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

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

In:

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

The foundations of evaluation and impact research

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

Reproduction is authorised provided the source is acknowledged

Additional information on Cedefop’s research reports can be found on: http://www.trainingvillage.gr/etv/Projects_Networks/ResearchLab/

For your information:

• the background report to the third report on vocational training research in Europe contains original contributions from researchers. They are regrouped in three volumes published separately in English only. A list of contents is on the next page.

• A synthesis report based on these contributions and with additional research findings is being published in English, French and German.


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

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

For further information contact:
Cedefop, PO Box 22427, GR-55102 Thessaloniki
Tel.: (30)2310 490 111
Fax: (30)2310 490 102
E-mail: info@cedefop.eu.int
Homepage: www.cedefop.eu.int
Interactive website: www.trainingvillage.gr
Impact of education and training

Preface

The impact of human capital on economic growth: a review
Rob A. Wilson, Geoff Briscoe

Empirical analysis of human capital development and economic growth in European regions
Hiro Izushi, Robert Huggins

Non-material benefits of education, training and skills at a macro level
Andy Green, John Preston, Lars-Erik Malmberg

Macroeconometric evaluation of active labour-market policy – a case study for Germany
Reinhard Hujer, Marco Caliendo, Christopher Zeiss

Active policies and measures: impact on integration and reintegration in the labour market and social life
Kenneth Walsh and David J. Parsons

The impact of human capital and human capital investments on company performance Evidence from literature and European survey results
Bo Hansson, Ulf Johanson, Karl-Heinz Leitner

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

The foundations of evaluation and impact research

Preface

Philosophies and types of evaluation research
Elliot Stern

Developing standards to evaluate vocational education and training programmes
Wolfgang Beywl; Sandra Speer

Methods and limitations of evaluation and impact research
Reinhard Hujer, Marco Caliendo, Dubravko Radic

From project to policy evaluation in vocational education and training – possible concepts and tools. Evidence from countries in transition.
Evelyn Viertel, Søren P. Nielsen, David L. Parkes, Søren Poulsen

Look, listen and learn: an international evaluation of adult learning
Beatriz Pont and Patrick Werquin

Measurement and evaluation of competence
Gerald A. Straka

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

Evaluation of systems and programmes

Preface

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

Evaluating systems’ reform in vocational education and training. Learning from Danish and Dutch cases
Loek Nieuwenhuis, Hanne Shapiro

Evaluation of EU and international programmes and initiatives promoting mobility – selected case studies
Wolfgang Hellwig, Uwe Lauterbach, Hermann-Günter Hesse, Sabine Fabriz

Consultancy for free? Evaluation practice in the European Union and central and eastern Europe
Findings from selected EU programmes
Bernd Baumgartl, Olga Strietska-Illina, Gerhard Schaumberger

Quasi-market reforms in employment and training services: first experiences and evaluation results
Ludo Struyven, Geert Steurs

Evaluation activities in the European Commission
Josep Molsosa
From project to policy evaluation in vocational education and training – possible concepts and tools
Evidence from countries in transition
Evelyn Viertel, Søren P. Nielsen, David L. Parkes, Søren Poulsen (1)

Abstract
This study follows the international trend in evaluation approaches in vocational education and training (VET), moving away from single projects towards policy evaluations. It examines the contribution of research and a number of examples from transition countries in central and south-east European countries with a view to helping clarify underlying concepts, methods and possible tools used when undertaking VET policy evaluations.

Using VET policy evaluations as a platform for identifying and designing reform proposals in transition countries, as well as monitoring reform progress, the study is as much concerned with aspects of system analysis, as it is with elements of implementation research.

The study argues that there is no holy grail in terms of conceptualisation or methodology related to VET policy evaluations. The engineer’s toolbox is of limited use. Similarly, management approaches based on refined analytical frameworks have been found to obscure rather than illuminate VET, including reliance on tools derived from system analysis. Thus, the only remedy seems to be the evaluator’s broad understanding of the essential components of VET, of the relationships between them, of the fundamental logic between the system and its environment and, of change levers. This is an understanding that develops only through many years of apprenticeship and first-hand experience with VET policy evaluations, and not least through a close dialogue with key local actors.

Evaluators need to develop an understanding of both the starting points for reform in the given country context and strategic levers of change. The study emphasises that process evaluation, which acknowledges the role of organisational or administrative learning as part of an overall evaluation activity, is as important as impact studies. The study also argues that any VET policy evaluation needs to pay specific tribute, amongst others, to the professionalisation and continuous development of teachers and other education specialists, as they are key to the success of any systemic VET reform effort.

Presented in this study are the first results of a kind of ground-clearing work in VET policy evaluations. As such, we see our study as a basis for discussion, which opens a number of fields for further research and debate.

(1) In cooperation with Jean-Marc Castejon, Henrik Faudel, Barbara Kuta and Piero Benazzo.
1. Introduction
   1.1. Methods of work

2. The limited transparency of concepts and methods used in VET policy evaluations
   2.1. OECD
   2.2. The building-blocks approach

3. Humanistic approaches to understanding VET systems
   3.1. The hermeneutic approach
   3.2. The structural-historical approach
   3.3. Understanding and explanation

4. VET as a (sub)system and issues of internal and external consistency
   4.1. System analysis
   4.2. VET as a subsystem
   4.3. Internal and external consistency
      4.3.1. Internal consistency
         4.3.1.1. Consistency between subsystems and clarification of pathways
         4.3.1.2. Regional consistency
         4.3.1.3. Consistency for the individual learner
      4.3.2. External consistency
      4.3.3. Regulating mechanisms
   4.4. Examples of lack of internal and external consistency of VET reforms in Bosnia-Herzegovina, Estonia and Romania
      4.4.1. Teacher training in Bosnia-Herzegovina (Durand-Drouhin and Bertrand, 1994)
         4.4.1.1. Pedagogical Academy in Sarajevo (PAS) in charge of education of primary school teachers
         4.4.1.2. Sarajevo University, Faculty of Philosophy, Pedagogical Department, in charge of pre-service VET teacher training
         4.4.1.3. Pedagogical Institute, Canton Sarajevo, in charge of in-service VET teacher training
      4.4.2. Curriculum reform in Estonia (Nielsen, 1999b)
      4.4.3. National dissemination of pilot project results in Romania (Nielsen, 1999a)
   4.5. Example of stressing system connections in a VET policy evaluation for Croatia
      4.5.1. Decentralisation and school headmasters’ and teachers’ training
         4.5.1.1. Governance
         4.5.1.2. Finance
         4.5.1.3. Curriculum (content)
      4.5.2. Curriculum reform and teacher training
         4.5.2.1. Design and development of new curricula
         4.5.2.2. New equipment
         4.5.2.3. Curriculum delivery
         4.5.2.4. Teams of teachers
         4.5.2.5. Curriculum assessment and revision
5. System change and the process and actors in VET policy evaluations 216
   5.1. The complexity of systemic reforms 216
   5.2. The path dependency of institutional change 217
   5.3. Accomplishing change 217
   5.4. The role of evaluations 218

6. Criteria, values and indicators underpinning VET policy evaluations 220
   6.1. Evaluation criteria 220
   6.2. Similarities and differences in VET reforms in western European countries 220
       6.2.1. Improving progression from initial VET to higher education 221
       6.2.2. Improving links between vocational education and employers 221
       6.2.3. Improving the status of VET teachers and trainers 221
       6.2.4. Improving the vocational curriculum 222
       6.2.5. Summary 222
   6.3. Potential and limits of the benchmarking approach 223

7. Conclusions and needs for further research 226
   7.1. Conclusions 226
   7.2. Needs for further research 229

List of abbreviations 231

Bibliography and references 232
List of figures

Figures

Figure 1: Basic concepts in evaluation 195
Figure 2: Denmark’s ranking – education sector 224
National and international donor agencies have a shared understanding of how and why to undertake evaluations. Evaluation in its classical definition aims at comparing the design and implementation of projects or programmes to actual outcomes by analysing:
(a) relevance to objectives and defined needs;
(b) efficiency in providing inputs promptly and at lowest cost;
(c) effectiveness in achieving planned outputs and immediate results;
(d) impact on higher level objectives to which the result should contribute;
(e) sustainability over time, and especially after the project/programme inputs have been provided and/or external assistance stops (Grootings, 2000) (²).

The following scheme summarises the relations between these different criteria and aspects of projects/programmes:

The scheme illustrates that:
(a) efficiency evaluation has to do with the relation of means and expected results. It assumes that means and expected results are clearly formulated;
(b) effectiveness evaluation deals with the relationship between achieved results and immediate objectives. Again, this assumes the existence of clearly formulated objectives;
(c) impact evaluation concerns the relationship between immediate objectives and overall objectives. This assumes that both immediate objectives and overall objectives are clear (Grootings, 2000) (³).

However, evaluation increasingly becomes a continuous cycle that takes place at all stages of a reform or change process. The aim is to reflect on policies and practice with a view to influencing them. Such evaluation is meant to have an impact on decision-making processes, on how reform actions are designed or redesigned, on costs, communications, etc.

According to Windham (2000), there are ‘three primary purposes of all assessment work in the field of education policy: to assess the nature and magnitude of the opportunities and constraints that face the systems that provide education and training; to assist the government, the private sector and the individuals in establishing priorities within a resource-constrained environment; and to specify options for exploiting the identified opportunities and for dealing with the identified constraints. Alternative goals and strategies will be proposed from which government and its

Figure 1: Basic concepts in evaluation

---

² Evaluation criteria and their definition are based on documents from the European Commission’s Joint RELEX service for the management of Community aid to non-Member States (SCR – now EuropeAid Cooperation Office). However, identical presentations can be found with the World Bank, the Organisation for Economic Cooperation and Development (OECD), United Nations organisations and others.

³ Ibidem
partners can make concrete recommendations for policy reform."

Evaluation in this sense is no longer just a reflection on the outcomes of some project or programme, but a type of consultancy work providing advice and guidance and the basis for informed decision-making. This makes criteria, such as relevance, effectiveness, efficiency, impact and sustainability relevant or irrelevant. However, project/programme evaluation and policy evaluation are often closely linked and feed each other.

Also, international literature suggests a trend away from evaluating the impact of single projects. A review of evaluation practice in education (OECD), VET (Grubb and Ryan, 1999) and labour market policy (Schütz et al., 1998, among others) signals a significant change in evaluation approaches towards target-oriented approaches in labour market policy and a systemic approach in VET. These shifts are caused by a growing awareness of the restricted contributions of single projects and the limited impact they can have, even if highly sustainable. Reviews of international cooperation (e.g. King and Buchert, 1999) similarly observe a shift away from project-based funding to policy assistance. As relevant publications suggest, the Danish international development assistance (Danida) and the UK Department for International Development (DFID) are increasingly following this approach, as is the European Commission in foreign aid policy vis-à-vis the countries and territories eligible for the MEDA programme (the Euro-Mediterranean partnership). However, while a pretty robust methodology for project or programme evaluation exists (even up to the point that there is now an evaluation of the evaluation methods), this is not the case for policy evaluation (Grootings, 2000).

We admit that the distinction between policy, which broadly encompasses political intention and may be expressed through several programmes, and particular programmes, which are themselves made up of many projects, may be somewhat arbitrary. Sometimes a policy is virtually identical with a programme. However, the rationale for this distinction is to emphasise that we are more and more involved in broader policy evaluation. The evolving human capital imperative and its interrelationship with education encourages decision-makers to expand or alter their views regarding the nature of useful evaluation. Policy-makers increasingly want information regarding systemic, not only operating unit, performance. This is also the case for the partner countries with which the European Training Foundation (ETF) (\(^{(4)}\)) is working. They ask for advice related to future policy directions rather than just support with the design of a certain project.

The ETF has both an analytical and a developmental role, i.e. directly or indirectly the Foundation is involved with the implementation outcomes of its evaluations. The main tasks include:

(a) to identify gaps and needs for intervention or assistance;
(b) to assist with the identification of needs or the formulation of policy objectives;
(c) to identify and design reform proposals in transition countries;
(d) to review progress and evaluate the outcomes of VET reform programmes;
(e) to manage the dialogue with policy-makers and other stakeholders from our partner countries.

The analysis of an individual project or a limited number of themes would be too restrictive and the strategic perspective too narrow for the above assignments. Hence, VET policy evaluations are of strategic importance for the Foundation or, generally, for anyone who assumes, or prepares for, a development role in a domestic or a foreign context.

We explored the contribution of research and examined a number of examples from transition countries with a view to illuminating, where possible, VET policy evaluations. In line with the remit of the ETF, we have concentrated on, and used examples from, transition countries in central, eastern and south-eastern Europe, as well as from our Mediterranean partner countries.

The aim of our study is to clarify underlying concepts, methods and possible tools used in

\(^{(4)}\) The European Training Foundation is an agency of the European Union whose mandate is to support VET reform in its 40 partner countries, comprising central, east and south-east European countries, the successor States of the former USSR and Mongolia, as well as our Mediterranean partner countries and territories
VET policy evaluations. Our search for more conceptual and contextual clarity and possible instruments is, not least, connected with the hope that a shared understanding among evaluators can contribute to an increased coherence of policy recommendations and a clearer rationale for establishing priorities.

