within the reform. These assessments and rankings might create more interest in learning from good practices than the peer review as such.

### **3.2 Peer Learning Within the Open Method of Coordination**

The EU introduced peer evaluation as an instrument of the open method of coordination (OMC) within different policy fields. The peer learning is one single instrument of many complementary ones from the OMC, e.g. like the more quantitative oriented benchmarking process (Arrowsmith et al. 2004). The OMC peer review aims at achieving greater convergence towards commonly- agreed EU Member States objectives. Peer review in this context is a voluntary mutual learning process involving the scrutiny of specific policies, programs or institutional arrangements presented as good practice.

Peer evaluation shall promote political innovations by discussing good practice examples, which are suggested by Member States or identified by the European Commission (EC). The chosen practices should be recently evaluated in order to create an evidence-based learning process. Contractors assist the EC in the coordination of the peer review process. The core of the peer review is a two-day workshop arranged in the Member State presenting good practices. Every participating Member State is represented by one governmental official and one independent (research) expert, each from the area under evaluation.

In comparison to the OECD peer review, the OMC peer review is even less known. Because the entire OMC is relatively unknown to the public, peer pressure will not translate into public pressure, nor is this an objective. This peer evaluation takes place on a very voluntary basis, and is intended to reflect the state's own competencies. Practitioners present their best projects or programs, which must have been under evaluation already, and which might be limiting factors for countries with a low level of evaluation culture. Peer learning activities can supplement the knowledge found in written evaluation reports. Since the evaluation, which is supposed to be a basis for the peer visit, has often been carried out for other purposes, additional information, especially for peers from other nations, becomes important. This includes the experiences from the practitioners, from the users, from the political community as well

as from the organizations. The OMC might encourage a better use of evaluations in policy making with the help of the peer learning activities.

Here the main role of the peers is to learn in order to disseminate and share information, which shall induce policy changes in the home country. It is difficult to trace the flow of information from the peers who participated in the peer learning activity on the one hand, to policy formulation or improvements in the home country on the other. Nonetheless, Nedergaard (2006) shows that countries learn the most from best-performing countries and that there is a tendency for countries to learn more from the “most similar” ones. The willingness to implement might depend on the economic and ideological costs of policy change and on the amount of pressure exercised by societal actors. Two other studies (Casey & Gold 2005; Peer Review in Social Inclusion 2005: 42) gave a mixed picture of the transfer to the home countries. The peer learning visits are partly instructive for the participants as interesting programs. The peers need to be powerful within the organizations in their respective home country in terms of being able to initiate a follow-up and implement the new information. Policy makers are not only seeking information from good practices but they also take into account the risk of failure respectively the political costs of implementation, and the consensus required for re-election.

### **3.3 The African Peer Review Mechanism**

The African Peer Review Mechanism (APRM) has been established in 2002 at the inaugural African Union Summit. It is part of the New Partnership for Africa’s Development (NEPAD) and its priorities on poverty reduction, democracy, human rights and corruption. The NEPAD (2003) released a declaration on Democracy, Political, Economic and Corporate Governance and these are also the areas to be reviewed. The APRM shall support standards and practices that promote political stability, economic growth and sustainable development. It is though not focusing on single policy fields as in the peer evaluation mechanisms described above. “The APRM process is designed to be open and participatory. Through a participatory process, the APRM will engage key stakeholders to facilitate exchange of information and national dialogue on good governance and socio-economic development programmes, thereby increase the transparency of the decision-making processes, and build trust in the pursuit of national development goals” (APRM 2003: 2f.). For the African situation it is unique that the civil society has an active role in such a peer evaluation mechanism. Meanwhile 29 of Africa’s 53 states have voluntarily joined the governance monitoring system and twelve of these countries have already undergone all stages of the peer review process (Gruzd 2009). The peer review process consists of five steps. The country under review prepares a self-assessment report and a draft program of action. National stakeholders shall be included in this process, but there are no detailed guidelines for that. Next, a peer team visits the country and conducts interviews with the widest range of stakeholders. A draft report is written and

sent to the APR secretariat. The government of the country under review can comment it, and then send it to the APR panel of eminent persons, which then in turn is presented by the head of state to the APR forum or an AU Summit or a NEPAD steering committee meeting. Due to large differences between the African states is the APRM not based on direct comparisons and rankings.

Ghana was one of the frontrunners to be reviewed and seen as a good example for putting emphasis on civil society engagement (Herbert & Gruzd 2008). The self-assessment process was not led by government officials but was managed through individuals working at independent research institutes. This form of organizing the self-assessment declined political influence and allowed for strong civil society groups involvement. Grimm et al. (2009:2) describe the APRM having a “*club mentality*”, where pressure is preferably exercised behind closed doors. Boyle (2008) reports for South Africa a process in which civil society was not much involved and the media was little informed. The government dominated and drove the APR process without broad public participation. Also Herbert & Gruzd (2008) emphasize the importance of media involvement in the African context.

#### **4. Results and Perspectives for the further development**

Three different peer evaluation systems have been discussed. All of them are oriented towards shared goals, or standards and at learning through good practices. Instead of sanctions the peer evaluation shall generate improvements in policies and institutions. The OECD peer review is partly exerting peer pressure, perhaps more in the circle of peers than in the public debate. The OMC peer learning activities open learning opportunities from good practices, which then can be transferred to the home countries. However, the prerequisites for a transfer are a “window for opportunity” and political entrepreneurs, who are willing to introduce new policies by convincing others that the status quo policies are not optimal. The African Peer Review Mechanism could show success in case the civil society has been strongly involved. This involvement may lead to process use of the peer review, empowerment of the stakeholders and is strengthened by information of the public. For all cases, it is difficult to attribute causality to changes due to the peer evaluations because many other factors will always largely influence policy shifts. However, some incentives could be identified.

The success of peer evaluation on the system-level can be highly influenced by the peers during the process as well as afterwards, the participation of stakeholder groups and the media coverage. Peer evaluation leads then to learning through a public debate and can even build up pressure from within the country (vertical accountability). But also external peer pressures may be strong elements for policy shifts (horizontal accountability). The peer pressure can occur either in form of formal recommendations or in form of comparisons and rankings. Self-commitment is strengthened by these forms peer pressure. The weakest form of peer evaluation is the peer learning, which only relies on the transfer of single peers. These



peers are dependent on other actors in their home country and cannot be seen as independent agents for transferring good practices. Future research should further analyze links between peer evaluation and other forms of evaluation, which can strengthen knowledge utilization and prevent peer evaluation from being a symbolic activity. In which ways can system-level benchmarking and system-level peer evaluation be combined? In which ways can system-level benchmarking and meso-level peer evaluation be combined?

## References

APRM Secretariat (2003), Guidelines for countries to prepare for and to participate in The African Peer Review Mechanism, November 2003.

Arrowsmith, J., Sisson, K. & Marginson, P. (2004), What can 'benchmarking' offer the open method of co-ordination? In: *Journal of European Public Policy* 11:2 April 2004, 311–328.

Boyle, B. (2008), Making the News: Why the African Peer Review Mechanism didn't, SAIIA Occasional Paper, No. 12, South African Institute of International Affairs, September 2008,

[http://www.saiia.org.za/images/stories/pubs/sops/saia\\_sop\\_12\\_boyle\\_20080925\\_en.pdf](http://www.saiia.org.za/images/stories/pubs/sops/saia_sop_12_boyle_20080925_en.pdf) [acce

Campanario, J.M. (1998a), Peer Review for Journals as It Stands Today - Part 1. In: *Science Communication* 19(3), 181-211.

Campanario, J.M. (1998b), Peer Review for Journals as It Stands Today - Part 2. In: *Science Communication* 19(4), 277-306.

Casey B. H. & Gold, M. (2005), Peer Review of Labour Market Policies in the European Union: what can countries really learn from one another? In: *Journal of European Public Policy*, Vol. 12, 23-43.

Cichetti, D.V. (1991), The reliability of peer review for manuscript and grant submissions: A crossdisciplinary investigation. In: *Behavioral and Brain Sciences* Vol. 14 pp.119-135.

Di Battista, G., Palomba, L. & Vergani, A. (2009), Evaluating the Quality of Italian Local Vocational Training Systems. In: *Evaluation*, Vol. 15, No. 2, 185-203.

ENQA-VET (2009), Policy Brief: ENQA-VET Peer Learning Activities,  
<http://www.engavet.eu/documents/PolicyBriefPLA000.pdf> [accessed November 09]

European Commission (EC) (2008), Recommendation of the European Parliament and of the Council on the establishment of a European Quality Assurance Reference Framework for Vocational Education and Training, SEC (2008) 440, SEC (2008), 441, Brussels.

Gans, J.S. & Shepherd, G.B. (1994), How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists. In: *Journal of Economic Perspectives* , Vol. 8, 165-179.

Grimm, S., Nawrath, K., Roth, R., Triebel, S. & B. Utz (2009), The African Peer Review Mechanism (APRM) as a Tool to Improve Governance?, DIE, Bonn.

Gruzd, St. (2009), African Peer Review: A Progress Update, in: Growth, November 2009, [www.growth.co.za](http://www.growth.co.za) [accessed November 2009]

Herbert, R. & Gruzd, S. (2008), The African Peer Review Mechanism: lessons from the pioneers, South African Institute of International Affairs, Johannesburg.

Kyriakides, L. & Campbell, R.J. (2004), School Self-Evaluation and School Improvement: A critique of Values and Procedures. In: Studies in Educational Evaluation, Vol. 30, 23-36.

Leeuw, F. L. (2002), Reciprocity and Educational Evaluations by European Inspectorates: assumptions and reality checks. In: Quality in Higher Education, Vol. 8, No. 2, 137-149.

Lehtonen, M. (2005), OECD Environmental Performance Review Programme: Accountability (f)or Learning? In: Evaluation, Vol. 11, No. 2, 169 - 188.

Liverani, A. & Lundgren, H.E. (2007), Evaluation Systems in Development Aid Agencies. In: Evaluation, Vol. 13, No. 2, 241-256.