Chapter 2 points to the fact that there is rarely transparency for the methodologies deployed in (vocational) education (and training) policy evaluations, quoting the example of the Organisation for Economic Cooperation and Development (OECD). However, in a number of evaluations carried out or commissioned by the ETF, there is evidence of a building-blocks approach which was used to structure the evaluation and which is further explained in this chapter.

Given the limits of methods derived from management thinking, we argue in favour of humanistic approaches to understanding VET systems. Describing the hermeneutic approach, Chapter 3 discusses how a particular system can be understood. It also refers to the need to put certain phenomena in their historical context (the structural-historical approach) which helps to understand why existing practices, functions and structures are as they are. Finally, we refer to the continuous interaction between understanding and explanation, as explanations are needed when we fail to understand.

Chapter 4 examines VET as a (sub)system and elaborates on both the concepts of internal and external consistency of VET systems and the inherent tensions between flexibility and consistency, which an evaluator must grasp when undertaking VET policy evaluations.

Chapter 5 places VET policy evaluations in the context of system change. Viewing policy evaluations as a consciously organised, critical reflection and learning process involving the social actors and, hence, as formative and constructivistic, we emphasise their indispensability in any development process.

Chapter 6 explores the possibility of establishing predefined lists of evaluation criteria, values, policy priorities, indicators or other assumptions underpinning VET policy evaluations and argues why these should be subject to negotiations with the social actors involved. Furthermore, we review first experience with the use of the benchmarking approach on an international scale and point to its potential and limits.

Chapter 7 includes the executive summary and outlines needs for further research.

Finally, there is a list of references and a structured bibliography of background literature on the topic.

1.1. Methods of work

A working group was set up within the ETF to examine and compare concepts and methods used for VET policy evaluations. Group members reached understanding through collaboration and mutual learning. The study method was a structured collaborative writing process where working group members generated the results of their analysis during writing. Group members engaged both in an extensive literature review that included the European Commission, OECD, Cedefop, the World Bank, ILO, BIBB, the Leipzig and Vienna groups of transformation researchers, plus selected evaluation reports from our partner countries and many others.

The project also benefited from workshops involving both working group members and international experts. Two expert meetings were organised: one in February 2001 bringing together representatives from major international organisations operating in transition countries, and a second in November 2001 with evaluation experts from European countries. Moreover, all group members were involved, over the past two years, in a number of VET policy evaluations in various transition countries.
2. The limited transparency of concepts and methods used in VET policy evaluations

VET policy evaluations cover a multi-disciplinary, highly complex terrain. Concepts and methods are derived from sociology, economics, psychology, history, pedagogy, philosophy, management and organisational theory, comparative education and VET research, etc., whose scientific theoretical positions and methods have to be integrated. Furthermore, using VET policy evaluations as a platform for identifying, designing or redesigning and monitoring policies or programmes/projects, it was clear that some kind of implementation research would be as important, since, unavoidably, some form of intervention, or a type of applied, forward-looking research, is involved.

However, the methods and tools that support first analysis and then implementation are explored in international literature only in outline and often remain implicit rather than explicit. That literature conveys other aspects: the values and assumptions that underpin policy; the information and analysis that feeds in and out of policy formulation and determination; and the research base that exists or is created to support policy determination and implementation.

2.1. OECD

The OECD as one of the major international players in the field of (education) policy evaluation suggests that theoretical underpinnings derive from the choice of experts rather than being an explicit part of an overall approach. At the Foundation’s workshop in November 2001, Whitman (2001) set out the approach of the OECD to system reviews: client country consensus; a framework of values and criteria; a careful choice of experts carrying their own cultural and research hinterland; and legitimacy from both the client country and the general secretariat of the OECD. The main purposes of the evaluations as such are:

(a) to link educational policy to national, regional and economic issues;
(b) to place each country’s system in an international comparative perspective;
(c) to make policy recommendations to government (being of greater benefit to partner countries and other donors by not itself representing a donor organisation).

While the two underlying themes (or values) are:

(a) the emphasis on active labour market policies and their effectiveness;
(b) improving labour force skills and competences through wide-ranging changes in education and training systems.

Whitman’s exposition is useful, since elsewhere in OECD literature (even in country review outcomes) there is rarely transparency for the methodologies deployed.

2.2. The building-blocks approach

Analysing evaluations that had been carried out or commissioned by the ETF, we found that some form of the building-blocks approach had been applied in certain cases. The notion of building blocks was first mooted by Parkes (1995a). The approach was used for evaluations both in Bosnia-Herzegovina and in Moldova as a basis for constructing the VET ‘green’ and ‘white’ papers. Furthermore, the approach was used in the Maghreb countries (Algeria, Morocco, Tunisia) as a common framework for comparison, comprehension and subsequent cooperation; and in Uzbekistan as a means of evaluating policy determination and implementation. The approach has also appeared in project terms of reference drawn up by the ETF, most recently for Armenia, Montenegro and Turkey, etc.

The metaphor ‘building blocks’ referring to an effective VET (sub)system was introduced first to key actors from Baltic countries and then to a review of curricular reforms of the 1990s in nine central and east European countries. In their initial use, building blocks were defined in terms of functions or process. It was suggested that all successful VET systems, independent of their
cultural and historic contexts, need to possess the following ingredients (ETF, 1996a) (Parkes et al., 1999, p. 27):

(a) ‘to be able to define occupational sector priorities (on the best possible evidence available);
(b) to be able to identify the appropriate occupational sector competences and skills required (and to construct the institutions and tools to do this);
(c) to be able to turn these into curricular profiles and programmes and measurable standards;
(d) to deliver these at school level (including the capacity to transfer from pilot to system level);
(e) to help make the processes attractive to students and teachers (transferability, visibility and portability of qualifications for students and working conditions for teachers);
(f) to provide for timely and effective feedback through evaluation, monitoring, quality control and tracer studies of school-leavers.’

It was recommended that these elements be undertaken in the context of establishing the balance between general and vocational provision and in the context of transparent and accepted approaches to standards, certification and qualification. They also have to be related to other factors: financing mechanisms; changes in the location of decision-making; a credible research base; and the development of management capacity, the acquisition of appropriate tools in management, as well as curriculum development per se.

While we can agree with the specified functions, which need to be undertaken in any modern VET system, there are some problems with the use of metaphors. It is, in essence, a functionalist analysis and should, hence, be called this way.

Added later on to the notion of function or process was the development of an architectural element. This was to offer an operational model for transition country working groups set up to analyse existing structures and practices and to make proposals for change in a consistent way, in order to have a common agenda among ministries and agencies. Eight topics (building blocks) were formulated:

(a) educational management and administration;
(b) curriculum, assessment and certification;
(c) financing of VET;
(d) the labour market and social partnership;
(e) educational standards and quality control;
(f) in-service teacher training;
(g) legislation;
(h) labour market and adult education.

Each of the building blocks was then analysed in detail, for example, by answering the following three key policy and action questions for the building block of legislation:

(a) should there be separate or integrated legislation for general education and VET?
(b) should existing legislation be left largely unchanged with modification to regulations only or should there be clear, new legislation, indeed an Education Reform Act?
(c) if new legislation, should it be short, simple and transparent, leaving the detail to regulation?

Judging from the evaluation report of the respective Phare project under which the exercise was undertaken (Nielsen, 2001), Bosnia-Herzegovina can be considered a very successful example, in which the building-blocks approach was used to structure and systematise the analyses by the actors involved and then to provide a framework for a concept paper for legislation. The building-blocks approach helped establish a common conceptual grasp of the issues at stake and a common language in which a relatively large group of key actors could discuss structures, functions and institutions of a VET system in transformation. According to the report, it turned out to be a useful tool for specifying well-grounded and specified VET reform strategies, encompassing all elements of VET systems.

In the cases we reviewed, the building-blocks approach was an attempt to provide a simple, transparent vehicle for managing the dialogue among country key actors, donor representatives and consultants, i.e. between the researcher and the practitioner. It was an attempt to balance actual complexity with simplicity and transparency in the use of tools, which is consistent with Grubb and Ryan (1999) who recognise the need to find a way to express a holistic approach in a manner convincing to policy-makers: ‘In the selective use of evidence, policy makers have avoided the complexities of theory and methodology. Instead they have often used social science results in simplified ways, to tell almost
commonsensical stories consistent with the particular ways in which they want to frame problems. In this habit they are reinforced by journalists, whose methods of personalising issues by concentrating on the experience of individuals – story telling – makes the issues more vivid in the public mind, but [...]’.

However, as the ending indicates, the approach also leaves a number of unresolved questions which will be discussed in the following chapters.
Instead of methods derived from management thinking, we agreed on a phenomenological approach, in which the emphasis is on describing phenomena, as they appear to the observer, instead of attempting to guess or theorise about what might be behind these phenomena. The basis is first and foremost to establish a genuine understanding of what you observe.

Having said this, we asked ourselves whether we can identify a number of methods or tools (Krücken für die Gedanken) actually made use of by experienced VET policy evaluators. Which methods do the latter use to come to an understanding of VET? And what is the correct relationship in evaluations between understanding and explanation?

3.1. The hermeneutic approach

In more specific terms, we may talk of a hermeneutic approach that covers a number of stipulations and various methods related to the understanding and interpretation of phenomena. Hermeneutics play an important role in human and social sciences; an important thinker in this field is Hans-Georg Gadamer (1972).

A hermeneutic approach seeks to elucidate what happens when you try to come to grips with a text or any other phenomenon in society. In the description of how to come to an understanding, a specific concept is normally used: the concept of the hermeneutic circle. This refers to the fact that the particular element is understood in terms of the totality, while the totality is, at the same time, understood through its constituent parts. Through our interpretation work we will achieve a more and more secure, detailed and varied understanding by way of a continued circular alternation between studying parts (system elements) and totalities (the total system). There is an additional circular process in the interpretation work: the interaction between the preconception (Vor-Verstehen) of the phenomenon, which the researcher/evaluator brings with him/her, and the new understandings worked out through the process.

An important point for discussion on hermeneutic approaches is the question of whether objectivity is possible. Some hermeneutics would argue so. Karl Popper, for example, claims that you can avoid wild interpretations and arbitrary postulations, which are a clear risk in hermeneutic understanding, by going through a number of well-conceived loops for testing your interpretation hypotheses. However, Gadamer questions this position.

Gadamer (1972) insists that it is not possible to come to an understanding of a phenomenon, for instance a foreign VET system, just by following certain methodical procedures. According to Gadamer, the interpreter cannot abstract him/herself from history and culture in trying to understand, as the researcher is her/himself part of history and culture. Understanding always takes place between two entities both placed in their own historical and cultural contexts. What the researcher/evaluator must do is openly recognise and clearly articulate this tension. Another conclusion for us is the involvement of evaluators coming from the reviewed country itself or, in the context of external peer reviews, from countries with similar education systems.

3.2. The structural-historical approach

Building blocks or ingredients may be a useful, simple tool to structure (or manage) evaluation, but the approach also leaves a number of unresolved questions. One is the use of metaphors instead of clearer concepts; but let us continue metaphorically. If these (eight) building blocks are the bricks, what then is the mortar holding the building together? Can we provide a conceptual framework for grasping not only the parts (building blocks) of the VET system, but also its historic roots and dynamic relationships? Which architecture structures the VET system under
observation and gives the constituent parts their specific meaning? Are there different systemic logics behind the way building blocks are put together in the different countries? How can we dig deeper into the inner driving forces behind the configuration of the building erected over many years?

A structural-historical approach (Strukturell-genetischer Zugang) could be applied. Here, the analysis would start with an overview of existing practices, functions and structures as they appear to the (experienced) evaluator. Essential structures are then traced back to their origin (genesis), and an understanding/explanation of the historical context is reached in which the specific phenomenon was born or established. Described through the comparative, structural-functionalist mirror, for instance, the overall Danish VET system may seem a wonderfully harmonic building that consists of highly refined building blocks (Cort, 2002). But analysed historically (through the crystal ball instead of the mirror), this VET system is a patchwork of political compromises reflecting what was achievable under given historical-cultural circumstances over the last 100 years. How do we come to such an understanding of other VET systems in evaluations, using the building-blocks approach?

Another question related to the understanding of the systemic logic behind the building blocks is the challenge of change. How do VET systems change and what are the driving forces behind change processes? Even if we get the description and understanding of the building blocks right in VET policy evaluations, we cannot at all be sure to have fully established the capacity to predict change or to formulate the right intervention strategies to achieve change. Is there a strategic lever at force in VET system change? What is (empirically and theoretically) known about dynamic forces, catalysts for change, etc., and what is the right balance between top-down (politics) and bottom-up (market) approaches? What follows is that the building-blocks approach must be accompanied by a deeper insight into organisational and institutional processes of change. We will come back to these issues in Chapter 5.

Yet another problem is how to cope with the risk of preconceptions and nationally biased understandings of what makes up a good VET system. It is not uncommon to see the foreign evaluator measure the distance between VET reform initiatives in a transformation society against the advanced state of his/her own system. The first question is: are the building blocks per se culturally bound? The next question is: how can we make use of them in evaluations, being nevertheless reasonably objective and transparent? One way around this dilemma is to make statements of values behind evaluations explicit, when using the building-blocks method. Chapter 6 will deal specifically with this issue.