Nedergaard, P. (2006), Which Countries Learn from Which? A Comparative Analysis of the Direction of Mutual Learning Processes within the Open Method of Coordination Committees of the European Union and among the Nordic Countries. In: Cooperation and Conflict, Vol. 41, No. 4, 422-442.

New Partnership for Africa's Development (NEPAD) (2003), Declaration on Democracy, Political, Economic and Corporate Governance, <http://www.aprm-international.org/> [accessed November 2009]

Pagani, R. (2002), Peer Review: A Tool for Co-operation and Change, OECD SG/LEG (2002)1.

Peer Review in Social Inclusion (2005), Final Report: Implementation of the Peer Review Programme 2005, Main Results of the Peer Reviews 2004 and 2005. [www.peer-review-social-inclusion.eu](http://www.peer-review-social-inclusion.eu) [accessed June09]

Rist, R. C. & Stame, N. (Eds.) (2006), From Studies to Streams. New Brunswick: Transaction Publishers.

Speer, S. (2001), Evaluation und Benchmarking – Ein Methodenvergleich am Beispiel der betrieblichen Personalarbeit, in: Geise, W. (Ed.), Ökonomische Bildung zur Bewältigung von Lebenssituationen, Bergisch Gladbach: Verlag Thomas Hobein, 51-67.

Speer, S. (2007), Evaluation Report of the Leonardo da Vinci- Project Peer Review in Initial Vocational Education and Training, Vienna, [http://www.peer-review-education.net/TCgi/Images/peerreview/20071218155137\\_Evaluation%20Report%20Peer%20Review.pdf](http://www.peer-review-education.net/TCgi/Images/peerreview/20071218155137_Evaluation%20Report%20Peer%20Review.pdf) [accessed November 2009]



Speer, S. (2010), Peer Evaluation – The Powerful Peer?, in: Eliadis, P., Furubo, J.-E. & Jacob, St. (Eds.), *Evaluation: Seeking Truth or Power?*, New Brunswick: Transaction Publishers [forthcoming].

Starbuck, W.H. (2005), How Much Better Are the Most-Prestigious Journals? The Statistics of Academic Publication. In: *Organization Science*, Vol. 16, No. 2, 180-200.

# Qualitative Techniques for Evaluating Vocational Education and Training (VET) Systems

*Dr. Benjamin A. Ogwo*

## **Abstract**

Many evaluation processes in vocational education and training (VET) entail translating thoughts and feelings (qualitative data) into quantitative data which invariably result in loss of important facts/details. The use of qualitative techniques ameliorates this challenge in VET systems' evaluation because it preserves the chronological flow of events; they are very fluid and adaptable. Agreeably, research designs and instrument development for collection and collation of qualitative data could sometimes be daunting. Fortunately, there are manual and computer aided qualitative data analysis techniques for evaluating different aspects of VET systems relative to behaviour engineering, program development, change management and policy analysis which analysts are unaware or underutilizing. This paper presents literature on best practices as well as practical suggestions on using different qualitative techniques for VET systems' evaluation in order to mitigate the loss of valuable information and poor perception associated with conduction qualitative evaluation.

**Keywords:** VET systems, Qualitative techniques, Systems evaluation, Policy analysis

## **Introduction**

Global skills dynamics has occasioned increased investments in vocational education and training (VET) which justifies the need for more effective VET systems evaluation. VET is generally described as that aspect of general education geared towards skills acquisition for gaining and retaining employment. UNESCO & ILO (2002) explained that VET is understood to be: a means of preparing for occupational fields and for effective participation in the world of work; an aspect of lifelong learning and a preparation for responsible citizenship and a potent for facilitating poverty alleviation. Evidences abound globally about the expanding scope of VET, examples include Europe's increased attention on recognition of prior learning (CEDEFOP, 2009), the United States of America's renaming it career and technical education (CTE) that includes degree programmes such engineering, architecture etc (Scott & Sarkees-Wircenski, 2008; Fretwell, 2009). In some African countries, there have been some inroads at providing interventions for the informal VET programmes (Ogwo, 2008, ILO, 2000). Increasing scope of VET, occupational changes, mobility/migration of labour and general unemployment justify the huge investment on VET. In the same token, evaluation of VET systems has equally gained more prominence in recent times since practitioners and funders would wish to ascertain their returns on investment: time, efforts, materials, and money.

VET systems evaluation can be done at various levels namely: school, district, state, national, regional and international. What constitutes a given VET system depends on the

evaluator's delineation criteria because what is regarded as a system in one instance could become a subsystem in another instance. For example the Federal VET system in Germany is a subsystem within European Union. Indeed, a system is a set of interrelated and sometimes directly interdependent things or components/elements that work severally and jointly within a given setting for the attainment of a definite goal (s) (Ogwo, 1996). Thus in evaluating a national VET policy issues an evaluator could discern its elements to include, Laws on VET, region linkages, national/state level administrators, school level administrators, teachers, other school personnel, and students. This approach is applicable to other forms of VET system evaluation such as program implementation/development, change management and behaviour engineering (instructional system). Hence in conducting VET systems evaluation, the evaluator has to select the most suitable model based on the purpose of the evaluation, relationship between the system's major elements as well as the supra system and the techniques for data collection (Gall, Gall, & Borg, 2007).

There are two fundamental techniques of data collection/analysis for VET systems evaluation namely quantitative and qualitative but this does not connote mutual exclusivity regarding the use of these techniques. In fact there is ample blending of one technique's features over the other in many instances of conducting evaluation studies. The mix methods designs have also been reported in literature (Brewer, 2001, Barbour, 1998). Qualitative evaluation/research is an umbrella term denoting research methods that use language data (Polkinghorne, 2005) as against quantitative methods that use mostly numeric data. The cardinal difference between quantitative and qualitative techniques in VET evaluation is that the former uses numbers to explore relationships while the later uses word/language and visual data. Hence the essential evaluation skill in quantitative technique is numeracy while qualitative technique is literacy. The justifications for using qualitative techniques in VET systems evaluation include but not limited to the following: they are very fluid, adaptable, preserve the chronological flow of events, and entail less numerical capacity of field researchers and are most useful where there is language translation barrier between the evaluators and the population being studied. Qualitative techniques enable in-depth study of complex issues in their natural settings through any of its approaches: case study, ethnography, phenomenology and biography.

### **Targeted Problem**

In spite of numerous reasons for the use of qualitative techniques in evaluating VET systems, many practitioners have been weary in using them. A couple of issues have been implicated on the use of these techniques namely, being too rigorous (Shuttleworth, 2008), data collection/collation processes are not truly transparent (Shenton, 2004), restricted to the study of small sample size (Gall, Gall, & Borg, 2007) and the subjective involvement of the researcher in the research process (Neill, 2007). The prevailing scenario is the under-

utilisation of qualitative techniques and prodigal use of quantitative techniques in VET systems analysis. Thus useful information is lost especially in translating words to figures and loss of natural sequence of events. As damaging as these negative comments may seem, the state of knowledge in qualitative techniques provide compelling reasons for revisiting earlier reached conclusions. The use of information communication technology (ICT) has positively influenced every facet of human endeavour including the use of qualitative techniques. There are many free, efficient and qualitative softwares that can analyze texts, picture and derive trends on qualitative data (Barry, 1998, Duriau & Reger 2004). Competency-based training is required in order to use them as also applicable in other quantitative equivalents. At the manual level of qualitative analysis, there have been remarkable improvements in procedures and reporting techniques (Barbour, 2001). Thus the paper presents literature on best practices as well as practical suggestions on using different qualitative analysis techniques for VET systems' evaluation in order to mitigate loss of valuable information and poor perception associated with conducting qualitative evaluation.

## **Methodology**

The paper was written based on literature review on current best practices on the use of qualitative techniques. Overviews on manual and computer-aided qualitative techniques were presented against the backgrounds of three VET systems evaluation issues such as policy analysis, programme development and behaviour engineering. Practical suggestions were equally presented in relation to various challenges in adopting qualitative techniques.

## **Results and Perspectives for further development**

### *Overview of Qualitative Techniques for Manual and Computer Applications in VET*

#### *Evaluation*

Numbers data representation is not superior to those in words. In most quantitative evaluation, the exercise commences in numbers and ends up in words (rights and wrongs, haves and have not's). For most evaluation processes within vocational education and training (VET), practitioners think, feel in words, and not in numbers. Routinely, they are required to translate these thoughts and feelings (qualitative data) into quantitative data which invariably result in loss of important facts/details. Forms of VET qualitative data includes transcript that is in text format; audio data which may be in form of recorded interview or events; video data that may comprise motion pictures with audio records; and visual data which includes pictures, and maps. However, the recent developments in ICT have further narrowed down the differences between the two techniques. Leedy & Ormarod (2010) listed issues to consider before opting for qualitative data techniques namely description, interpretation, verification and evaluation. Description refers to the evaluation objective entailing the revelation of the nature of relationships, cultural settings, and processes. Many training and systems' phenomenon (new concepts/theoretical perspectives) cannot be readily interpreted using quanti-

tative approach and as such needs qualitative verifications within real-world contexts. The evaluation factor in choosing qualitative technique relates to the advantageous position of the researchers in judging the effectiveness of policies, changes and practices from the perspective of end users and in unthreatening research environment. Shuttleworth (2008) listed focus group discussion, interview schedule, and documentary analysis as methods of collecting qualitative data. He listed these wide range of documents that are analyzed in VET systems evaluation to include course and module outlines, course materials, staff development documents, inspection reports, examiners' reports, minutes of meetings, memoranda, letters, diaries, students' journals, administrative records, and management information systems.