3.3. Understanding and explanation

It has been argued above that the phenomenological/hermeneutic method is primarily used to achieve an understanding rather than for explanations. If we accept this point of departure, then we would argue that we should not give up using the building blocks as an interpretative framework between the evaluator and the VET system under observation, but we should definitely put more emphasis on explanations.

Understanding and explanation are different ways of (re)cognition. Understanding is a more immediate experience or recognition of a phenomenon. Sometimes, if a phenomenon is difficult to understand, we need explanations to come fully to grips with, say, a component of a VET system and its precise role within the system. When you explain something, you establish some distance to the phenomenon under observation. On the basis of explanations, you may often understand the phenomenon better or view it from a different perspective. To explain something is to state causes behind the phenomenon under analysis, external to the phenomenon itself. To understand a phenomenon is to give a reason for it, which is internal; meaningful understanding presupposes a communicative community between the evaluator and what is being evaluated. Understanding a phenomenon, in casu the VET system, means that one can supposedly see (recognise, realise) how the elements fit together, and see the meaning of the phenomenon (intention, purpose, function), often in an immediate, even intuitive way.

The relationships between understanding and explanation are tricky; one form of recognition is
not better or more correct than the other. The argument here is simply that we should reflect on these connections and deliberately seek to establish a continuous interaction between understanding and explanation.

The following are but a few examples of useful explanatory frameworks:

(a) economic or labour market or sociological laws and other forms of determination of the concrete conditions under which a phenomenon is active, for example, the conditioning factors behind VET systems, such as demography, trends in labour market developments and broader educational policies;

(b) functionalist explanations referring to the totality in which the phenomenon to be explained is placed and a description of the precise function it serves; for example the role of a national VET system for societal functions such as qualifying, socialising, sorting, and being a depository of young people;

(c) historical-cultural explanations behind the occurrence of different institutions and practices in contemporary VET systems, for example: why is there such an assortment of independent VET providers (production schools, etc.) in Denmark compared to Sweden? Because of Grundtvig and the free school and folk high school tradition;

(d) structural(ist) explanations of phenomena which can only be uncovered through an in-depth analysis, such as the specific connections between the economic, social and political systems in social structures, making transition of VET systems in many former communist countries very difficult;

(e) system-analytical explanations showing how changes in one component will have spill-over effects on other system components or on the total system, cf. the following discussion on internal and external consistency, such as the role played by access to higher non-university education as a determinant for the attraction of VET.

Central to this reasoning is that understanding and explanation feed each other, that understanding is as important as ever. Four examples may illustrate the problems evaluators face:

(a) the trendy catchwords broadcasted by the OECD, the European Commission, the World Bank, etc., in recent years have mainly focused on modern competences, such as adaptability, changeability, responsibility for one's own learning, etc.;

(b) the cultural school of thought in European VET research (around Anja Heikkinen from Jyväskylä University, Finland) argues that this is a frontal attack against VET institutions which the labour movement has been able to establish only after more than 100 years of hard struggle. The change of emphasis from the sociological qualification to the psychological competence is bound to remove social control from VET. And it undermines the social category of the skilled worker;

(c) the lifelong learning concept is now recommended everywhere. As evaluators we tend to measure VET system developments against the indicators found in or around the European Commission’s memorandum on lifelong learning (European Commission, 2000a). We tend to do so even in countries where the most rudimentary conditions for initial VET are not in place;

(d) in our thinking on modern and adequate VET provision, we tend to promote output-based, modularised and credit accumulation-based models with little emphasis on process factors (curricula, teachers, pedagogics, time-servicing requirements). We promote the Anglo-Saxon models without remembering why they had no choice but to configure the system the way it is;

(e) the complete destruction of the East German polytechnic education system was the result of a lack of understanding of the intrinsic qualities of that system. Entirely lost is the expertise developed in work-based learning principles (Lernen in der Arbeit), didactics and practices, which were further developed in the socialist countries at the time than anywhere else.

Therefore, besides developing further the building-blocks approach by including explanations, and in order to avoid hobbyhorse explanations, we have to cultivate, individually and in a community of VET evaluation practitioners, a new attitude towards argumentation. We need a more open discussion practice: a discourse.
In our effort to understand VET systems (or the system context of VET policies), it seems obvious that we have to apply some kind of system analysis. However, it is quite difficult to define a VET system:

(a) how is the system universe to be delineated?
(b) what are the constituent components, parts, units?
(c) which relationships exist between units?
(d) what matrix of dominance patterns the interplay of units?
(e) where are the boundaries of the system located?
(f) which relationships (metabolism) exist between the system and its environment?
(g) what is the prime mover or the strategic lever for change?

The set of questions is very complex and forms part of a broader scientific context. It is helpful to go back to the roots of system analysis and to analyse, from a theoretical standpoint, issues related to internal consistency and external consistency of VET systems and their consequences for VET reforms, and hence VET policy evaluations.

4.1. System analysis

The VET field is short of a comprehensive theory and in many ways remains a very young research field. Perhaps one could classify the research tradition as follows. First, there is work within the tradition of general educational theory. A second group of topics relates to the study of those organs of the VET system which exist by law and are recognised in official documents about VET structure, functions and actors. Third are studies which take these formal institutions as given and explore the forces, including the informal institutions, which set them in motion and link their activities.

A whole range of topics is opened for discussion, as soon as one asks the question: what makes the VET system tick? One can then explain things in terms of functions; we can look for systemic interrelations between the data, with a view to discover what goes with what, what changes in parts produce what reactions in the whole and in other parts. We can treat functions as heuristic, as a notion that will help us to spot relationships and so to define structure as a system of parts with internal and external relations. An interesting example of this approach was the early 1990s research component of the European Union’s (EU) PETRA programme on innovation of VET Programmes. The research question asked was ‘through which mechanisms are VET Programmes renewed?’ instead of asking ‘who does what when?’.

General systems theory is a field of thought that has been growing in strength over more than 50 years. Its origins go back to biologists who opposed reductionism, i.e. an analysis of living things merely as a sum of parts (which was a view contradicted by observation). Parallel movements existed in the psychology of perception and learning (Köhler, 1959) (5), and in social anthropology, in which structure/function analysis already began to flourish in the 1920s. But the systemic approach is more than just an example of the use of biological analogies to describe society. Much impetus was given by mathematicians and electrical engineers during World War II and immediately after.

General systems theory would not claim to be empirical; rather it claims to be axiomatic, in that the mathematics conceived first for the electronic systems has been so far generalised that it is applicable to any systems whatever, provided that the rules are followed: systems are isomorphic.

(5) Interestingly, the Gestalt concept has been applied very fruitfully by German VET researchers; cf. Felix Rauner, University of Bremen/Institut Technik und Bildung (ITB), and Gerald Heidegger, Universität Flensburg/ Berufsbildungsinstitut Arbeit und Technik (BIAT), who make use of the Prinzip Gestaltung as a way to conceptualise innovation in VET.
The meaning of the term ‘system’ is often confused. The most general definition is formulated by the founding father of the general system theory, von Bertalanffy (1950), after he had noticed that in physics, biology, psychology and social sciences it was no longer acceptable ‘to explain phenomena by reducing them to an interplay of elementary units which could be investigated independently’. ‘A system is a set of units with relationships among them; the word “set” implies that the units have common properties – the state of each unit is constrained by, conditioned by, or dependent on the state of other units. The units are coupled. Moreover, the system as a whole has “got something” which its components separately have not got. Systems may be concrete or abstract, systems may be “open” or “closed”, and systems can be analysed on the principle of systems and sub-systems – systems within systems within [...] ultimately a wholly general system. There are theoretical difficulties about this ultimate system: but the problems of supra-systems and sub-systems, levels and boundaries, etc., are generally manageable in a practical way.’

The German sociologist/philosopher Niklas Luhmann (2001) has built up a general system theory of society, where the educational system is a subsystem having its own logic, laws of motion, discourse, etc., and is even further differentiated.

4.2. VET as a subsystem

VET as a subsystem can be seen like other living organisms as an open system in a steady state; it depends on self-regulating mechanisms to maintain its boundaries and its continued existence within these boundaries (meaning primarily boundaries which distinguish the VET subsystem from other (sub)systems of the same society). The steady state depends on a balance of inputs and outputs. The inputs are demands and support: support makes the system strong enough to process demands and to produce outputs in the form of qualifications. The VET system sits in an environment – the total social system – and there is a continuous feedback into the VET subsystem, signalling whether outputs produce good or adverse effects in the environment. A systemic approach to VET evaluations focuses on the analysis of relationships, communication channels, and responsiveness and adaptability, based on the fundamental understanding that changes in one component lead to changes in other components and in the system as a whole.

It is important to understand the specific logic of the VET system, to explore whether the system is internally and externally consistent, and in which way the system needs to be regulated to balance available tensions.

4.3. Internal and external consistency (6)

Improving VET is possible only through a systemic approach, and we provide arguments below why this is so. Evaluations would have to establish to what extent VET policies have followed a systemic approach, applying in particular the concepts of internal and external consistency. At the same time, the evaluator needs to be aware of important system-inherent tensions that are not at all easy to reconcile.

In transformation societies the main driver for reforms is the fact that the whole logic of VET systems has radically changed. Before the collapse of the planned economies, there was a close relationship between the human resource needs of companies and vocational schools. Based on a precise manpower forecasting (planning) model and a vocational guidance system whose main aim was to steer young people into certain occupations in line with available job positions, schools could almost exactly cater for the needs identified by companies. People tended to remain in their jobs for a lifetime. With the change to market economy principles, this system no longer worked. The labour market underwent dramatic changes within a short period, calling for different skills than those catered for by schools. Schools had been slow to adapt, results

(6) Section 4.3 and subsections 4.3.1–4.3.3 draw heavily on Durand-Drouhin and Bertrand, 1994.
include a large number of young unemployed and inadequate skills upgrading or retraining system for adults. There are tremendous conceptual, political, financial, structural, organisational and content-related problems to be solved at the same time. This concerns internal factors of VET, but above all the need to establish new mediating mechanisms between education and work to replace the former administrative ones.

Most countries have placed new emphasis on the need for more responsiveness of VET systems to changing needs for education and training. Two key concepts, flexibility and consistency, are associated, to variable degrees and in variable combinations, with the search for responsiveness.

Responsiveness is needed with regard to:
(a) constantly changing and largely unpredictable requirements of labour markets for skilled labour;
(b) growing demands by young people for higher levels of education and for more diversified and individualised learning processes and pathways.

Responsiveness can be interpreted and may be achieved in a variety of ways. In institutional and management terms, three major developments can be identified.

First, there is a clear trend towards more decentralised governance and management of education systems in general and of VET in particular. Examples include France, which in the early 1980s, delegated authority and responsibility for the location and equipment of vocational schools largely to the regional level, as well as several Scandinavian countries which transferred responsibility to local authorities and schools. The latter were given a great deal of initiative, including curriculum design and implementation, in order to satisfy diverse needs in local economies and of individual students.

Second, the role of central government in educational planning is declining, while industry is called upon to participate more actively in the design and provision of VET and of related certification and qualification systems. While the participation of employers and their organisations is expected to contribute to improved responsiveness of VET to both quantitative and qualitative demands in the labour market, the involvement of trade union organisations and workers' councils varies from country to country.

A third trend is the development of training in enterprises and in private institutions and the related emergence of a training market where public and private institutions may be competing. This trend is related to the growing importance of adult training, which can itself be seen largely as a response to the demand for a more flexible workforce.

In terms of the structure, content and organisation of VET, the search for responsiveness has in particular led to:
(a) the diversification of training programmes and pathways;
(b) the broadening of curricula;
(b) the increased duration of initial training; and in some countries:
(a) the modularisation of courses and certificates;
(b) the introduction of competence-based learning and assessment.

All these developments are assumed to contribute to the adaptability of individual workers and to the flexibility of education systems. However, they also raise serious questions about the types of qualifications which young people acquire, about the costs and effectiveness of education pathways and about the consistency of the rules of the game prevailing in different sectors of national education systems and in the labour market.

These issues seem to be of interest to central and east European countries, which are going through a process of transition, from centralised and rigid education and training systems towards other types of systems, expected to meet the demands of market economies. Considering that, in the present context of uncertainty, these demands are to a large extent unpredictable, the need for adaptability and flexibility would seem to be even more pressing in countries in transition than in most west European countries.

The implications of such orientations concerning the structure, organisation and management of VET systems should be carefully examined. There are certain tensions between objectives of flexibility in VET on the one hand and consistency on the other. The underlying question is whether flexibility and consistency can be conceived of as complementary rather than contradictory features of education and employment systems and which regulatory mechanisms deserve particular attention in this respect.
This question will be considered from three points of view: the internal flexibility and consistency of education and training systems; the external responsiveness of education and training systems and the consistency between their internal organisation and objectives and the functioning of labour markets; and, finally, different approaches to system regulation.

4.3.1. Internal consistency

4.3.1.1. Consistency between subsystems and clarification of pathways

Introducing flexibility by diversifying education pathways has also meant the creation of ever more sinuous, long and unfocused educational careers. Moreover, while countries have moved closer to flexible and open systems, their strategies of reform have tended to concentrate on one type and/or level of education and training at a time, rather than explicitly aiming at complementary transformations of all stages and in all relevant areas. And yet the consistency of learning opportunities and pathways across types and levels of education and training is essential.

One important aspect of the internal consistency issue is the relative attractiveness of VET vis-à-vis academic education. Many efforts to attract young people into vocational streams do not take a sufficiently broad view of the relationship between the different streams. For instance, one of the reasons for the preference for academic streams is the wider opportunities that they provide for further studies. But attempts to open up access to higher education to vocational students often fail, because universities operate on a completely different logic and use different selection criteria.

The two examples illustrate that improving VET is possible only through a systemic approach, encompassing all the components of the system. Such an approach may conflict with a policy which would promote a high degree of decentralisation and of privatisation.

4.3.1.2. Regional consistency

When far more autonomy is given to schools, not only from the point of view of management but also of teaching content, the problem of possible imbalances and inequalities between regions, schools and individual students might arise. For the central and east European countries, with their strong egalitarian tradition, this raises the question of how to overcome such problems.

Most English-speaking countries, with traditionally highly decentralised systems, shared a concern about the multiplicity of assessment and certification systems. This was one of the reasons why the UK and other English-speaking countries engaged in the development of national vocational qualifications. They are not, in contrast to many continental European countries, related to remuneration systems through collective bargaining at national or sectoral levels. This is considered as an additional element of flexibility. However, the value of such qualifications and their motivating effect on young people and adult learners may also be weaker than that of more holistic and rigid qualification systems, such as exist in Germany.

4.3.1.3. Consistency for the individual learner

In countries like the UK, the definition of national vocational qualifications is associated with a modular approach to training, allowing students to choose their own personal range and sequence of course/modules (which can be provided by different institutions) and to accumulate corresponding certification units. Young people are thus – to some extent – encouraged to define their own individual learning pathways. Modularisation is expected to allow for more rapid responses to changing labour demand. Flexibility is pursued both from the point of view of students’ interests and of labour market requirements.

One of the possible drawbacks of modular approaches is the risk that young people will acquire only loose collections or incomplete puzzles of narrowly defined skills or competences. In such systems, responsibilities for the level of skills and the relevance and completeness of skill profiles with which young people leave the education system is largely left to young people whose information, strategic planning capacities and financial possibilities vary greatly.

4.3.2. External consistency

External consistency refers to the relationships between educational objectives and the economic and social context. In reality, there are potential problems, such as the fact that the employment system is not always able to articulate its needs and that there are tensions between
the pedagogical logic of school systems and competence needs of employment. Furthermore, the drive to maintain a holistic approach of training for an occupation, and to develop national qualification systems as a means to increase internal and individual learner consistency, may conflict with deregulated labour markets and the short-term recruitment and training practices of companies. It is also realised that education cannot by itself solve the problems of the employment system. See, for example, Hodgson (2001), Paquet (2001), Miegel and Nölke (1996), as well as Grubb and Ryan (1999): ‘Education without suitable employment, and specific skill training without jobs requiring such skills may be valuable in their own right but they cannot enhance economic conditions. And so the other conditions necessary for education and training to be effective – the employment necessary, the capital required, the institutions that can give these arrangements some permanence – also need to be carefully understood, and the most successful programmes carefully consider the nature of local employment.’

In evaluations commissioned by the ETF we have found a number of examples of internal and external consistency lacking in the VET policies implemented by transition countries of central and south-eastern Europe. We include some of them in Section 4.4. In contrast, in an evaluation in Croatia, we have tried to highlight system connections which need to be considered when designing a consistent (in the given case, EU-funded school headmasters’ and teachers’ training) programme (Section 4.5).

4.3.3. Regulating mechanisms
The foregoing observations underline the need for regulation of VET systems themselves and of the interface with their social and economic environment in order to find the narrow path between two dangers. On the one hand ‘a centralised, uniform and rigid system would not meet the needs of a modern economy, would limit adaptability and threaten the ability to innovate’ (Wolf, 1993). On the other hand, leaving individual schools and training institutions without any common framework of objectives and organisational conditions would not satisfy the long-term economic goals of quality and effectiveness and is likely to increase geographic imbalances and social inequalities. Ideally, successful system regulation requires both clear policy orientations and effective regulation mechanisms, which need to be reviewed by VET policy evaluations. Clear policy orientations are particularly important in a context of strongly decentralised systems. As far as possible, policies should be based on both a sufficient capacity of technical analysis and a high degree of social dialogue. In institutional terms, this implies the need for:

(a) ‘a framework for consultation among the various actors, which, at the national, local and/or sectoral level, guarantees some degree of continuity, coherence and consistency, especially between education systems and labour markets;

(b) some kind of research and development structure, providing information and technical support to decision-makers; [...] this includes [...] the development and monitoring of labour force information systems, the development of learning theory and appropriate teaching and training methods, etc.;

(c) clearly defined and agreed financing agreements which are a major element of the system regulation. This raises policy – or political – questions about the respective role of governments, enterprises and individuals; [...] 

(d) finally, there is the broader issue of evaluation. The more decentralised the system, the more necessary it is to set up mechanisms allowing for the evaluation of its output and effectiveness. Certification is an important element of evaluation and it has to be seen, not only from the point of view of recognition on the labour market. The above comments concerning the limitations of a market-driven approach do not deny the fact that the ultimate value of qualifications depends on the market and not on administrative decisions. But the need for consistency pleads for some degree of cooperation between the actors and for some form of institutionalisation in order to regulate the market and transcend its most immediate and specific demands.’ (Durand-Drouhin and Bertrand, 1994).

This brief overview of requirements or conditions of successful regulation refers to an ideal situation. In reality, it is clear that such requirements cannot easily be fulfilled, especially in the present context of transition countries. To
mention only a few points, the institutions and the identity of the various actors are only emerging and the inadequacy of technical instruments is reinforced by the lack of resources. But, in one way or the other, responses to these problems, adapted to national cultures and traditions will have to be found.

4.4. Examples of lack of internal and external consistency of VET reforms in Bosnia-Herzegovina, Estonia and Romania

To understand the importance of systemic consistency (and to which phenomena the concepts refer) when evaluating VET policies, examples shall be provided from three transition countries. More examples are available from the ETF.

4.4.1. Teacher training in Bosnia-Herzegovina (Durand-Drouhin and Bertrand, 1994)

Part of the questions during the evaluation were addressed to actors and institutions outside the EU Phare VET reform programme with a view to assessing whether framework conditions would be conducive to the longer-term sustainability of the pilot project. Obviously, one central factor here is vocational teacher training, which had not originally been a project component. The pilot schools had been asked only to specify teacher-training needs in relation to newly introduced curricula. This was changed later, and a budget reserve of EUR 200 000 allocated to run a programme for teacher and management training.

The following questions were asked in three institutions visited: were existing teacher training institutions involved? And were they familiar with the curriculum principles applied in the Phare VET reform programme?

4.4.1.1. Pedagogical Academy in Sarajevo (PAS) in charge of education of primary school teachers

The PAS department for Chemistry was visited. The reason for choosing chemistry was that in almost all pilot VET schools visited, it was argued that pupils are below standard and that particularly the new Phare module in chemistry is very problematic and probably the most difficult subject. The two assistant professors interviewed had never heard of the Phare VET programme and its renewed module in chemistry, neither were they aware of poor end results of pupils in primary schools.

The conclusion is that a clear problem exists here in terms of the vertical integration of the education system, which has a negative impact on Phare programme results. Only the Ministry of Education can solve this problem by promoting internal consistency measures.

4.4.1.2. Sarajevo University, Faculty of Philosophy, Pedagogical Department, in charge of pre-service VET teacher training

An interview was carried out with the Dean, Prof. Dr Vlado Sucic. Pre-service vocational teacher training is subject-based and covers the full range of subjects taught in secondary schools. Teacher training is carried out according to a consecutive model: first, students graduate in a subject (e.g. engineer), then there is a six-month pedagogical-psychological-didactical-methodical training course, which includes school practice. After two years of teaching practice at a VET school, the candidate takes the final state examination and becomes a certified teacher. This model is also offered for teachers of practical subjects.

Until now, the University's Pedagogical Department had always been involved in the design of new curricula for all subjects taught. This was done to ensure that new subject matter and innovative methods were integrated immediately into teacher training. However, the department was not involved in the Phare programme and did not even know of its existence.

The fact that no inputs from changes in VET equipment, curricula and methodology are fed forward and channelled directly into pre-service teacher training is a serious problem. The development model is not optimal: too little emphasis is placed on the teacher training system and the training of teacher educators. In addition, VET teacher training would capitalise on becoming an important research field in Bosnia-Herzegovina. Any support programme should focus on the teacher educators to ensure a multiplying and lasting effect and to ensure that activities are 'anchored up' within existing providers of pre-service (and in-service) teacher training.
4.4.1.3. Pedagogical Institute, Canton Sarajevo, in charge of in-service VET teacher training

Tasks of the pedagogical institutes in Bosnia-Herzegovina in general include:
(a) advising ministries of education on curricula and new legislation;
(b) collecting statistical and other information from schools;
(c) evaluating the quality of schools;
(d) advising on the appointment of teachers and assessing and confirming their qualifications;
(e) providing professional development opportunities for teachers and setting up in-service training courses;
(f) evaluating school development plans.

The institute has traditionally promoted a subject-based approach to curriculum design and delivery, which led to some opposition with respect to the newly introduced module-based curricula in Phare pilot schools. An interview was held with Mr Vreto Sadik, inspector in the field of mechanical engineering at VET schools. His specific responsibility covers subject matter and training practice in these schools. He is familiar with the Phare programme and had participated in two seminars in the previous year, one on curriculum development and teacher training, the other on the dissemination of the Green Paper. He formed part of the team who approved the new equipment specifications. But he is only one of very few.

Pedagogical institutes need to be involved in development projects along the lines of the Phare VET programme. The latter focused on 12 pilot schools, while the pedagogical institutes were totally forgotten. Subject inspectors did not receive any training; they see their subjects disappear in the long run or being parcelled out into new modules, and they are less and less able to perform their supposed roles when going to the schools. This is a serious barrier to broad acceptance of the reform. In fact, it may give ground to a large group of key actors who veto reforms.

The above three examples demonstrate a sub-optimal reform environment resulting from a lack of internal consistency of reform measures introduced.

4.4.2. Curriculum reform in Estonia (Nielsen, 1999b)

The EU-funded Phare 1995 VET Reform programme in Estonia aimed at assisting the country in kick-starting VET reform, which would be key to developing a workforce required for a modern market-driven economy. The specific activities to be undertaken embraced the following main components:
(a) curriculum development;
(b) teacher training;
(c) provision of equipment;
(d) establishment of links with partner schools in European Union countries;
(e) support to the design of a VET reform strategy and dissemination of results.

The curriculum design principle chosen in Estonia is a flexible, modular-based and employer-led model. In the development of curricula, Estonia follows a model similar to the Irish (FAS) one, but adapted to the Estonian context. 13 selected pilot schools developed modularised curricula to nationally approved industry standards. Social partners, in particular the Chamber of Commerce and Industry, are actively involved, at national level, in the definition and approval of occupational profiles and, at local level, through school-based curriculum design groups developing course modules to national standards.

The system logic of VET has changed. A methodology has been developed and successfully piloted to identify skill needs and educational goals and to translate occupational profiles into curricula. This has functioned well. Employers and their representatives have started to influence VET provision by articulating their skill requirements and requiring schools to produce skilled workers according to these needs. The dynamics of the VET system have been increased and self-regulating mechanisms built into the very core VET.

The results achieved show that the adopted strategy served as a catalyst for building the thrust and the momentum needed in the new approach to curriculum development. The basic institutional frameworks are in place. However, the curriculum support infrastructure is extremely fragile. Human resources in the VET Department of the Ministry of Education are severely limited. Some years ago the Ministry had already transferred its responsibilities for curriculum innovation to the national centre for examination and qualifications (NCEQ). The centre, currently engaged in other assignments such as handling national examinations, is methodologi-
cally and technically clearly too weak. It has not been in a position to develop sufficient capacities to absorb Phare programme results, such as endorsing the newly developed modules or driving the reform process forward. In 1999, about 130 modular descriptors that had been submitted to the centre were still awaiting their approval and accreditation. The centre, in its function as the national vocational qualification institution in the Anglo-Saxon philosophy, constitutes the basic cornerstone in the whole VET system logic. If the system is to work well, the capacities of the NCEQ to set up a national qualification structure and to act as the national award and accreditation body will have to be built up quickly. Staff needs to be recruited and trained.

This is a clear threat to sustaining VET reform results. The NCEQ needs to be given a clear mandate and to attract new experienced staff. In addition, it requires immediate technical support with a view to building up its capacity. Simply not endorsing the modular descriptors developed under the Phare programme, as was the case, puts both the momentum of the whole curriculum innovation process and the still fragile interaction between the employment and the education system seriously at risk.