From practical perspectives, the use of interview, focus group discussion, and documentary analysis should be guided by evaluator's skills, type of interpretation desired, audience for which the report is meant, and the available type of software in case of computer analysis. In order to ameliorate researcher subjectivity during interviews and focus discussions, a good interview/discussion guide should be developed for the exercise. Interviews and focus group discussions are highly interactive socio-intellectual exercises and developing a good guide is as difficult as developing attitudinal scales used in quantitative analysis. The use of research team is another functional way of reducing researcher subjectivity since team members compare notes on different sets of observations/exercises. In designing interview/discussion guides, extreme care should be exercised in wording the interview questions (Polkinghorne, 2005) in relation to cultural ethos, gender sensitivity and semantics. The cardinal technique in documentary analysis is the use of content analysis technique. Concept mapping, theme selection and relationships schema should be done at the commencement of qualitative evaluation. These activities would afford the study thorough pre-data collection and interpretation readiness and establish some measure of internal reliability/validity. Other measures to improve the quality of qualitative evaluation include: giving adequate description of assumptions, methods and data analysis (Mays & Pope, 1995).

Computer softwares have alleviated much of the difficulties of manual manipulation of VET quantitative and qualitative data without compromising the quality of research findings. The qualitative softwares can be categorized into mono and multimedia qualitative data analysis software. The mono-media qualitative softwares handle single data set – textual, audio, video or visual kind of data, while the multimedia softwares can manipulate two or more kinds of qualitative data simultaneously. Some multimedia qualitative softwares have the advantages of performing two or more of the following tasks – translating data from one form to another, importing and analyzing rich or plain text documents, audio files, video files, and digital photos. For example, Dragon Naturally Speaking converts voice to text and vice versa while NVivo, has the advantages of converting six languages including English, French, German, Spanish, Simplified Chinese and Japanese, working with(out) transcripts, analyzing

materials straight from audio and video files, importing and analyzing PDF files, working with materials in many languages, displaying project information in graphic form, merging separated projects and identifying who did what by each team member, and sharing of files and findings (QRS International, 2007) When qualitative data are in multimedia form, multimedia softwares become ideal for the analysis of multimedia data. Table 1 below shows some qualitative softwares, their scope, developer, website and technical requirements.

**Table 1: Some Qualitative softwares, developer, scope, website and technical requirements**

<b>SOFT-WARE</b>	<b>Type of Data (SCOPE)</b>	<b>DEVELOPER</b>	<b>WEBSITE</b>	<b>PLAT FORM</b>
Dragon Naturally Speaking	Coverts voice to text and back	Dragon SDK	<a href="http://www.nuance.com/naturallyspeaking/products/default.asp">http://www.nuance.com/naturallyspeaking/products/default.asp</a>	Windows 2000, 2003, XP, Vista
Aculine	Digital Automated Scoring	Harpe & Associates Ltd.	<a href="http://www.harpe.ca/">http://www.harpe.ca/</a>	Windows PC
AnSWR Freeware	Primarily text analysis	(CISSS)	<a href="http://www.cdc.gov/hiv/software/answer.htm">http://www.cdc.gov/hiv/software/answer.htm</a>	Windows: 95,98, NT, 2000, ME, or XP
Atlas.ti	Text, digital Audio and Video Analysis	Thomas Muhr	<a href="http://www.atlasti.com/index.php">http://www.atlasti.com/index.php</a>	Windows 95, Windows 98, Windows NT 4.0
CDC EZ-Text Freeware	Primarily text analysis	Centers for Disease Control	<a href="http://www.cdc.gov/hiv/software/ez-text.htm">http://www.cdc.gov/hiv/software/ez-text.htm</a>	Windows 95,98, NT, 2000, ME, or XP
ELAN	Max Planck Institute for Psycholinguistics Developer Site:	Max Planck Institute for Psycholinguistics Developer Site:	<a href="http://www.la-mpi.eu/tools/tools/elan">http://www.la-mpi.eu/tools/tools/elan</a>	Windows 2000/XP/Vista, MacOS X, Linux
HyperRE-SEARCH 2.8.2	Text, images, audio, video, code maps,	Research Ware Inc	<a href="http://www.researchware.com/">http://www.researchware.com/</a>	MacOS, Windows

	auto coding,			
INTERACT	Multimedia Video	Mangold Software & Consulting	<a href="http://www.behavioral-research.com">www.behavioral-research.com</a>	Windows ME, NT, 2000, XP
NVivo 8	Text audio and video Multimedia	QSR International	<a href="http://www.qsrinternational.com/products_nvivo_features-and-benefits.aspx">http://www.qsrinternational.com/products_nvivo_features-and-benefits.aspx</a>	Windows XP, Vista Window 7

No software can do all the analysis for the evaluator no matter how sophisticated. For example, it cannot analyse jokes within an interview. The qualitative evaluator has to do serious concept mapping, theme selection and objectively interpret the data. With the advent of softwares in multiple languages, conducting interview as well as transcribing the information has become a lot easier. Yet first speakers of the language should be charged with actual interpretation of the audio data. With careful planning and diligent implementation, research and bias in qualitative research can be mitigated with the use of computer applications.

### **Qualitative Frame work for Policy Analysis, Programme Development and Change Management in VET Systems Evaluation**

Effecting any major policy change is like relocating a cemetery. Many memories, sentiments objective reasons, different languages, living and dead people are associated with existing cemeteries and policies. For such complex issues as policy development and behaviour management in VET, numbers alone cannot capture the whole attributes of the picture: colour, texture, smell that are vital in analyses. In policy analysis, the systems approach and cost effectiveness analysis are vital tools that have proven capabilities that rely reasonably on the use of qualitative techniques. The potency and sophistication of human language in sustaining human civilization through education, behaviour engineering and policy formulation (Ritchie & Spencer, 2000) is well document. It becomes preposterous to have language/words downgraded in data collection and analysis. The least that can be asked is to conduct VET policy evaluation, programme development and change management processes with the most sophisticated research design, software, and objective interpretation that would capture the feelings, perception and objective dispositions of the population.

### **Potentials and challenges for the use of Qualitative Techniques in VET: International Perspectives**

Globalization is challenging the earlier held notion as well as characterization of nation states because many regional bodies are assuming more nation-like responsibilities. With this



paradigm shifts, comes ethnocentric pride; language war and national loyalties. Thus many of the underlining issues to VET systems policy dynamics at the regional levels would centre on which language the data was collected, how reflective of the constituting nation were the evaluators, and did the process respect all cultural idiosyncrasies. It is basically qualitative techniques that can aptly meet these subtle but important challenges in providing regional policies. In effect, data can be collected in as many languages as needed, in the socio-cultural context of the sample and very large samples can be handled using appropriate software. The European Centre for the Development of Vocational Training (CEDEFOP), UNESCO, ILO and other multi-national organizations provide good qualitative framework by publishing their reports in need languages.

Some policies and programmes associated with VET systems failures in Africa, Asia and Latin America could in part be blamed on the purely quantitative approach as well as “classical western world views” adopted in their development process. By impugning that qualitative technique is too subjective and grants recipients too naive to contribute to VET systems policy process would be good recipes for systems failures. Numbers cannot well describe human feelings and the ownership of policies is directly proportional to the extent the process is humanized through qualitative data collection techniques. As many developing countries continues to battle literacy and numeracy of the illiterate population, their literate workforce are more recently saddled with e-illiteracy (Ogwo, Onweh and Nwizu, 2010) and this aspect of workforce development would depend on qualitative techniques to develop suitable softwares for these countries. Invariably, qualitative techniques would not answer all research questions but would provide unquantifiable clues to most questions.

## References

- Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *British Medical Journal*, 322 (7294), 1115 – 1117 Retrieved July 31, 2008 from <http://www.bmj.com/cgi/content/full/322/7294/1115>
- Barry, C. A. (1998) 'Choosing qualitative data analysis software: Atlas/ti and nudist compared' *Sociological Research Online*, vol. 3, no. 3, Retrieved on August 17, 2009 from <http://www.socresonline.org.uk/socresonline/3/3/4.html>



- Brewer, E. W. (2001). Mixed Method Research Designs. In E. I. Farmer & J. W. Rojewski (Ed.), *Research Pathways: Writing professional papers, theses, and dissertations in workforce education*. Lanham, MD: University Press of America. Center for Disease and Control (CDC), (2006). Retrieved on November 9, 2009 from <http://www.cdc.gov/hiv/topics/surveillance/resources/software/answer/use-qda.htm>)
- Duriau, V. J. & Reger, R. K. (2004). Choice of Text Analysis Software in Organization Research: Insight from a Multi-dimensional Scaling (MDS) Analysis. *Journées internationales d'Analyse statistique des Données Textuelles (JADT)* Reterieve on August 17, 2009 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.124.2237&rep=rep1&type=pdf>.
- European Centre for the Development of Vocational Training (CEDEFOP), (2009). The development of national qualifications frameworks in Europe. Retrieved on October 26, 2009 from [http://www.trainingvillage.gr/etv/Information\\_resources/Bookshop/publication\\_details.asp?pub\\_id=562](http://www.trainingvillage.gr/etv/Information_resources/Bookshop/publication_details.asp?pub_id=562)
- Fretwell, D. (Oct., 2009). Exporting U.S. CTE goods and services. *Technique*, 84 (7), 16 – 21.
- International Labour Organization (ILO) (2000). Employment and social protection in the informal sector. Retrieved on November 08, 2009 from <http://www.ilo.org/public/english/standards/relm/gb/docs/gb277/pdf/esp-1-1.pdf>
- Gall, M. D., Gall, J. P. & Borg, W. R. (2007). Educational research: An Introduction (8<sup>th</sup> edition. Boston: Pearson Educational, Inc.
- Leedy, P. D. & Ormrod, J. E. (2010). *Practical research planning and design (9<sup>th</sup> edition)*. Boston: Pearson Educational. Inc.
- Mays, N. & Pope, C. (July, 1995). Qualitative Research: Rigour and qualitative research. *British Medical Journal (BMJ)*, 311, 109 – 112.
- Neill, J. (2007). Qualitative versus quantitative research: Key points in a classic debate. Retrieved on August 27, 2009 from <http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html>
- Ogwo, B. A. (2008). Sustaining the Mechanic Village (out-of-school Vocational) Schools in South Eastern Nigeria Using Information Communication Technology. *Network for Policy Research Review and Advice on Education and Training (NORRAG) News*, (40), 79 – 81.
- Ogwo, B. A., Vincent C. Onweh, V. C. & Nwizu, S. C. (2010). globalized workforce development and responsible global citizenship through e-literacy capacity building programs for low income african countries. In B. M. Maumbe & Owei, V. T. (Eds.) *E-agriculture and E-government for Global Policy Development: Implications for Future Directions* (pp 162 - 175). Hershey (USA): IGI Global Publishers.
- Ogwo, B.A. (1996). Systems approach in education. In B.A. Ogwo (Ed). *Curriculum development and educational technology* (pp 280 – 314). Makurdi: Onaivi printing and publishing Co. Ltd.

Polkinghorne, D. E. (2005). Language and meaning: Data collection in qualitative research. *Journal of Counseling Psychology*. 52 (2), 137 – 145.

QRS International (2007). NVivo 8: Overview. Retrieved on November 2<sup>nd</sup> from [http://www.qsrinternational.com/products\\_nvivo.aspx](http://www.qsrinternational.com/products_nvivo.aspx)).

Scott, J. L. and Sarkess-Wircenski, M. (2008). *Overview of Career and Technical Education*. USA: American Technical Publishers.

Ritchie, J. & Spencer, L. (2000). Qualitative data analysis for applied policy research. In A. Bryman & R. G. Burgess (Eds) *Analyzing qualitative data* (pp 173 – 208). London: Routledge Publishers.

Shenton, A. K. (2004). The analysis of qualitative data in LIS research projects: A possible approach. *Education for Information* 22 (2004), 143–162.

Shuttleworth, M. (2008). Choosing your research methods. Retrieved on November 02, 2009 from [http://www.pginfo.uhi.ac.uk/pick\\_research\\_methods.htm#questionnaires](http://www.pginfo.uhi.ac.uk/pick_research_methods.htm#questionnaires)

UNESCO and ILO (2002). *Technical and vocational education and training for the twenty-first century*. Paris: UNESCO.

# Combining quantitative and qualitative research methods in comparative study on VET cost-sharing mechanisms and regulatory instruments

*Simonas Gaušas and Rimantas Dumčius*

**Abstract:** The paper is based on the study commissioned by CEDEFOP “Sharing the costs of vocational education and training in the newer EU Member States” which was carried out by Public Policy and Management Institute in 2008-2009. The study had a multiple aim – to map and compare vocational education and training cost-sharing mechanisms and regulatory instruments across the 12 newer EU Member States, evaluate their influence on private investment and participation in VET, determine which ones are more successful and identify shortcomings in their governance, and also to analyse the influence of contextual factors on cost-sharing mechanisms and regulatory instruments and private investment and participation in VET. Multiple-aim has challenged to combine both quantitative (surveys, regression analysis) and qualitative (literature review, interviews, qualitative comparative analysis, qualitative analysis of statistical data) research methods. This paper shortly describes our journey from the method selection to their final application.

**Keywords:** comparative analysis, cost-sharing, triangulation

## Introduction

This paper is based on the study commissioned by CEDEFOP “Sharing the costs of vocational education and training in the newer EU Member States” carried out by Public Policy and Management Institute in 2008-2009. A number of initiatives at the EU level advocated developing balanced and shared funding and investment mechanisms. EU Member States have been experimenting with and introducing various cost-sharing mechanisms<sup>37</sup> and regulatory instruments<sup>38</sup> to ensure fair distribution – between public authorities, employers and individuals – of vocational education and training (further – VET) costs and responsibilities. Private cost-sharing, however, was not considered as hot topic in policy agendas of these countries. Furthermore, little was known about the practice and evaluation of cost-sharing approaches across the 12 most recent members of the EU<sup>39</sup> (further – EU12). The study aimed to fill this gap. It presented, for the first time, exploratory comparative analysis of cost-

---

<sup>37</sup> Cost-sharing mechanism is method and source through which funding is made available to VET participants. Cost-sharing mechanisms include collective financing arrangements (where employers and/or employees and/or social partners share the costs) and public-private cost-sharing. VET cost-sharing mechanisms analysed in the study are tax incentives, training funds, subsidy-based mechanisms, loans and savings schemes.

<sup>38</sup> Cost-sharing regulatory instrument sets out the general rules of cost-sharing. The regulatory instruments either help to reduce the risk of investment and secure the benefits of investment in training for employers (payback clauses) or to secure equitable access to training for employees (training leave). Both are analysed in the study.

<sup>39</sup> Bulgaria, the Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Romania, Slovenia and Slovakia.

sharing mechanisms and regulatory instruments to finance VET in the EU12 and has built strong arguments for the further improvement of private cost-sharing.

The aim of the study was multiple – to map and compare VET cost-sharing mechanisms and regulatory instruments across the EU12, evaluate their influence on private investment and participation in VET, determine which ones are more successful and identify shortcomings in their governance, and also to analyse the influence of contextual factors<sup>40</sup>. At first, the aim of the study looked unrealistic: no internationally comparable statistics and data on mechanisms and instruments, few academic papers in this field, very fragmented sources of information and limited budget and time period of the project. It would be impossible to carry the assignment by the standards of academic research. Therefore we opted for policy analysis – an applied research discipline which aims to solve practical problems and helps to inform decision making, in this case in the area of VET cost-sharing at the EU level (Martinaitis 2008). Although being more practically oriented, it uses tools and results of academic research (Stephen and Peters 1984).

Multiple aim of the study has challenged us in terms of what methods we should select for the analysis. We needed to use a combination of both qualitative and quantitative research methods. This paper shortly describes our journey from the choice of methods to their final application. Paper is structured in three parts. First details the methodological challenge we faced in the study. Second part explains our methodological approach. The last part summarises the results and provides insights on how to develop research approach further.

## **1. Targeted Problem**

Initially the study looked like mission impossible. This was due to several reasons. First was related to the aim of the study which was very complex and consisted of multiple questions. In our study we had to answer several interconnected research questions related to number and characteristics of cost-sharing mechanisms and regulatory instruments, their influence on private investment in VET as well as interrelationship between other variables of the study (see 2.1-2.2).

There was also a number of factors that considerably restricted scope and depth of our assignment. Firstly, internationally (i.e. across the EU12) comparable statistical data was often not available. For example, many EU countries cannot break their VET expenditure figures by programme orientation (general versus vocational education), type of VET (initial, continuing or VET for unemployed) or to indicate share of different stakeholders in VET funding. To make analysis possible we had to substitute the missing statistical indicators with the closest available proxy indicators. Furthermore, some statistical indicators had only been measured

---

<sup>40</sup> Such as the capacity of spending on education from public sources, the certainty of employers about their returns on investment in VET, the technological advancement of the economy, and the balance of supply with demand for labour with VET qualifications.

once or twice over the period analysed, which greatly limited the application of the regression analysis in this study. Secondly, academic literature analysing these mechanisms and instruments was very scarce. Thirdly, the information sources in this specific field were very fragmented in the EU12. In some countries it was even hard to identify representatives of policy-makers, social partners (trade unions and employers' representatives) and academics focusing on VET cost-sharing. Finally, the above-mentioned questions had to be answered within limited time and financial resources.

The above-mentioned limits challenged us when we had to choose research methods. Multiple aim of the study required using complex methodological approach and at the same time the choice of methods was restricted due to above-mentioned limits. We opted for using both qualitative and quantitative research methods and tried to apply triangulation techniques. This is further explained in the next section.

## 2. Methodology

### 2.1. Selecting variables

Three groups of variables were singled out in the study: dependent, independent and contextual variables. Dependent variables were two – private investment in VET and participation in VET. There was no specific single statistical indicator which could be used to measure each dependent variable. Therefore, to make the analysis possible, we substituted the missing statistical indicators with the groups of closest available proxy indicators (as shown in Table 1). The same was done in case of some contextual variables.

**Table 1: Measurement of dependent variables**

Dependent variable	Measurement indicators
Private investment in VET	Expenditure on educational institutions from private sources as percentage of GDP, for all levels of education combined;  Cost of CVT courses as percentage of total labour cost (all enterprises); and  Cost of CVT courses per participant (PPS - purchasing power standard)
Participation in VET	1. % of employees (all enterprises) participating in CVT courses; 2. Number of hours in CVT courses per employee (all enterprises); 2005; 3. Training enterprises as percentage of all enterprises; and 4. Percentage of labour force from ISCO 4-9 main groups, which received education and training during previous four weeks (formal and non-formal)

Source: prepared by the authors on the basis of statistical data from EUROSTAT

The dependent variables are interrelated: higher levels of private investment in VET should lead to greater participation in VET. Therefore, participation in training was a dependent variable for investment in training.

The independent variables were characteristics of VET cost-sharing mechanisms and regulatory instruments. The study has analysed 5 cost-sharing mechanisms and 2 regulatory instruments as shown in Table 2.

**Table 2: Independent variables (mechanisms and instruments) analysed in the study**

<b>VET cost-sharing mechanisms</b>	<ol style="list-style-type: none"> <li>1. Tax incentives for individuals and legal entities (types: tax allowances and tax credits);</li> <li>2. Training funds (types: national/multisectoral and sectoral);</li> <li>3. Subsidy-based mechanisms (types: grant schemes and vouchers/individual learning accounts);</li> <li>4. Loans (types: conventional and income-contingent);</li> <li>5. Saving schemes.</li> </ol>
<b>VET cost-sharing regulatory instruments</b>	<ol style="list-style-type: none"> <li>1. Payback clauses for individuals and for future (next) employers;</li> <li>2. Training leave (types: paid and unpaid).</li> </ol>

Source: prepared by the authors.

The analysis of causality between the characteristics of VET cost-sharing mechanisms and the levels of private investment in VET, as well as VET participation rates, involved several contextual variables which could influence both independent and dependent variables of this causal chain. On the basis of literature review and the regression analysis, four contextual variables correlating with the dependent and independent variables of the study were identified (Table 3).