It follows that there is a lack of external consistency, owing to problems of consistency between educational objectives on the one hand and the economic and social context on the other. Although attempts to increase the responsiveness of VET to the changing needs of the economy were a stated priority, there was only limited evidence of the implementation of this objective, and the educational infrastructure has continued to operate primarily according to its own logic.

4.4.3. **National dissemination of pilot project results in Romania** (Nielsen, 1999a)

The Phare VET Reform project in Romania selected 75 VET schools to act as pilot schools, some 10% of all schools. At the end of the reform programme the country was faced with the challenge of generalisation of results. The government then decided that the new curricula were obligatory in all schools, without any new equipment or any proper training. The national inspectorate, whose actors were themselves only half-familiar with the new curriculum principles, was supposed to support this vast implementation effort.

The decision to generalise the pilot school curricula indicated an intention to take the Phare VET reform methodologies from the experimental to the systemic level. However, it is difficult to see how this could be achieved without a considerable investment in equipment and training of both teachers and school managers, in short strengthening material as well as human resources. Pilot schools and teachers are not in a position to generate their own resources for this huge effort. The necessary critical mass of change agents may be in place to consolidate what has been achieved. The teachers and curriculum development officers trained under the Phare programme could have played a crucial role in training the staff of the majority of vocational schools not involved in the Phare programme. However, a broad nation-wide implementation effort requires careful planning and substantial resources.

In all types of pilot projects there is a risk of encapsulation. As a rule, the best institutions are selected and, during the piloting phase, they are granted good development conditions, i.e. they are freed from the difficulties that exist in the ordinary structure. When it comes to the transfer and broad systemic implementation of pilot results, all the real world problems and barriers suddenly turn up again. The risk of barriers in the ordinary structure is always a serious factor to be reckoned with when going from the phase of ‘VET reform (pilot) project’ to the phase of ‘national VET reform strategy’.

One other impeding factor in the model school approach is the fact that a number of elitist schools are given all the equipment, all the coaching, all the study tours, all the development assignments, etc., leaving the initially backward schools another couple of years behind. It is not clear how Romania (or any other country) will cope with this problem.

Another threat to the wider dissemination of curricular reforms, as initiated under the Phare programme, is the relatively low level of teacher qualifications. A major national in-service teacher training programme is required to make VET reforms happen at school level. Modularisation, as the basic pedagogical principle, implies a radical change of working conditions for Romanian vocational school teachers. In many ways, the teaching of modules could be seen as reduced qualification from the point of view of the professional teacher. So what is needed is not only technical up-skilling
but also awareness-promoting initiatives on a broad scale.

The reform of VET systems is more than a single act establishing a new legal framework. A change in the logic of the system can only be realised gradually by those who are involved in VET themselves, especially at the local level and in the schools. Since there are many people involved, a tremendous learning process is necessary, which needs to be facilitated and intensified through proper intervention and guidance.

After a long period of central regulation and bureaucratic administration, it is particularly important to depart from the traditional top-down conception of VET reforms and promote self-responsibility and self-initiative for local actors. But such bottom-up strategies, that give the initiative for educational change to the individual schools, may easily lead to fragmentation and competition between schools, as a consequence of the pilot-school based reform strategy.

Romania shares this challenge with other central and east European countries. It will be necessary to integrate the various changes introduced in a more coherent change strategy. From a strategic point of view, one should reflect once again on the interrelationships between the various aspects and issues and ask the fundamental question: where should one start in order to trigger a development process that will lead to the desired outcomes? In particular, it is important to explore whether the applied strategy of bottom-up curriculum-modernisation-through-pilot-schools, which has been forced upon the country by the EU Phare programme, has had the intended impact. And how, given the experience so far, is it possible to transcend from the level of piloting to the level of systemic reform.

Based on the fact that changes to one part of the system have an impact on other parts of the system, the challenge for us was to design a reform process that builds upon the internal links and develops both the framework structures and the actors at the same time. We highlighted the following system connections.

4.5. Example of stressing system connections in a VET policy evaluation for Croatia (7)

While the examples given in Section 4.4. indicate a lack of consistency in one form or another, a further example shall show how emphasis has been placed on system connections in a VET policy evaluation for Croatia. The original assignment given to the ETF was to analyse the training of VET teachers and trainers in Croatia and, on this basis, to come up with proposals for an EU-funded development programme.

Basic problems identified in Croatia include:

(a) the non-existence of a genuine and holistic pre-service VET teacher training system (would-be VET teachers read a technical subject plus general pedagogy; there are no in-company training practices);
(b) the in-school training year of would-be teachers is not well supervised;
(c) the central Institute for Educational Development offers, through a centrally published catalogue, a list of teacher training courses established centrally on the basis of perceived needs for teacher training and available capacities;
(d) in-service teacher and school headmaster training schemes are highly limited in general; training is neither linked to school development nor to teacher performance evaluation or career progression.

Based on the fact that changes to one part of the system have an impact on other parts of the system, the challenge for us was to design a reform process that builds upon the internal links and develops both the framework structures and the actors at the same time. We highlighted the following system connections.

4.5.1. Decentralisation and school heads’ and teachers’ training

4.5.1.1. Governance

School heads, administrators and teachers will have to share the burden for a number of new educational management issues. This implies a radical change, as schools must be managed in a

different way. Apart from pedagogical innovation, schools will have to establish a strategic planning and budget execution capacity. Both school headmasters and selected teachers will be involved at the strategic level of the new school organisation. This calls for training to develop the necessary skills.

4.5.1.2. Finance
Money will (partly) be transferred in the form of block grants, and schools will have to strengthen their capacity for planning and economising in the context of budgetary restrictions. They may also have to look for new income sources and will have to form local partnerships with parents, education authorities, employment services and employers. This will have immediate consequences for school headmasters, administrators and teachers alike: all of them will have to learn how to act as entrepreneurs.

4.5.1.3. Curriculum (content)
A modern VET system is based on a framework rather than detailed regulations, to give some freedom of action to the institutions providing the education and training services. Curriculum development and delivery will, therefore, be decentralised to some extent. At the central level, both clear objectives in terms of learning achievements (competences) and frameworks for content will be formulated, while schools will have to determine, to a higher degree, both the more specific content elements and how learning objectives are to be achieved (learning methods). As emphasised also in both the White Paper and the Ministry’s concept paper, modern learning theory calls for a shift from teaching to learning and from a teacher to a learner-centred focus. In this context, headmasters will play a key role in promoting new pedagogical approaches in their schools, and VET teachers and trainers have to be equipped with the necessary skills to master these new assignments.

We are aware that the needs for training in relation to the change processes are enormous, and that training should start as soon as possible. Without such training and school development processes being initiated on a massive scale, curriculum change and VET modernisation will not be accomplished. As this is also a never-ending process, resources have to be freed up in the budget for this task.

4.5.2. Curriculum reform and teacher training
A decentralised VET system ascribes new roles and responsibilities to all involved. New curriculum principles will involve all teachers at a vocational school. Curriculum-led VET reform has considerable implications, including the following.

4.5.2.1. Design and development of new curricula
Within given objectives and frameworks defined at national level, teachers will be actively involved in curriculum development: they will have to translate overall learning goals into specific subject matter, and determine the actual teaching plans and the methods of teaching/learning. In addition, selected teachers will also have greater responsibilities in other fields as outlined above.

4.5.2.2. New equipment
New curricula imply the use of new or updated learning (pedagogical) equipment. The modernisation of school workshops calls, first, for teachers and instructors to master the new machines themselves and, second, for the ability to make good pedagogical use of it, i.e. to create meaningful learning opportunities for students. An intensive training in vocational didactics is necessary. The (practice-related, partly in-company) programme for the training of teachers and trainers within the dual VET system, which was implemented by the Chamber of Crafts and Trades in collaboration with western partners, may serve as a good example in this respect. Although we are aware that it is impossible to send all teachers and trainers abroad for a certain period, learning on the spot from the best vocational schools or companies is probably the most effective way to learn about the new, often tacit knowledge and skills required.

4.5.2.3. Curriculum delivery
New student-activating methods must be introduced, such as problem-solving organised as group work, project work, etc. The focus is on training teachers in developing the capacity to organise new learning processes, to stimulate the learning motivation of students and to establish better learning environments. This would have to go hand-in-hand with the development of new teaching material.

4.5.2.4. Teams of teachers
The work of teachers and instructors will follow new principles of organisation. The need to
develop key occupational competences on the basis of a much better integration of general subjects, subjects of vocational theory and subjects of vocational practice requires a much closer cooperation between all teachers and trainers in a vocational school. We have witnessed an excellent illustration of this new approach in a private economics gymnasium, where various subject teachers work together to cluster a number of formerly separate, but related subjects into broader subject fields. The new approach presents a radical challenge vis-à-vis current school practices: the teacher becomes part of a collective planning body composed of several teacher colleagues who altogether feel responsible for the holistic (more broadly occupation-oriented rather than narrowly subject-oriented) competence development of an individual student. To achieve this goal, a lot of school-based teacher and trainer training and development will be required in the coming years, as an inherent feature of true curriculum reform. Again, the school headmaster has to understand these processes to be able to provide direction and leadership, orchestrating the efforts of the teacher teams.

4.5.2.5. Curriculum assessment and revision

In a decentralised VET system leaving substantial freeway to the schools, mechanisms will nevertheless have to be introduced to assure quality to the level of agreed national standards and a purposeful and efficient use of public funds. As a result of pressures from the financing authorities, but also a number of civil and economic actors to which the school becomes increasingly accountable (students, parents, employers, etc.), vocational schools will have to reflect on the range of their programmes and the quality of their services. The best schools in Europe have developed a genuine evaluation culture as a shared preoccupation by all school staff, which encourages feedback from their external clients with a view to improving learning processes and outcomes continuously. Such an evaluation culture can only be established on the basis of mutual trust and must reflect the values of the school, which can vary considerably from school to school. Once more, school headmasters would have to play a pivotal role in establishing such a culture.

As not all the necessary changes in the Croatian VET system could be fully implemented immediately, a careful choice had to be made of elements that are critical to reform. Our choice of priorities was based on the following criteria: instead of a series of separate interventions in the fields of pre-service and in-service VET teacher training, we recommended trying out a coherent, systemic approach based on partnerships between schools, universities, teacher training institutions and companies. Furthermore, our approach focused on the principles of learning on the job or action learning, meaning that learning should be as closely related to the changing day-to-day work processes as possible. This way, theoretical knowledge would be combined with experience-based learning as part of the change process and become applied knowledge or action competence (Handlungskompetenz). Rather than taking teachers out on external courses (as is the current practice in Croatia), learning outcomes are optimised when teachers and school headmasters learn through first-hand experience, being supervised by teacher-mentors in their original work environments. Having gone through such a process themselves, teachers are more likely to practice the same new style of learning with their students in day-to-day teaching practice afterwards. Moreover, involving all school staff in a collective learning process, rather than training school headmasters and teachers individually, seemed the most promising way forward towards school development. Headmasters and teachers need to share the same vision and see a role for themselves in a future reformed system. Also teacher trade unions needed to be taken aboard.

The activities which were finally recommended for inclusion in the EU-funded development programme were, amongst others:

(a) the development of change agent team strategy and the training of a sufficient number of change agents as a way to introduce new VET reform concepts and train teachers at school;

(b) the development of the mentor as one of the profiles of education specialists, the selection of excellent teachers specialising in certain fields and their training as mentors;

(c) the design of a training programme for school headmasters which should be owned and multiplied later on by the School for School Headmasters in Croatia;

(d) the design of new methods for analysing training needs, the pilot-testing of training courses delivered in VET schools and the implementation of a system to monitor the
relevance and effectiveness of training before, during and after the courses;
(e) the implementation of a number of school development projects as the principal approach to training;
(f) the identification of a central teacher training institution that would accompany the development process, and own and disseminate further all innovations introduced into the system.
Apart from the systemic approach, the evaluator needs to develop an understanding of how systems change in the interactions between actors, that innovations need to be taken forward not only to implementation but also to institutionalisation, and what are possible strategic levers for change.

5.1. The complexity of systemic reforms

To illustrate the complexity of systemic reforms in VET and the role of evaluations in them, let us first have a closer look at how they are defined. While, currently, there is no agreed definition of systemic reforms, Grootings (1993) describes them in the context of the following post-socialist stages of reform in transition countries (the first stage was added later by Bîrzea, 1997):

(a) corrective reforms that are initiated with immediate repairing objectives in countries emerging from a war or deep economic regression period;
(b) modernising reforms that are interventions aimed at reducing gaps and catching up with western institutions. They are especially active at the level of curricula, teaching and learning methods, examinations, and school textbooks;
(c) structural reforms that are targeted at the structures, legal framework and management of educational systems;
(d) systemic reforms that are deeper and have a global character because they call for a genuine change of paradigm in terms of educational policy. They are aimed not only at the curricula or the legislative framework, but at the internal logic of education and its relationships with the global social system. A systemic reform examines the key elements of every educational policy: the role of the state, relations with the labour market, the financing system, efficiency control, the normative role of national standards, etc.

A study by the US Department of Education (1995) offers the following definitions for systemic reforms in education:

(a) ‘systemic reform addresses all of the mutually reinforcing structures, processes, and activities within the educational system, recognising that altering any one part of the system necessarily impacts on all other parts;
(b) systemic reform requires system coherence through the integration of policy and practice;
(c) systemic reform constitutes a mainstream activity of all organisations involved, not an alternative or special programme;
(d) systemic reform requires strategies that help develop and mobilise the conceptions, skills, and motivation in the minds and hearts of scores of educators;
(e) systemic reform requires the development of routine mechanisms for bringing people together across roles, within and across organisations, for developing and maintaining shared direction and understanding; and to maintain strong communication among all of the constituent parts of the system;
(f) systemic reform in education addresses the preparation, continuing learning and working conditions of [...] educators in all roles – teachers, principals, counsellors, specialists, paraprofessionals, central office and higher education personnel.’

What is interesting to note in the above definitions is that:

(a) the professionalisation and further development, in a structured and continuous way, of teachers and other education specialists are considered key to the success of any systemic reform effort;
(b) higher education plays a central role in implementing reforms of the VET system. Much will therefore depend on new school-university partnerships.

In other words, changing the VET system – as any social system – is a complex, multi-layered process, comprising all system levels:

(a) the primary learning process;
(b) the organisational level;
(c) the institutional frame;
(d) the policy and legal frame.
5.2. The path dependency of institutional change

The organisational settings and arrangements, and related social values and behavioural patterns, have grown historically as the result of long, non-linear social development processes. According to Grootings (2002a), ‘institutions result from social relations among people and imply therefore power and trust. They are embedded in historical contexts and part of wider institutional frameworks in which different institutions are related to each other’. Taking the example of curriculum reforms, Parkes et al. (1999) illustrate their institutional embeddedness as follows:

‘Curriculum is perceived as a balanced relation between the potential and interests of individuals and the requirements of society. From the individual point of view “it” can be seen as the totality of measures, interactions and experiences within an organised learning process. However, individuals can make up their curricula only within patterns laid down by organised bodies and social institutions within society. These institutional frameworks are anchored in the specific social system of a society. They offer opportunities to the individual who in turn is constrained by requirements and standards to become a “responsible” member of society, in whatever way this is defined. The educational field has its own bodies and institutions to develop a framework to respond to the temporary and long-term needs of society, as well as the wishes of the individual to promote his/her personal development. … The main actions, actors and institutions involved in the development of a curriculum are part of “institution-building” involving developing “institutional homes” for key functions and activities.’

Grootings (2002b) argues that there is an institutional logic that is country-dependent, and reform through institutional change is determined by the institutional context of the country. They represent what was possible to achieve at a certain point in time rather than being a perfect making. They are unique for each country, which explains the uniqueness of reform designs for each country. Similarly, current OECD economic thinking (the development of social capital) suggests that reform, and economic development, is best undertaken through the mutation of historic institutional structures.

5.3. Accomplishing change

It is also important for the evaluator to understand the ways in which understanding and support by the actors necessary for implementation can be gained, as a lack of this has often led to failures of reform projects. Reforms will have to be carried out by existing staff, and despite widespread agreement with the global policy objectives and the recommended improved policies, there may still be great discomfort felt by the reform actors at the prospect of changing traditional ways. Accomplishing change is about reversing deeply embedded policies and strongly held beliefs. It is about preventing resistance to change or attenuating its impact.

The prevailing political culture will determine the accentuation of approaches to reform – with top-down and bottom-up reform approaches at the two ends of the spectrum. Again, there is no wrong or right in favour of the one or other approach. Our experience tells that contrasting organisational forms are required for interventions at the central level and the level of organisations, respectively. A high degree of bureaucracy is likely to make structural reform easier, while non-bureaucratic reforms are needed for curricular innovation and learning.

While, in transition countries, system change was traditionally designed by the centre and decreed from the top, the political culture is now changing. Local agents claim a stake in the design of reform concepts and, as we have seen, resistance may build up at an early stage when Governments, or foreign consultants, present ready-made concepts rather than involving local change agents in reform design from the very outset.

According to Paquet (2001), the defining characteristics of democratic change include bottom-up governance schemes and a new form of transversal coordination, while collaboration becomes the new categorical imperative. The centre will focus on norms and the periphery on delivery; for the two to function together, a clear sense of public purpose is needed, as well as new partnerships and skills. The agencies need to be granted the necessary powers to organise activities and become negotiating arenas, providing space for interaction, the (re-)definition of activities, as well as monitoring. A shift is
taking place in the policy management of public services: the stress is no longer on inputs but on outputs. The government defines objectives and sets frameworks while allowing a flexible style of decentralised management that is capable of rapidly adapting to new demands. This in turn calls for a decentralised structure, and for new forms of horizontal accountability, for the system of governance to be effective.

Such a process of cultural transformation and the translation of policy into practice is almost always an extremely lengthy process. McLeish argues in her introduction to *Process of transition in education systems* (McLeish and Phillips, 1998) that the completion of the transition process at the structural-legislative level in no way implies that educational transition at the micro level has been achieved. To change a label is easy, to effect a comprehensive change in practice is very difficult. Peter Drucker (1999), whose speciality, since the 1950s, has been to distil current thinking and research for the practical manager at the level of the organisation, makes the following remarks in his last book: ‘There is an enormous amount of work to be done in organisation theory and organisation. There are only organisations which have distinct strengths, limitations and specific applications. A given organisation fits certain tasks in certain conditions and at certain times [...] One cannot manage change; one can only be ahead of it [...] Change is necessarily a series of small incremental changes.’

We conclude that system change will have to build on the given historically grown institutional structures. It is likely to be achieved only through small, incremental change in narrow and targeted areas and only where there is equilibrium between radical change and traditional forces. Change requires, according to Paquet (2001, see above) a clear sense of public purpose, new partnerships and new skills, as well as careful policy coordination, compensatory mechanisms and collaboration in adequate forums for consultation and decision making.

5.4. The role of evaluations

Planning system change would be the art (rather than the science) of designing a multi-purpose, multi-actor, multi-stage process comprising coherent actions across many interdependent areas. This process must be based on a sound understanding of the institutional logic of the system and a number of hypotheses about the reasons for lack of internal and/or external consistency and how they could be overcome. Due to the multitude of factors impeding a system that is in a steady state of development, and the unpredictability of human behaviour, such planning models and their outcomes must always remain imprecise. OECD (2000a) argues that ‘the outcomes of educational processes are multi-dimensional, underlying factors remain imperfectly defined.’

Being aware of the imperfection of any planning model, it becomes essential to build into the change process, strategic phases where social actors come together and reflect and learn with the ultimate aim of verifying or rectifying hypotheses. This explains the need for evaluations not as one-off events, but as a continuous cycle.

Evaluations take on different roles in such a change process. They can be commissioned by the (national or international) funders of reform projects with a view to providing feedback on progress and establishing whether they have met their targets. In this case, evaluations would be summative (evaluation of results/impacts); they would serve as accountability tools and be largely limited to the projects in question. However, as already argued in the introductory chapter, project evaluations provide too limited a scope for assessing the complex nature of changes in VET within their social/ institutional contexts. That is why also ‘summative’ evaluations need to be complemented by formative elements. What formative literally means is having influence in forming or developing.

Evaluations can be organised as a conscious, critical reflection process at various stages of the reform process. They involve the reform actors, arrange for learning and negotiating opportunities and help shape or reshape policies. As such, they are part of the development process. They are constructivistic in that they deny that there is an objective knowledge of the world (the VET system) and in that they put actors at the centre of the evaluation rather than its outcomes.

Stern (2003) describes the following steps of such an evaluation:
Consensus creates one priority. A second need is capacity building by means of a dialogue, in simple and transparent terms, with the key actors so that it is possible for them to internalise and own the outcomes of evaluation. This will, subsequently, allow them to have the understanding and political will to move to strategic implementation both at macro and micro levels.

Inspiration and guidelines can be found in the so-called empowerment evaluation school in US university evaluation research. But this approach is also well known in Europe. For example, the UK Tavistock Institute of Human Relations (TIHR) makes use of an approach described as ‘development evaluation’ (Stern, 1989). It involves a number of elements:

(a) it offers participants in the setting being evaluated a voice in shaping the evaluation agenda;
(b) it involves an active feedback policy;
(c) it sets a clear focus on utilisation and implementation.

The approach is based on clusters of shared professional values, one of which is the commitment to action research.

Nyhan (1998) outlines the lack of a research tradition in VET (as opposed to general education). He reviews the approach of Kuhn (1996) in rooting rationality in particular political and cultural contexts, and Carr and Kemmis (1986) in accepting that the social actors involved in transformation need to be engaged in a kind of action research. In doing so, he points the Leonardo da Vinci Programme firmly in the direction of action research: ‘how can we change things at present to ensure a better future?’

Gisela Shaw (1999) further elaborates the Nyhan position: ‘This type of research (action research) is based on two fundamental assumptions. The first is that social research, including research in VET, is ill-advised, if it tries to copy the positivistic model of the natural sciences, relying primarily on quantitative and empirical methods and aiming at research results of universal validity allowing scientific predictions; instead, social research ought to accept that human actions and behaviour can only be interpreted and illuminated and that such interpretation and illumination requires taking into consideration the social context in which they occur (Hans-Georg Gadamer’s hermeneutic/interpretative approach; Gadamer, 1972). The second goes beyond this position by aligning itself with a claim embedded in Jürgen Habermas’ critical theory: human science being rooted in the emancipatory human knowledge interest (rather than in the technical or practical knowledge interest) cannot stop short at interpreting and illuminating actions and behaviour but ultimately aims to equip individuals with an understanding of how best to pursue their objectives in a rational way (Habermas, 1972). Although acknowledging the problem inherent in the notion of action research, i.e. a deliberate blurring of the boundary separating the researcher from the practitioner, Nyhan nevertheless highlights the fruitful implementation of action research in educational and training programmes in a number of countries [...]’.

The action research strategy offers channels for continuous dialogue with the relevant stakeholders (policy-makers, actors) about the validity of implementation of the concepts being studied, throughout the project: ‘An action-research project attempts to build concurrent dissemination activities into the research process’ (Nyhan, 1998, p. 28).

This means that the process of undertaking an evaluation – i.e. the way in which it is negotiated and managed and how it impacts on the subjects of an evaluation – is itself seen as a legitimate object of study. It follows that there is a thin line between evaluations, or action research, and quite specific consultant activity.

It follows also that the experience of the evaluators is key (as the OECD also suggests – see Section 2.1). Long apprenticeships seem to be necessary, whereby the junior evaluator would try to learn from the senior evaluator, thus going through the Dreyfuss & Dreyfuss five-stage process from beginner to expert.
6. Criteria, values and indicators underpinning VET policy evaluations

6.1. Evaluation criteria

In the introduction to our paper we have mentioned evaluation criteria, such as relevance, effectiveness, efficiency, impact and sustainability, which are important to project evaluations. It is clear that, for the purpose of policy evaluations, we would have to alter or redefine the set of evaluation criteria.

The OECD (Windham, 2000), for example, has used the following evaluation criteria for reviews of national education policies/systems in southeast European countries:
(a) appropriate political foundation and support for the reform process;
(b) equity in access, attainment and achievement;
(c) flexibility in planning and implementation;
(d) support for an interactive planning model involving cooperation among national, local, and institutional levels;
(e) incorporation of intersectoral and intrasectoral coordination;
(f) development of appropriate regional, national and international emphases;
(g) affordability;
(h) sustainability;
(i) efficiency;
(j) encouragement of supplementary resource mobilisation.

To evaluate reform policies meant to aid systemic change, as defined in Section 5.1, the following criteria may be useful:
(a) the political foundation and support for the reform process;
(b) the visions guiding the reforms;
(c) the leadership driving them;
(d) the structures and resources for planning and implementation, including the knowledge and research foundation upon which they are built;
(e) the opportunities for learning needed to sustain them;
(f) the mechanisms for communication used for intrasectoral and intersectoral coordination;
(g) the organisational arrangements designed to support them;

(h) the strategies used to implement them;
(i) the structures and mechanisms used to monitor and evaluate pilot implementation, and take corrective action, where necessary;
(j) the affordability/economic viability of interventions, as well as the structures and mechanisms used with a view to including innovative results in mainstream provision.

However, as illustrated above, a universal manual is difficult to assemble; a list would depend on the scope and purpose of the evaluation exercise. We suggest that, following the development evaluation approach, these evaluation criteria were subject to negotiations with the actors concerned prior to the start of the evaluation exercise.

6.2. Similarities and differences in VET reforms in western European countries

The same applies to the question of underlying themes and values. Evaluators often come with preconceived notions: VET policies in transition countries tend to be evaluated against western standards and practices. Can we, indeed, identify some common policy priorities (and indicators) which, sooner or later, lead to some convergence of systems and which could, hence, facilitate VET policy evaluations in transition countries?