**Table 3: Measurement of contextual variables**

<b>Contextual variable</b>	<b>Measurement indicators</b>
Capacity of spending on education from public sources	<ol style="list-style-type: none"> <li>1. Total public expenditure on education as percentage of GDP, for all levels of education combined</li> </ol>
Certainty of investment in training	<ol style="list-style-type: none"> <li>1. Percentage of all enterprises who establish the training needs of their personnel as percentage of training enterprises;</li> <li>2. Percentage of all enterprises with a training budget for CVT as percentage of all Enterprises; and</li> <li>3. Percentage of training enterprises using external advisory services.</li> </ol>

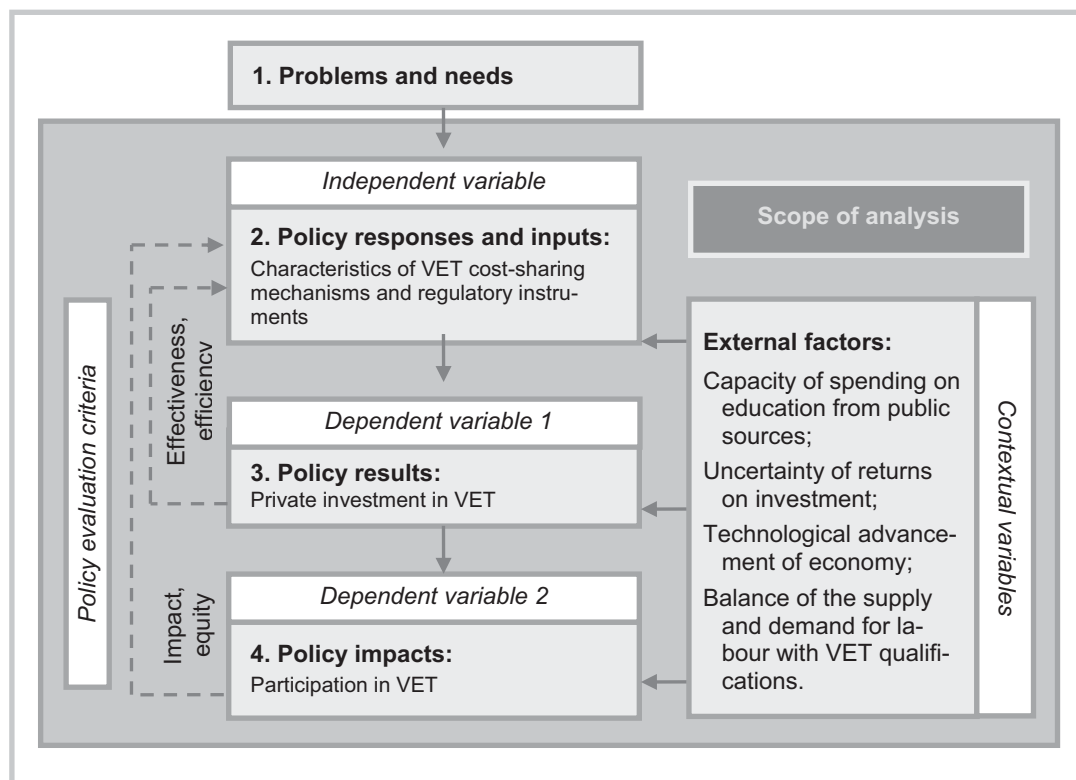
<b>Contextual variable</b>	<b>Measurement indicators</b>
Technological progress of the economy	<ol style="list-style-type: none"> <li>1. Employment in hi-tech sectors: employment in high/medium technology manufacturing and knowledge intensive services as percentage of total employment;</li> <li>2. Share of research and development (R&amp;D) staff and researchers in total employment;</li> <li>3. Abstract innovation index (data from European innovation scoreboard, EIS); and</li> <li>4. Labour productivity: GDP in purchasing power standards (PPS) per person employed relative to the EU-27</li> </ol>
Balance between supply and demand for labour	<ol style="list-style-type: none"> <li>1. Unemployment rate (yearly averages);</li> <li>2. Population activity rate;</li> <li>3. Expenditure on labour-market training as percentage of GDP; and</li> <li>4. Expenditure on labour-market services as percentage of GDP</li> </ol>

Source: prepared by the authors on the basis of statistical data from EUROSTAT

## *2.2. Building framework for analysis*

Most VET cost-sharing mechanisms and regulatory instruments are conceived as government initiated, regulated, sponsored or otherwise promoted interventions into the existing situation and intended to improve it. Our analytical framework was built on the basis of a simplified policy cycle, linking the perception of problems and needs, formulation of policy responses and their outcomes (results and impacts). The outcomes in turn often affect the perception of problems and result in policy changes. This cycle provides a clear structure of variables of the study (Figure 1).

**Figure 1: Framework for analysis**



Source: prepared by the authors.

The main research questions of the study were the following:

- What VET cost-sharing mechanisms and regulatory instruments are applied in the newer EU Member States? How they operate?
- Do more effective and efficient mechanisms and instruments increase private investment in VET?
- Does higher private investment in VET increase participation?
- To what extent and how did external factors influence mechanisms and instruments, private investment and participation in VET?

One of the aims of the study was to establish which mechanisms and instruments are more successful (*effective, efficient*) in stimulating private investment, and subsequently, which ones have greater *impact* on participation, while taking due regard to *equity* concerns. Our analytical framework shows these evaluation criteria as relationships between the variables of the study.



### 2.3. Choosing and applying research methods

Challenges of the study (see section 1) have limited the choice of research methods. Literature review has provided us little information about VET cost-sharing mechanisms and regulatory instruments. Therefore we had to carry out express and detailed surveys to gather all necessary data on mechanisms and instruments. There were many cases when data provided by survey respondents had to be further clarified through telephone or e-mail interviews. The survey database provided the basis for answering the first two questions of the study (section 2.2). However the detailed survey could not be made representative for all national VET experts in the field. It was only a pilot survey and its findings could not be used for quantitative analysis. Thus we opted for qualitative comparative analysis (further – QCA) to explore the influence of VET cost-sharing mechanisms and regulatory instruments and contextual variables on private investment in VET. The relationship between private investment and participation in VET as well as the influence of contextual variables on the remaining variables of the study was analysed using the regression analysis and qualitative analysis of statistical data available from EUROSTAT. Table 4 summarises research methods and their application.

**Table 4: Application of research methods in the study**

<b>Method</b>	<b>Application in the study</b>
<b>Literature review</b>	Used to initially map the existing VET cost-sharing mechanisms and regulatory instruments, to identify and operationalise contextual variables, to investigate academic arguments regarding the relationships between variables of the study.
<b>Express survey of national ReferNet coordinators</b>	Used to map the existing VET cost-sharing mechanisms and regulatory instruments (verify the results of literature review) and to identify national VET experts in each EU12 country.
<b>Detailed survey of national VET experts</b>	<p>Used to gather detailed factual information on the characteristics of mechanisms and instruments in each country and to collect expert evaluations on their effectiveness, efficiency, equity and impact. The intention was to receive at least three opinions from different experts who would include: the representatives of public authorities, social partners (trade unions and/or employers') and independent experts (e.g. academia).</p> <p>The expert evaluations of the same type of mechanisms/instruments (such as different types of tax incentives or training leave) were included in the analysis after an average of their expert evaluations (points given by the national experts on each indicator) was calculated. This was done to prevent the overall evaluation of the VET cost-sharing system being unbalanced due to dominance of one type of cost-sharing mechanism.</p>

Method	Application in the study
<b>Interviews</b>	Used to clarify and to check the responses to the questionnaire of the detailed survey. Interviews were carried out via telephone and/or e-mail.
<b>Qualitative Comparative Analysis (QCA)</b>	<p>Used to identify factors influencing private investment in VET.<sup>41</sup></p> <p>The dependent variables were dichotomised, namely all countries were divided in two groups according to the median value as high or low performing.</p> <p>The independent variables for the QCA were chosen from the VET expert survey. Following the initial analysis of survey responses, only the clearest and the least controversial responses were included in the analysis framework.</p> <p>Contextual variables were included in the QCA only if statistically significant preliminary correlations were observed between them and private financing variables.</p> <p><i>Tosmana</i> software<sup>42</sup> was used to conduct the QCA. Illustration of the results obtained using <i>Tosmana</i> software is provided in the Annex to this paper.</p>
<b>Regression analysis</b>	<p>Used to select the contextual variables demonstrating the highest correlations with the remaining variables of the study and to test relationships between all variables of the study. Sufficient data from official statistics was available for the regression analysis of relationships between private investment and participation in VET, and between contextual variables and private investment in VET. But even in these cases the regression analysis encountered difficulties due to lack of some data. An attempt was made to reduce these problems by extending the sample to the EU27 instead of focusing only on the EU12. In the regression analysis <i>Pearson's correlation</i> was considered 'significant' at 0.01 level (two tailed) and correlation coefficient was 'low' if under 0.5, 'medium' between 0.5 and 0.7 and 'high' if over 0.7.</p> <p>Regression analysis used <i>SPSS</i> software. The analysis was supplemented with qualitative analysis of statistical data.</p>
<b>Qualitative analysis of statistical data</b>	Used to initially select contextual variables, to compare the trends in private investment and participation in VET and external factors between the EU12, EU15 and EU27.

Source: prepared by the authors.

<sup>41</sup> Note: due to data limitations, statistical analysis could not be carried out, so a qualitative comparative analysis was applied.

<sup>42</sup> The software does not test the strength of the relationship, but it reveals covariations between relevant characteristics of cost-sharing mechanisms/regulatory instruments and private financing indicators in terms of formal logic. The result of the qualitative analysis is a list of sufficient and necessary conditions. This list has to be further analysed to establish whether the relationships are meaningful in addition to being logical. However, if sufficient and necessary conditions coincide, it can be presumed that the relationship is not accidental. Software is available from Internet: <http://www.tosmana.net>.

#### *2.4. Triangulating research methods*

Triangulation of methods involves using more than one method to gather data, such as interviews, questionnaires and documents (Denzin 2006). Triangulation is often used to indicate that more than two methods are combined in a study of the same phenomenon (Bogdan and Biklen 2006). This process is also known as ‘cross examination’ or ‘cross verification’.

Triangulation of methods in this study was used to verify the results of the influence of contextual variables on private investment in VET. Firstly, the relationships between variables of the study were indicated using literature review and analysed with the help of QCA. Secondly, regression analysis was carried out to cross verify the results obtained using literature review and QCA. Finally, we concluded that regression analysis and qualitative analysis of statistical data support the findings of the QCA about the importance of contextual factors in explaining the patterns of private investment in VET.

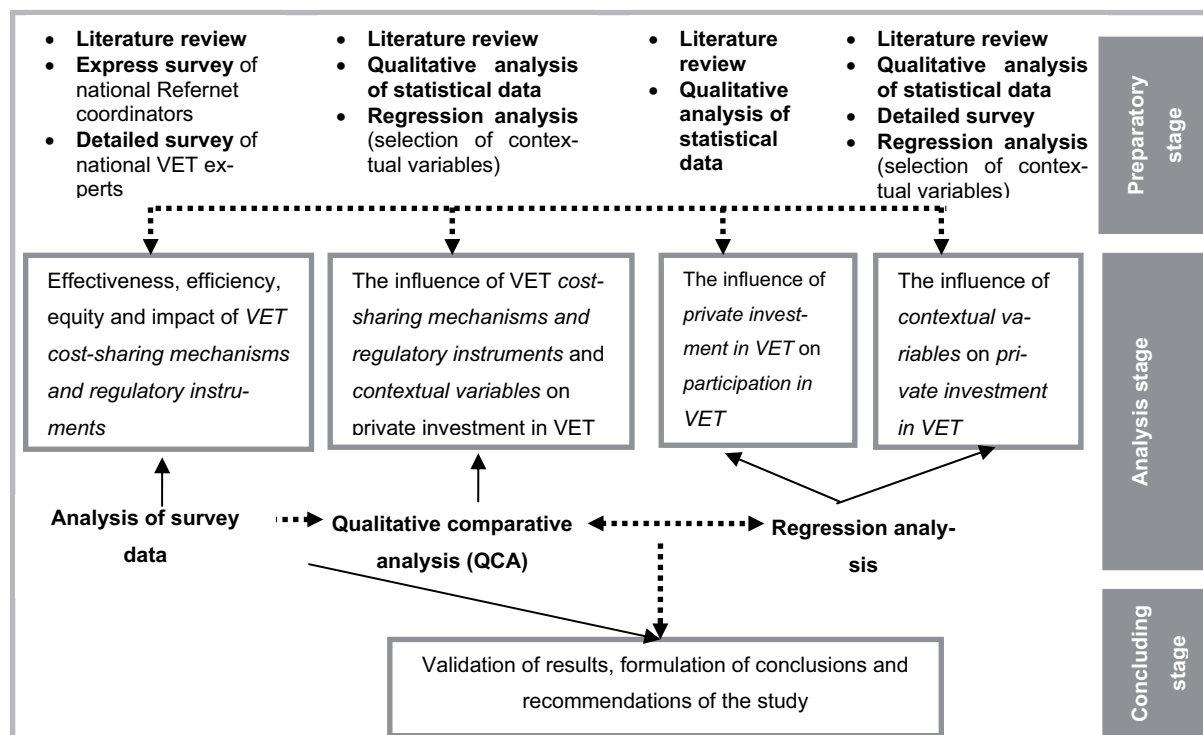
Furthermore, some results obtained with the help of survey and literature review were used to support findings of comparative and regression analysis. To mention just one example, regression analysis provided evidence that the overall public spending capacity and the level of public expenditure on education influence both private expenditure and participation in VET. This was supported by the example of Hungary taken from the detailed survey of national VET experts – country has recently abandoned several cost-sharing mechanisms due to its high budget deficit. Meanwhile when explaining the findings of QCA, it was assumed that if monitoring systems are sufficiently elaborate<sup>43</sup>, experts’ opinions are more likely to be based not only on general estimations or impressions but also on hard evidence.

Patterns of triangulation of methods are showed in dotted lines in Figure 2 below.

---

<sup>43</sup> This was estimated by the availability of monitoring and/or evaluation reports in the country.

**Figure 2: Triangulating research methods**



Source: prepared by the authors.

### 3. Results and perspectives for further development

#### 3.1. For the described instrument/approach as such

There are a number of methodological and practical/policy implications stemming from application of the research approach described above. The most important achievement, however, is that the study was the first attempt to make a cross-country list of VET cost-sharing mechanisms and regulatory instruments in EU12 and to conduct their comparative analysis. Additionally, the study was the first attempt to test the relationships between mechanisms and instruments existing in the EU12, context factors and private investment and participation in VET. The findings of the study could inform decision making while the data to carry out a rigid comparative analysis will be missing for the years to come. The study helped draw attention to poor governance standards and equity concerns, and urged to collect monitoring information, continue research and improve statistics in the field.

There are two main methodological limitations that limited scope and depth of the study. Avoiding or improving the following limitations would allow for a better quality of forthcoming comparative studies in this or similar field:

1. Comparative analysis can never ensure that all the relevant factors and conditions have been considered and studied. For example, it was impossible to access complete data on all key financial and other potentially important characteristics of mechanisms and instruments in the EU12 (i.e. no statistics, no literature and sometimes even no information from experts);

2. The number (and diversity) of mechanisms and instruments studied was not matched by sufficient number of observations of each mechanism. VET expert opinions on a particular mechanism were not representative of all VET experts of a particular country. Thus the limited number of observations can lead to identification of accidental, rather causal relationships;

### *3.2. Potentials and challenges for the use in/for international comparisons*

This research approach, although being complex, could potentially be very useful for exploratory comparative analysis investigating similar, under-researched areas which are important for policy decision making. However tools used to apply certain research methods could be different depending on the availability of data and nature of variables. For example, *Tosmana* is not the only tool for conducting comparative analysis. *Multi-criteria scoring method (MCSM)*<sup>44</sup> could be also applied in this particular case. Furthermore, even the approach could be different. For example, instead of carrying out comparative analysis it could be possible to use principles of counterfactual analysis<sup>45</sup>. However both MCSM and methods used for counterfactual analysis require large and comprehensive sets of statistical data/observations. Consequently the greatest challenges in using this or similar research approach would be probably related to the lack of internationally comparable statistics (i.e. lack of indicators, short periods of available data) and lack of research (i.e. short number of observations, little or no research regarding the relevant factors and conditions).

---

<sup>44</sup> Multi-criteria scoring method (MCSM), sometimes called multi-criteria decision making (MCDM), is a discipline to support complex system analysis and aid decision makers who are faced with making numerous evaluations in connection with these systems. Measurements in MCSM are executed on different scales and therefore provide opportunity to include subjective indicators, qualitative opinions and various preferences. Preferences in MCSM are essential; outcome of complex system analysis depends on who is making the decision and what their goals and preferences are. Because of these features this methodology is perfectly suitable for large scale policy assessment and strategy analysis. This type of analysis can be supported by several software tools, such as Criterium DecisionPlus by InfoHarvest. For more information see, for example Saaty 2001.

<sup>45</sup> Counterfactual analysis helps researcher to identify what would have happened without intervention (i.e. to establish the counterfactual). In this type of analysis it is essential to randomly select a control group or to identify a suitable comparison group with the help of matching methods. Quasi-experimental methods for the design of comparison groups such as Propensity Score Matching (PSM) or Difference-in-Difference method could be applied. For more information see Tavistock Institute 2003.

## References

Bogdan, R. C. and Biklen, S. K. 2006, Qualitative research in education: An introduction to theory and methods, Allyn & Bacon.

CEDEFOP. 2009, Sharing the costs of vocational education and training. An analysis of schemes in the newer EU Member States, CEDEFOP Panorama series, 182. Luxembourg: Publications Office of the European Union.

Denzin, N. 2006, Sociological Methods: A Sourcebook, Aldine Transaction, 5<sup>th</sup> Edition.

Martinaitis, Ž. 2008, Problemos apibrėžimas politikos analizėje (Problem definition in policy analysis). Vilnius: Vilniaus universiteto leidykla.

Saaty, T. L. 2001, Fundamentals of Decision Making and Priority Theory. Pittsburgh, Pennsylvania: RWS Publications.

Stephen H. L. and B. G. Peters. 1984, From Social Theory to Policy Design, Journal of Public Policy, Vol.4, No.3.

Tashakkori A. and C. Teddlie. 2003, Handbook of Mixed Methods in Social & Behavioral Research, SAGE.

Tavistock Institute. 2003, The GUIDE – The Evaluation of Socio-Economic Development. Available at: <http://www.evaled.info>.

## Annex

Table below provides an illustration of the results obtained using Tosmana software.

**Table: Characteristics of VET cost-sharing mechanisms and regulatory instruments, contextual variables and private expenditure on educational institutions\***

Achieving objectives	Efficiency (combined)	Deadweight effect	Employment in high-tech sectors (a)	Labour productivity (b)	Public expenditure on education (c)	Planning investment in human resources (d)	Unemployment rate (e)	Private expenditure on educational institutions (f)	Cases
low	high	low	high	high	low	high	high	high	CZ
high	high	high	high	low	low	high	low	low	EE
high	high	high	low	high	high	low	low	high	CY
low	low	high	low	low	high	low	low	high	LV
high	low	low	low	low	high	low	low	low	LT
high	high	low	high	high	high	low	high	low	HU
high	high	high	high	high	low	low	high	low	MT
low	high	high	low	low	high	high	high	high	PL
low	low	low	low	low	low	high	high	low	RO
low	high	low	high	high	high	high	low	high	SI
low	low	low	high	high	low	high	high	high	SK

(a) Employment in high technology sectors is 'high' if it is above the median value, which is 34.25.

(b) Labour productivity is 'high' if it is above the median value, which is 70.4.

(c) Public expenditure on education is 'high' if it is above the median value, which is 5.

(d) Planning investment into human resources is 'high' if it is above the median value, which is 28.27.

(e) Unemployment is 'high' if it is above the median value, which is 7.1.

(f) Private expenditure on educational institutions is 'high' if it is above the median value, which is 0.6.