Over the past decade, the VET systems of many west European countries have come under reform pressures as a result of persistent or new problems, such as:
(a) the academic drift;
(b) the poor motivation of students on VET programmes;
(c) the declining investment of time and money of employers in VET (even in countries with strong traditions of employer involvement, such as Denmark, Germany and Austria);
(d) global competitive pressures which are putting strains on the traditional social partnership approaches;
(e) vocational teacher training (polarised between German university studies obligatory for all teachers and the zero solution in the UK);
(f) the lack of parity of esteem between general education and VET;
(g) the failure to provide future employees with new types of knowledge and skills which they will need in a fast changing global economy;
(h) the problem of organising learning environments and learning processes which stimulate context-free learning of generic skills; etc.

Lasonen and Manning (2001) have shown, in an interesting typology, that west European countries have actually applied very different strategies to address these problems. They are summarised as follows.

6.2.1. Improving progression from initial VET to higher education
There are two distinct approaches to improving progression from initial VET to higher education; the access approach and the diversification approach. Denmark, Austria and Finland, on the one hand, have emphasised the diversification of higher education and the upgrading of existing vocational higher education by creating a new tier of vocational higher education institutions, where programmes have been extended in length. Many countries, such as Denmark, Finland and Norway have combined the diversification of higher education with creating opportunities for students on initial vocational courses to progress to higher education by obtaining dual qualifications.

England and Scotland, on the other hand, have concentrated on improving access to a single unified university system. However, a clear hierarchy remains between the old or traditional universities, which recruit almost entirely from students with academic qualifications, and the new universities which are more open to students with vocational qualifications. Local provision of further education colleges, often in partnership with universities has been expanded.

Most countries want the proportion of students in higher education to be expanded to at least 50 % of each cohort. In unified stratified systems, as found in England and Scotland, nearly all higher education institutions have equivalent status as universities. In divided systems, such as those emerging in Denmark, Austria and Finland, and already found in Germany, some higher education institutions are specifically designed to have a vocational role and not be universities.

The general conclusion is that the question of parity of esteem of VET needs to focus on post-compulsory education as a whole and not just the provision for 16-19 year olds. As the numbers opting for general education at upper secondary level continue to rise, the parity of esteem issue shifts towards higher education.

6.2.2. Improving links between vocational education and employers
Major differences between the countries include:
(a) those with a dual system of apprenticeship based on social partnership (Denmark, Germany and Austria) in which employers have a direct role in decision making about VET, which is based initially on their recruitment of apprentices;
(b) the state-led systems of the Nordic countries in which governments have actively involved employers in VET curriculum issues through consultative committees at local and national levels;
(c) the more voluntarist systems, such as England and Scotland and, to a lesser extent, Spain, where governments have attempted to involve employers through work-based training schemes, and through employers’ membership in governing bodies of schools or colleges and regional training councils.

Voluntarist systems rely on creating the conditions for new types of learning partnerships between employers and educational institutions. The problem is how to increase the involvement of small and medium enterprises with little spare staff capacity for supporting partnerships.

A variety of ways for ensuring greater employer involvement in VET is being explored. In Finland, France, Norway and Spain a greater regional emphasis is allowing individual institutions to develop closer links with local employers. Work experience inherent in curricula is seen, especially in Finland, France and Spain, as an important driver for establishing new links with employers and enhancing the future employability of young people.

6.2.3. Improving the status of VET teachers and trainers
Improving VET is intrinsically linked with the quality of teachers. Various patterns are emerging
in European countries. First, a much greater emphasis is being placed on improving the quality of vocational teachers rather than workplace instructors; only in countries with a dual system are serious steps being taken to upgrade the qualifications of trainers. In England and Scotland, the focus on workplace trainers is limited to their (students’ performance) assessment role. Second, general subject teachers in vocational schools are typically better qualified than vocational subject teachers (the dual system-based countries are, again, exceptions here). Most countries are trying to standardise qualifications and limit traditional regional differences (as in Spain).

Finland is the only country with a statutory in-service training requirement (five days a year). Most progress in terms of structural parity has been achieved in the three countries with a dual system (Denmark, Germany and Austria) and more recently in Norway, where the qualifications for VET and general subject teachers are equivalent and programmes are shared.

In England and Scotland, staff development and initial training of vocational teachers is left to individual institutions. Despite an increase in the numbers of qualified VET teachers, formal qualifications are still not a requirement for vocational subject teachers in colleges, though they do have to pass a test if they want to assess student performance rather than just teach certain types of vocational courses. Steps are being taken to establish national standards. One important step was not to introduce fees for the training of (would-be) vocational (or further) education teachers, placing them in an equal position with general subject teachers at primary and secondary level.

6.2.4. Improving the vocational curriculum
School-based systems, such as those in France and the Nordic countries, have increased the general education component of vocational courses, as well as the choices available to students. In Finland, this increase is expressed through the concept of integrated learning, whereby:
(a) schools can decide to increase the general education content from one fifth to one third;
(b) students can compose their personal study plans;
(c) all programmes include work placements;
(d) the vocational content has been strengthened by integrating it with activities from working life.

Mixed systems which are becoming more linked or unified, such as in England and Scotland, give priority to generic skills such as numeracy, literacy and team working, which employers claim are important, and to the mixing of general and vocational knowledge.

Systems with strong apprenticeship traditions, such as in Germany, try to enhance the vocational knowledge component of VET programmes through work process knowledge. Denmark is introducing greater flexibility into the dual system though the overarching concept of pathways. The Danish pathways approach stresses the support of student choices through the development of individual course plans within a modular curriculum and the need to complement vocational specialisation with broader-based studies. Vocational curricula in Spain seem to be developing against the European trend with a reduction in the amount of general education, shorter courses (more attractive to students and employers) and, in line with other European countries, more opportunities for local partnerships between schools and employers.

6.2.5. Summary
What the above analysis illustrates is that many of the west European countries have not tackled some of the fundamental problems in VET themselves; so there is no blueprint for a good model. EU experience has shown that neither the common factors pressing on national policy formation for education and training, nor the priorities that most EU Member States share, have led to a uniform pattern of convergence (even if some common trans-European trends have emerged). We are aware that the non-existence of clear-cut policy priorities or indicators may create a dilemma for the objective evaluator. A particular donor may impose them, but more easily on the evaluator than the specific government concerned. A possible way around the dilemma of preconceptions or nationally biased understandings is to start an evaluation by presenting the building blocks and formulating ideal practices against which existing practices would then be evaluated. Finally, the analysis shows that adopted strategies for VET system renewal or reform reflect the different educational
histories of each country rather than the problems these strategies are designed to address. They cannot be detached from their institutional contexts, as we have also argued in Section 5.2.

This is confirmed by Young (2001) who analyses why some countries have so far remained partly immune from the pressures to develop outcome-based qualification frameworks of the kind found in the United Kingdom and other anglophone countries. This is so despite the need to rethink the role of qualifications in light of global economic changes and the related changes in skill and knowledge demands (Lasonen, 1996) (Lasonen and Young, 1998) (Young, 2001). He shows that there are built-in contradictions with EU countries having quite divergent philosophies and practices, making in particular a distinction between an outcomes-based approach (English-speaking countries) and an institution or process-based approach (most notably those associated with the Germanic and Nordic traditions of education and training).

6.3. Potential and limits of the benchmarking approach

Despite the diversity of systems at EU level, attempts have recently been made by individual countries and at EU level to apply a benchmarking approach, first, to employment policy and subsequently to VET policy. What are the defining characteristics of this tool, what are initial results, and does benchmarking generally point the way forward in VET policy evaluations?

There are numerous definitions of benchmarking, the simplest being that benchmarking is the process of learning by making comparisons. Although simple, it captures some of its essential features: namely, that it is deliberate learning process and that it involves comparison. From the early days of benchmarking, companies have used it as a way to compare their own processes, practices and performance with other similar companies with a view to making improvements. More specifically, benchmarking in its classic form is a process of studying, adopting and adapting successful processes and practices from other organisations to improve processes, products, and services in one’s own organisation.

Since then benchmarking has become an umbrella concept used in modern management of both private and public organisations. The increasing use of benchmarking has also meant that some of its basic features have, in some cases, been dropped, which makes it difficult to distinguish it from comparison or comparative analysis.

To talk about benchmarking as a meaningful activity we need to view it as a formal and deliberate process of comparison, aiming at detecting weakness and creating ideas for improvement. Benchmarking is normally distinguished from purely analytical methods of comparison in that specific performance indicators must be developed and units must be found which can be considered as best performers concerning the chosen indicators.

The comparison then aims at finding options to make progress towards the position of the best performer (Grootings, 2000). Essential features include:
(a) the need for improvement;
(b) identification of a benchmark, i.e. agreeing on a benchmark that can serve as a reference point for undertaking the comparisons;
(c) the translation of the analysis into concrete steps for improvement.

For several years the Danish Ministry of Finance (1999) has undertaken benchmarking exercises for various sectors to assess where Denmark stands compared to other countries with respect to a number of issues that determine prosperity and welfare (8). A best practice approach was followed by choosing a fixed group of countries consisting of France, Germany, Great Britain, Japan, the Netherlands, Sweden and the US who are considered to be among the world leaders in one or more relevant fields. Major areas of prosperity and welfare were identified, and each area compared, using a number of indicators. A ranking between the two best or the two worst scoring countries for each indicator was judged as being above or below average. In the area of education the indicators below were used, resulting in a ranking for Denmark which is shown on the left.

(8) Similar evaluations have been done for Finland: Benchmarking Finland – an evaluation of Finland’s competitive strengths and weaknesses and Sweden: Quality comparisons in education and the employment offices.
As the Danish example illustrates, benchmarking is used solely to compare educational systems on the basis of their results, whereas a qualitative analysis of the practices, processes and methods that have led to certain results is not made. In the given example, no rationale was offered for why, for example, the reading skills of nine-year-old students are poor. Also, it is taken for granted that scoring high is good, while good reading skills are no guarantee that they could be readily applied to learning in everyday life. That is why the whole debate about which indicators to choose for a benchmarking process is so important and needs to be considerably advanced (9).

To us it is doubtful whether the Danish benchmarking exercise was more than just an international comparison of selected educational indicators. It should be mentioned that the Danish Ministry of Finance clearly recognises that softer educational objectives, which are valued highly in the Danish context, are not sufficiently covered by the benchmarking exercise (10). However, the ranking has been useful as a political exercise, as it triggered a national debate on education and

### Figure 2: Denmark’s ranking – education sector

<table>
<thead>
<tr>
<th>Above average</th>
<th>Average</th>
<th>Below average</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Number of hours of teaching per teacher in basic vocational education</td>
<td>- Total expenditure per pupil in the educational system</td>
<td>- Number of hours’ teaching per teacher in primary and secondary (general and vocational) education</td>
</tr>
<tr>
<td>- Pupil/teacher ratio in primary and secondary education</td>
<td>- Annual number of hours of teaching in primary and lower secondary school – 12-year-olds</td>
<td>- Annual number of hours’ teaching in primary and lower secondary schools – 14-year-olds</td>
</tr>
<tr>
<td>- Small proportion of young people who have been unemployed for at least one year</td>
<td>- Proportion of population with upper secondary education</td>
<td>- Natural science skills after 8 years’ schooling</td>
</tr>
<tr>
<td></td>
<td>- Proportion of population holding tertiary education qualification</td>
<td>- Reading skills of nine-year-olds</td>
</tr>
</tbody>
</table>

Source: Danish Ministry of Finance, 1999.

---

9) The discussion on identification of relevant benchmarks is not new and not only linked to educational issues. The quote below points to reservations and illustrates how sensitive to the whole benchmarking exercise the identification of relevant features and benchmarks is. ‘Companies often waste time benchmarking non-critical functions or struggling to raise the performance bar an insignificant amount. Does it really matter to your customers if you shrink your delivery time by two days? If you deliver packages, yes. If you deliver antiques, maybe not. And how much does it cost to cut out those days?’ (Dahle, 1996).

10) ‘In addition there are other important objectives. At primary and lower secondary school level one objective is to have pupils develop their creativity and self-confidence, as well as interpersonal and communication skills, and for the school to prepare pupils for life in a democratic society. These soft values need to be measured using new methods that are still in the developmental stage, though as yet there is still no sound body of data on which to make international comparisons. Indicators for the softer values are not included below, therefore’. Strukturovervågning – International benchmarking af Danmark, Finansministeriet May 1999, section 14.
helped identify areas for improvement or reorien-
ting the focus of reforms. The benchmarking
exercise was not used, therefore, to learn directly
from others, but to stimulate a national debate by
letting the figures speak for themselves.

The example points to the complexity of
applying benchmarking to education systems.
The approach presupposes that it would be
possible to disaggregate a VET system into
smaller units and that a benchmark could be
singled out that could then be adopted by the
institutional frameworks of other VET systems.
However, as we have seen before, the highly
context-linked nature of VET systems presents a
major barrier for the benchmarking approach. The
question is whether the benchmarking approach
can be applied to social systems at all.