NB: Indicators appearing in formulas of the qualitative comparative analysis and thus considered as influencing the dependent variables are marked in blue if the relationship is direct and in orange if the relationship is inverse (i.e., which characteristics are present in countries where private expenditure on educational institutions is low).

Source: prepared by the authors.

# Qualitative Evaluation of Governance Structures in VET

*Dr. Ludger Deitmer, Dr. Lars Heinemann, Prof. Dr. Felix Rauner and Dr. Wolfgang Wittig*

**Abstract:** In this contribution we present a theoretical framework for the classification of governance structures in VET and a workshop tool for the qualitative evaluation of VET systems according to this framework. VET systems can be classified along two dimensions, namely the integration of the system, which can be high or low, and the mode of governance, which can be predominantly input or output oriented. These two dimensions constitute a coordinate system and allow for a taxonomy of four types of VET governance (fragmented output-oriented governance, coordinated output-oriented governance, fragmented input-oriented governance, coordinated input-oriented governance). The two dimensions are represented by a set of main criteria, which are broken down into sub-criteria that are rated by experts in a questionnaire. The ratings allow for the positioning of VET systems within the coordinate system.

**Keywords:** governance, public value, evaluation

## Introduction

The present contribution presents a workshop tool for the qualitative evaluation of governance structures in vocational education and training that allows for the mapping of VET systems within a two-dimensional taxonomy and for the identification of possible shortcomings. This instrument has been designed to support the comparative analysis of VET systems by means of expert workshops and serves two major purposes: (1) to collect qualitative data on the current state of affairs concerning the management and governance of selected VET systems, and (2) to provide a basis for the in-depth discussion and evaluation of specific factors and dimensions related to the governance structures. In the following sections we first discuss the theoretical background of VET governance and the related research problems. We then present the rationale behind the development of the criteria for the evaluation of governance structures and explain the design and application of the instrument. In the final section we present selected results from a comparative study in which the evaluation tool was applied and discuss the challenges and potentials for the further development and application of the instrument.

## Targeted Problem

In vocational education and training three ideal types of regulation and governance are usually distinguished on the basis of the roles of the agents and the underlying rationale of agency. The dominant influence may come either from the state, the market or professional groups. On the basis of the categories of social regulation that have been commonplace in sociology since the time of Max Weber – tradition, market and bureaucratic rationality – the prevalent typology in the social sciences distinguishes three models of governance, which



can be termed market-driven, state-controlled and occupation-driven or corporatist VET governance (see Greinert 1998, 19–22; Clematide et al. 2005, 3–4).

The market-driven model of VET governance is characterised by the immediate control of vocational qualification by the employment system and the demand on the labour market. Vocational qualification is oriented towards the requirements of employers and takes place on the job and in a private sector of training providers offering job-related learning modules. The responsibility for the training process rests with the learners, who are expected to acquire the qualifications required by employers on their own. Typical examples of this model are the United States and Japan, where the relative absence of a regulated VET system is associated with a large number of students attending upper secondary schools and higher education. In this system VET as well as the access to VET are controlled by employers as “customers”, whose needs and demands determine the contents of training so that the transfer of qualifications from one company to another is difficult (cf. Greinert 1998, 20–21). On the one hand this system is regarded as quite flexible and adapted to the needs of the employment system, on the other hand the dependence on the private supply of training opportunities and the risk of underinvestment in vocational education are seen as serious flaws of this model (cf. Clematide et al. 2005, 3).

The state-controlled model of VET is characterised by a dominance of school-based vocational education, which is subject to a relatively tight regulation by state authorities. In this model, which is prevalent, for instance, in France or China, the regulation is based on the school’s logic of action and includes a focus on civic education. Enterprises do not have an institutionalised role in this system, but serve as suppliers of internships while all regulatory functions – planning, management and control – are concentrated in the public sector. The contents of vocational education are typically based on theoretical and academic types of education (cf. Greinert 1998, 21–22). Due to the integration into the state-controlled education system there is a relatively close connection to general education. Moreover, the supply of training opportunities is independent of the provision of training places by private companies. The major difficulty of this system is the weak linkage to the labour market (cf. Clematide et al. 2005, 3)

The third model is usually referred to as traditional occupation-based or corporatist regulation (cf. Greinert 1998, 19–20). Historically this model is derived from the apprenticeship tradition in the craft trades. It is characterised by a strong influence of the training companies and the chambers (i.e. the corporate bodies or associations that represent the business community at the local or regional level). This concerns the access to training as well as the definition of training contents and the responsibility for examinations. Today occupation-based regulation is part of “mixed” systems of cooperative governance in which the regulation of vocational education takes place in a plural network of state bodies, enterprises or employers’ associa-

tions as well as trade unions or professional associations. Variations of these mixed models of regulations can be found in systems of cooperative (dual or alternating) VET as they exist in Austria, Denmark, Germany and Switzerland

In practice the most important examples of these mixed VET systems are the models of alternance and dual apprenticeship training. The common feature of these models is the combination of in-company training and school instruction. In the case of alternating training the phases of school instruction and practical training alternate in relatively long periods, and the vocational school or college remains the dominant learning venue. Curricula are usually fixed by state authorities, and despite the involvement of enterprises the state has a dominating role in this model.

The situation is somewhat different in the model of dual apprenticeship training. This model is characterised by the fact that the vocational education and training system is composed of two independent but interrelated subsystems, namely, an in-company training sector organised by private enterprises and a corresponding sector of vocational school instruction for which the state is responsible (cf. Greinert 1998, 23–24). Although this model appears at first glance as a combination of market and state regulation, there are also considerable elements of occupation-based and corporatist governance. In Germany, for instance, the traditional strong role of the occupational principle (*Berufsprinzip*) entails a control of the access to vocational training by the occupational groups concerned. Following the tradition of the guilds, they participate in the formulation of training curricula and influence the organisation of vocational examinations through the chambers, which are the bodies that officially represent the companies.

These considerations lead to the problem how the systems of dual or alternating apprenticeship training can be characterised and situated with regard to their governance structures, and how this affects the performance of the systems. The epistemic interest is to identify examples of good practice in plural administration that can serve as a basis for policy recommendations.

## **Methodology**

Plural governance systems in which state-controlled and market-driven or corporatist types of governance overlap can be classified on the basis of two dimensions of the governance process. The first dimension is the degree of coordination between the different agents with their respective internal logic or, to put it differently, the integration of the system. At one end of the scale the “plural administration” may be completely fragmented. In this case the public and private or corporative agents act autonomously within the legal framework and follow their own internal logic of agency without coordinating their activities. Each class of agents fulfils the tasks assigned to them in the context of the VET system independently. The re-

sponsibilities are not allocated according to functions, but according to subjects and domains, which means that the administrative functions of rule-making, execution and monitoring are dispersed across all types of bodies in varying constellations.

The second dimension is the mode of governance, that is, the rationale that underpins the behaviour of the different agents and thus shapes the governance process. It is common in public management to distinguish between an input oriented type of management by rules and resources, and an output oriented management by means of the products and services to be achieved by the management process. Input control is typical of the traditional bureaucratic model of public administration, which is primarily concerned with the implementation of the law. Output control, on the other hand, is one of the cornerstones of the New Public Management approach, which claims to improve the efficiency of the public sector by means of management techniques adapted from the private business sector (cf. Osborne and Gaebler 1993; Spicer 2004).

These two dimensions allow for the construction of a coordinate system whose four quadrants represent the different types of plural corporatist governance system in vocational education and training. In the case of a fragmented input control the management processes follow the paradigm of the implementation of norms as expressed in the principle of the rule of law. The responsibilities are dispersed among different institutions or sub-systems of the political system. This includes rule-making as well, which is carried out by different bodies for their respective domains. The distinctive feature of fragmentation is that competences are allocated according to policy areas and that a vertical integration takes place at best within these areas. The result is that the institutions operate relatively independently of each other and have few incentives to coordinate their actions. A coordinated input control, on the contrary, is also characterised by a primacy of rules, but institutional arrangements such as the concentration of legislative powers and a consistent responsibility of government departments allow for a coordination of the bodies involved. Coordinated input control therefore features a more systematic structure of the legal framework and a consistent and coordinated implementation of the rules. The third model is fragmented output control, which combines a highly decentralised set of administrative bodies with a management by objectives. As this type of management automatically entails a relatively high autonomy on the part of institutions, the integration of the system as a whole can be secured only by means of a coordinated or centralised definition of the objectives in question. In the absence of such a centralisation or coordination there is the risk of the ultimate disintegration of the VET system and its replacement with a market of qualifications. Accordingly the fourth model, which can be termed coordinated output control, aims to secure the integration of the system by coherent objectives, which are formulated by a central body or developed jointly by the bodies involved. The following table summarises this conceptual framework:

**Table 1: Types of governance in vocational education and training**

	Integration of the system	
	Low	High
Rationale of agency	Fragmented output control	Coordinated output control
Input Output	Fragmented input control	Coordinated input control

The classification of existing VET systems according to the taxonomy described above allows for the development of policy recommendations if a type of governance can be identified that can reasonably be considered the optimum for dual or alternating vocational education and training. Governance within the public sector faces the problem that a simple adaptation of evaluative criteria that were originally developed for the private business sector is not possible. The reason is the difference between the internal logic of the economic system on the one hand and the state or public sector on the other. It can be expected that the “public value” model of plural and deliberative governance (cf. Spicer 2004) is a promising alternative to state-controlled or market-driven types of public policy. The integration of different stakeholders into complex decision-making processes is discussed in public policy research under the concept of *governance network*, which can be regarded as a particular way of implementing plural administration. The current debates on the democratic anchorage of governance networks (for a Abstract, see Sørensen 2005) show that the measures that may be taken for this purpose correspond to relevant elements of the public value approach and are capable of promoting the implementation of the latter. The ideas of public deliberation, multiple accountability, publicity and the continuous renewal of legitimacy are taken up in these democratization strategies. To put it differently, public management within a network – and the VET system may well be considered a network due to its plural structure – must be based on objectives defined by public deliberation as well as on norms and thus combine input and output control if it is to meet the requirements of democratic instead of merely efficient governance. A governance model for the VET system that is optimal in this sense should therefore include a high degree of coordination between the bodies involved and should combine elements of input orientation like participation and deliberation with elements of output orientation such as performance orientation, efficiency and quality assurance.

This model is operationalised by an evaluation tool with several indicators that are listed below. In addition to desk research carried out on the basis of the theoretical framework the set of criteria opens the opportunity to carry out expert interviews with a view to situating the different VET systems within the coordinate system described above. There are seven main criteria, of which five relate to the integration of the system (i.e. coordination and fragmenta-

tion) and two to the dimension of input and output orientation. These main criteria are the following:

Dimension 1: Integration of the system

Category 1: Consistent legal framework

Category 2: Cooperation of the various bodies

Category 3: Innovation strategies

Category 4: Balance of relevant policy areas

Category 5: Allocation of strategic and operational functions

Dimension 2: Input and output orientation

Category 6: Outcome orientation

Category 7: Input orientation

These criteria are operationalised by approximately 30 sub-criteria or items that are evaluated and discussed by experts in the course of evaluation workshops. Respondents are asked to judge the items on a scale from 1 (= not realised) to 10 (= fully realised). The aggregated answers determine the position of the VET system within the matrix described above. The position on the horizontal axis “integration of the system” is defined by the mean of the values for the main criteria 1 to 5 with increasing numerical values indicating a higher degree of coordination. As regards the second dimension, the value is calculated on the basis of the mean of the two remaining main criteria 6 and 7. Given that the two main criteria have a reciprocal relationship so that a system is situated halfway between the poles of input and output control if the two criteria are equally realised, the values are standardised before the mean is calculated. Therefore the value for the position on the vertical axis is calculated according to the following formula:

$$(n_{\text{Outcome}} + 10 - n_{\text{Input}})/2$$

The value expresses which of the two modes of governance has a stronger influence on the VET system in question.

The expert workshops in which the instrument is applied can be characterised as follows. The focus of the methodology is on national experts who are concerned with management and governance issues of their VET systems. The experts participate in a workshop where they express their views on the current state of affairs and rate the governance activities by means of the questionnaire described above. The role of the moderator is to support the evaluation process and to make sure that all participants have the opportunity to explain their views. Both the diverging and the converging ratings on the various criteria are clarified and

discussed. The final discussion, in which the participants look for a consensus, offers the opportunity for a deeper understanding of the issues for all participants.

## **Results and Perspectives for further development**

### *For the described instrument/approach as such*

The governance evaluation tool was applied in a comparative study of the VET systems in Austria, Denmark, Germany and Switzerland. The expert workshops were carried out in Vienna, Copenhagen, Berlin and Zurich in November 2007. Together with national case studies based on desk research and expert interviews the study provided interesting insights into the VET governance structures of these countries (Bertelsmann Stiftung 2009; Rauner, Wittig & Deitmer, forthcoming). The positioning of the four countries shows that in Germany the fragmentation of governance is particularly strong while in Denmark and in Switzerland there is a remarkable degree of coordination. With a score of 7.8 on the axis “integration” and 2.8 on the “input/output” axis Denmark exemplifies the type of coordinated output control. For Germany (4.4; 6.3) the analysis shows a weak coordination and a clear dominance of input orientation. Austria (6.0; 5.4) shows a stronger, but still relatively weak coordination and a balanced ratio of input and output control. In Switzerland (7.0; 5.1) the coordination is already very strong and almost reaches the Danish figure. In addition there is an almost equal distribution of input and output control, which suggests that the Swiss VET system comes closest to the ideal model of governance as discussed above.

In all four workshops the instrument was successfully used by the participants and stimulated interesting group discussions. However, it was also observed that the formulation of some of the sub-criteria posed difficulties for some participants and were not always easy to evaluate. The conclusion drawn from this observation was that the criteria have to be reformulated in a more neutral way as the current version still reflects the particular features of the German VET system to some extent.

### *Potentials and challenges for the use in/for international comparisons*

As already described in the previous sections, the methodology for the qualitative evaluation of governance structures by means of expert workshops is based on a theoretical framework that captures the major dimensions of the administrative and supporting structures that are in place in vocational education and training, and it offers the opportunity to locate VET systems within the taxonomy that is constituted by these dimensions. This mapping of governance systems allows for the identification of their particular strengths and weaknesses and gives orientation with regard to further research needs.

## References

Bertelsmann Stiftung (Hrsg.) 2009, Steuerung der beruflichen Bildung im internationalen Vergleich, Gütersloh: Verlag BertelsmannStiftung.

Clematide, B., Dahl, A., Vind, A. & Joergensen, C.H. 2005, Challenges for the Danish VET system – on the path towards a future model, Berufs- und Wirtschaftspädagogik online, 7.

[http://www.bwpat.de/7eu/clematide\\_etal\\_dk\\_bwpat7.pdf](http://www.bwpat.de/7eu/clematide_etal_dk_bwpat7.pdf)

Greinert, W.-D. 1998, Das "deutsche" System der Berufsausbildung. Tradition, Organisation, Funktion, 3<sup>rd</sup> ed., Baden-Baden: Nomos.

Osborne, D. & Gaebler, T. 1993, Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector, New York: Plume.

Rauner, F., Wittig, W. & Deitmer, L. forthcoming, Plural Administration in Dual Systems in Selected European Countries. In: F. Rauner & E. Smith (eds.), Rediscovering Apprenticeship, Dordrecht: Springer.

Sørensen, E. 2005, The Democratic Problems and Potentials of Network Governance, European Political Science 4, 348–357.

Spicer, M. 2004, Public Administration, the History of Ideas, and the Reinventing Government Movement, Public Administration Review 64, 353–362.



## **Contributors**

### **Kathrin Hoeckel**

*Organisation for Economic Co-operation and Development  
Directorate for Education, Education and Training Policy Division 2,  
rue André-Pascal, 75775 Paris CEDEX 16, France*

### **Ute Hippach-Schneider**

*Federal Institute for Vocational Education and Training (BIBB)  
Robert-Schuman-Platz 3, 53175 Bonn, Germany*

### **Dr. Torsten Dunkel**

*Cedefop, PO Box 22427 GR-551 02 Thessaloniki*

### **Stefan Thomas**

*German Agency for Technical Cooperation (gtz)  
Dag-Hammarskjöld-Weg 1-5, 65760 Eschborn, Germany*

### **Noela Eddington and Ian Eddington**

*Department of Education and Training, Queensland,  
P.O. Box, City East Brisbane, Qld 4001, Queensland, Australia  
  
University of Southern Queensland,  
P.O. Darling Heights, Toowoomba, Qld, 4350, Queensland, Australia*

### **Dr. Xiaorong Nie, Mr. Ziyuan Wang and Mrs. Zhibao Pan**

*China National Training Base for Vocational Education Teacher  
Jiang Xi Science and Technology Normal University  
Feng Lin Da Dao, 330013, Nan Chang, Jiang Xi, P. R. of China*

### **Dr. Uwe Lauterbach**

*Independent researcher at Deutsches Institut für Internationale Pädagogische Forschung  
(DIPF), Frankfurt a. M. & Institut Technik und Bildung (ITB) der Universität Bremen  
Darmstadt, Germany*

### **Prof. Dr. Marc Schütte and Prof. Dr. Georg Spöttl**

*Faculty Life Sciences, Hamburg University of Applied Sciences  
Lohbrügger Kirchstr. 65, 21033 Hamburg, Germany  
Institute Technology and Education, University Bremen,  
Am Fallturm 1,-28359 Bremen, Germany*



**Dr. Lorenz Lassnigg and Dr. Stefan Vogtenhuber**

*Institute for Advanced Studies (IHS), Vienna  
Stumpergasse 56, 1060 Vienna, Austria*

**Prof. Jim Stewart and Victoria Harte**

*Leeds Business School and Institute for Enterprise, Leeds Metropolitan University,  
Civic Quarter, LS1 3HE, Leeds, UK*

**Dr. Geoff Hayward**

*Department of Education, University of Oxford  
15 Norham Gardens, Oxford OX2 6PY, UK*

**Dana Stroie**

*National Centre for TVET Development,  
Spiru Haret 10 – 12, 010176 Bucharest, Romania*

**Dr. Theodorah. U. Ogali**

*Department of Agricultural Economics & Extension, University of Port Harcourt  
East/West Road, PMB 5323 Choba, Rivers State, Nigeria*

**Prof. Pilar Pineda, Carla Quesada and Victoria Moreno**

*Facultad de Ciencias de la Educación, Universidad Autónoma de Barcelona Despacho 148,  
Edificio G-6, 08193 Cerdanyola del Vallés (Barcelona), Spain*

**Sandra Speer**

*Independent Evaluator, Wiesbaden, Germany*

**Dr. Benjamin A. Ogwo**

*Department of Vocational Teacher Preparation, State University of New York,  
Oswego, NY 13126, USA*

**Simonas Gaušas & Rimantas Dumčius**

*Public Policy and Management Institute,  
Gedimino ave. 50, LT-01110 Vilnius, Lithuania*

**Dr. Ludger Deitmer, Dr. Lars Heinemann, Prof. Dr. Felix Rauner and Dr. Wolfgang Wittig**

*Institute Technology and Education (ITB), University of Bremen  
Am Fallturm 1, 28359 Bremen, Germany  
TVET Research Group (I:BB), University of Bremen  
Leobener Str./NW2, 28359 Bremen, Germany*

## **Abstract**

There is a high demand for methods and instruments that assist VET experts in taking stock of the state of VET systems for recommending appropriate measures to policy makers in order to face current challenges. The existing methods and instruments can be applied for a systems' overview, in order to recognise strengths and weaknesses as well as to identify possible areas for system development and improvement.

This paper gathers the contributions from a workshop has been conducted in December 2009 in Königswinter by BIBB, bringing together experts from international and national organisations that develop, apply or use such tools. During this workshop, an overview has been provided of existing approaches. Experience with their application as instruments for analysing and supporting policy and practice has been shared. Topics were discusses from three perspectives: System evaluation, Indicators & benchmarks, Peer review and peer learning.