In a study for the European Commission on the
possibility of benchmarking for employment poli-
cies (Schütz et al., 1998), a team of German
researchers makes a distinction between bench-
marking as an analytical inventory and as a policy
tool for improvement. The first mainly implies a
choice of benchmark areas for analysis and
policy recommendations; the second is of a
higher order and involves the understanding of
performance gaps and the implementation, moni-
toring and evaluation of interventions. Bench-
marking in the second sense has to be based on
benchmarking as an analytical inventory.

The authors argue that, in principle, bench-
marking for (employment) policies is possible, if a
number of conditions are met:
(a) a thorough understanding of the bench-
marking approach;
(b) a wide-ranging database and multi-method
approaches;
(c) adequate resources (time, money, personnel);
(d) commitment of involved key actors to organi-
sational learning.

They conclude that benchmarking is easier for
organisational entities (such as employment
offices) than various (employment) policies,
whose success depend in part on other policies
and institutions.

Grubb and Ryan (1999) see international or
system level benchmarking as valuable but
weakly experimental and consistent with fairly
weak research methods. Their comment is: ‘It is
more difficult to evaluate systems with their
multi-dimensional goals and multiple intertwined
characteristics, than it is to evaluate specific
public programmes with typically unidimensional
goals and simple attributes. It is also the case
that no social experiment involving apprentice-
ship has ever been conducted. Consequently,
while quasi-experimental evidence remains valu-
able, weakly experimental methods (notably inter-
national comparisons) become more important
for the evaluation of apprenticeship than for other
VET categories.’

The European Commission and the Member
States’ attempts to develop benchmarking proce-
dures, following the Amsterdam Treaty goal to
harmonise VET policies, are still in a very early
stage. In employment policy, where the process
is more advanced, the European Commission has
defined benchmarking as: ‘a formalised process
by which employment performances of different
countries are compared with each other in order
to highlight the best performing ones, to set
global targets for progress in the employment
situation over the medium/long term and to iden-
tify which policies have been most effective in
raising the level of performance’ (European
Commission, 1997).

The process implies:
(a) setting verifiable objectives and deadlines,
both at European and national level;
(b) agreeing on common performance and policy
indicators, based upon a country-comparable
statistical basis;
(c) undertaking peer reviews of major national
employment policies and programmes.

Experience with the benchmarking of VET
systems at EU level so far shows that the main
problem is not so much the general acceptance,
among Member States, of the approach, but the
development of appropriate indicators and
securing that data are available to measure these.
However, there are also words of warning raised
against believing that strategies of the best
performers can be easily copied under different
contexts. We have been aware of this danger
during our previous evaluation and policy advi-
sory work.
7. Conclusions and needs for further research

7.1. Conclusions

There is evidence of a shift in international review practice in VET from project-based to systemic approaches. This shift is related to the growing awareness of the restricted impact of single projects and their sometimes adverse effects on the system.

The aim of our study was to reach a better understanding of the underlying concepts and methods used for VET policy evaluations by exploring the contribution of research, as well as a number of examples from transition countries.

One of the central conclusions is that there is no holy grail in terms of conceptualisation or methodology related to VET policy evaluations. The engineer’s toolbox is of limited use. Likewise, management approaches based on refined analytical frameworks often tend to obscure rather than illuminate VET, including reliance on tools derived from systems analysis. It appears that the only remedy is the evaluator’s broad understanding of the essential components of VET, of their relationships, of the fundamental logic between the system and its environment and of strategic levers for change. This understanding develops only through many years of apprenticeship and first-hand experience of VET policy evaluations.

Using the hermeneutic approach, a particular system can be understood by alternating between studying parts (system elements) and totalities (the total system), as well as creating continuous interactions, in our interpretation work, between our preconception (Vor-Verstehen) and the new understandings worked out through the process. However, understanding always takes place between two entities both placed in a historical and cultural context, and Gadamer (1972) insists that the interpreter cannot abstract himself from personal history and culture. What the evaluator must do is openly recognise and clearly articulate this tension.

According to Habermas (1972), human science cannot stop short at interpreting and illuminating human actions and behaviour but, ultimately, aims to equip individuals with an understanding of how best to pursue their objectives in a rational way.

We explored the question of whether we can provide a conceptual framework for grasping not only the parts (building blocks) of VET, but also its historic roots and dynamic relationships and, thus, for managing the dialogue with key actors from inside and outside the system. A number of essential functions and processes of a successful VET system and architectural elements were described which are, initially, sufficiently simple to engage in the dialogue. However, for developing a better understanding of VET, the building-blocks approach has to be accompanied by a deeper insight into the systemic logic of the VET system.

A structural-historical approach (Strukturell-genetischer Zugang) could be used. Here, the analysis would start with an overview of existing practices, functions and structures, as they appear to the (experienced) evaluator. Essential structures are then traced back to their origin (genesis), and an understanding/explanation of the historical context is reached in which the specific phenomenon was born or established.

The phenomenological/hermeneutic method is primarily used to achieve an understanding rather than for explanation, which is to state causes behind the phenomenon observed. There is a continuous interaction between understanding and explanation; they feed each other. Understanding is the point of departure for VET policy evaluations, and explanations are needed when we fail to understand. Useful explanatory frameworks include:

(a) economic or labour market or sociological laws and other forms of determination of the concrete conditions under which a phenomenon is active;
(b) functionalist explanations referring to the totality in which the phenomenon to be explained is placed and a description of the precise function it serves;
(c) historical-cultural explanations behind the occurrence of different institutions and practices in contemporary VET systems;
(d) structural(ist) explanations of phenomena which can only be uncovered through an in depth analysis;
(e) system-analytical explanations unfolding how changes in one component will have spill over effects on other system components or on the total system (including discussions on internal and external consistency).

We conclude that, besides developing further the building-blocks approach by including explanations, and in order to avoid hobbyhorse explanations, we have to cultivate, individually and in a community of VET evaluation practitioners, a new attitude towards argumentation, a more open discussion practice. This we term a discourse.

VET policy evaluation requires a systemic understanding of VET. But how can the VET system universe be delineated? What relationships exist between the units? Which matrix of dominance patterns the interplay of units? What are the boundaries of the system? What relationship (metabolism) exists between the system and its environment? And what is the prime mover or strategic lever for change?

VET as a subsystem can be seen as an open system in a steady state; it depends on self-regulating mechanisms to maintain its boundaries and its continued existence within these boundaries (meaning primarily boundaries which distinguish the VET subsystem from other subsystems with the same society). The steady state depends on a balance of inputs and outputs. The inputs are demands and support: support makes the system strong enough to process demands and to produce outputs in the form of qualifications. The VET system sits in an environment – the total social system – and there is a continuous feedback into the VET subsystem, signalling if outputs produce good or adverse effects in the environment.

A systemic approach to VET evaluations focuses on the analysis of relationships, of communication channels, of responsiveness and adaptability, based on the fundamental understanding that changes in one component lead to changes in other components and in the system as a whole. So, it is important to understand the specific systemic logics of different VET systems. Hence, the interest in internal and external consistency, the preoccupation with systemic regulatory frameworks, and the search for the prime mover or the strategic lever related to VET system change.

Internal consistency indicates that improving VET (including increasing their attractiveness for both students and enterprises) can only be done through a system approach, encompassing all the components of the education system. The policy of a high degree of decentralisation and privatisation, and the modularisation of curricula, were discussed as two examples bearing the risk of lacking internal consistency.

External consistency refers to the relationships between educational objectives and the economic and social context. There is a range of potential problems connected to that: the employment system is not always able to articulate its needs; there are tensions between the pedagogical logic of school systems and competence needs of the employment system. Furthermore, the drive to develop national qualification systems as a means to increase internal and individual learner consistency may conflict with deregulated labour markets and short-term recruitment and training practices in companies.

Finally, it is realised that education cannot by itself solve employment problems.

To balance system-inherent tensions, there is a need for some form of regulatory framework that would allow flexibility and adaptability while satisfying the need for high quality and social equality. Such a regulatory framework needs to guide reform efforts.

A systemic approach to VET evaluation, as described above, is the way forward. The ultimate aim of all VET reform endeavours is to better balance tensions and ensure internal and external consistency of the system. But there is no uniform model that could be applied.

Apart from the systemic approach, the evaluator needs to develop an understanding of how systems change in the interactions between actors, that innovations need to be taken forward not only to implementation but also to institutionalisation, and of the possible strategic levers for change.

Our experience with VET reforms in transition countries so far, as well as definitions of the term systemic reform, demonstrate that two aspects are key to the success of any systemic reform effort in VET:
(a) the professionalisation and further development, in a structured and continuous way, of teaching and other education staff;
(b) new school-university and local partnerships.
Change is a complex, multi-purpose, multi-layered, multi-actor, multi-stage process whose outcomes are partly unpredictable. System levels affected by each change comprise the primary learning processes, the organisational level, the institutional frame, and the policy and legal frame. According to Grootings (2002b), reform is constrained by the institutional context of the country, and there is an institutional logic that is country-dependent. Reform can only be undertaken through the mutation of historic institutional structures.

Another aspect that the evaluator must understand concerns suitable ways to gain understanding and support from the actors necessary for implementation. Accomplishing change is about reversing deeply embedded policies and strongly held beliefs. It is about preventing resistance to change or attenuating its impact.

System change is likely to be achieved through small, incremental change in narrow and targeted areas and only where there is equilibrium between radical change and traditional forces. Change requires, in a new political culture, a clear sense of public purpose, new partnerships and new skills, and careful policy coordination, compensatory mechanisms and collaboration in adequate forums for consultation and joint decision. The prevailing political culture will also determine whether reforms are implemented through top-down or bottom-up approaches (or a mixture thereof). Our experience has shown that contrasting organisational forms are required, i.e. a high degree of bureaucracy for structural reform and rather non-bureaucratic reforms for curricular innovation and the learning process.

Owing to the multitude of factors impeding on a system that is in a steady state of development and the unpredictability in particular of human behaviour, planning system change must always remain imprecise and its outcomes imperfectly defined. As a consequence, it becomes essential consciously to build into the change process strategic phases where social actors come together and reflect and learn with the ultimate aim of verifying or rectifying hypotheses as input into the next stage of the reform. This explains the need for evaluations not as one-off events, but as a continuous cycle.

Such evaluations are formative in nature, as they have influence in forming or developing policies. They are constructivistic in that they deny that there is an objective knowledge about the world (the VET system) and in that they put the actors at the centre of the evaluation rather than its outcomes. Nyhan (1998) outlines the lack of a research tradition in VET (as opposed to general education) and highlights the fruitful implementation of action research in education and training, which engages the social actors involved in transformation. The action research strategy (here used in its empowerment evaluation version) offers participants a voice in shaping the evaluation agenda, it involves an active feedback policy, and it sets a clear focus on utilisation and implementation. Action research offers channels for continuous dialogue, with the relevant stakeholders, about the validity of implementation of the concepts being studied, throughout the project.

This means that the process of undertaking an evaluation – i.e. the way in which it is negotiated and managed and how it impacts on its subjects – is itself seen as a legitimate object of study. It also follows that there is a thin line between evaluation and consulting.

Evaluations call for substantial resources in terms of staff, time and money. It is important to remember that, with continually evolving reform efforts, little time spent in the country and few cases explored, evaluations represent only a snapshot and general conclusions must always be considered tentative.

Policy evaluations require different evaluation criteria from project evaluation. These depend on the scope and purpose of the evaluation and should also be subject to negotiation with the actors involved.

There are no absolute standards for a good VET system, which the evaluator could use. Concepts and practices of VET systems of west European countries all have their positive and negative aspects; they cannot be detached from their institutional contexts. VET systems of west European countries have come under reform pressures themselves, as many fundamental problems have not been tackled, and new challenges have emerged. The adopted strategies used by west European countries for VET system renewal or reform reflect the different educational histories of each country rather than the problems these strategies are designed to address.
In these circumstances, a universal manual is difficult to assemble. EU experience has shown that neither the common factors pressing on national policy formation for education and training, nor the priorities that most Member States share, are necessarily leading to a uniform pattern of convergence (even if some common trans-European trends have emerged). The non-existence of clear-cut policy priorities or indicators may create a dilemma for the objective evaluator. A particular donor may impose them, but more easily on the evaluator than the specific government concerned. A possible way around the dilemma of preconceptions or nationally biased understandings is to start an evaluation by presenting the building blocks and formulating ideal practices against which actual practices would then be described and evaluated.

Benchmarking becomes a meaningful activity (or tool), when we view it as a formal and deliberate process of comparison aimed at detecting weaknesses and creating ideas for improvement. Benchmarking is distinguished from purely analytical methods of comparison in that specific performance indicators must be developed and units must be found which can be considered as best performers concerning the chosen indicators. The comparison then aims to find options to make progress towards the position of the best performer. However, we have to take seriously the warning that strategies of the best performers cannot easily be copied under different contexts.

A team of German researchers (Schütz et al., 1998) makes a distinction between benchmarking as an analytical inventory and as a policy tool for improvement. The first mainly implies a choice of benchmark areas for analysis and policy recommendations; the second is of a higher order and involves the understanding of performance gaps and the implementation, monitoring and evaluation of interventions. Benchmarking in the second sense has to be based on benchmarking as an analytical inventory. The latter is especially relevant for our purposes.

Essential conditions for using policy benchmarking include a thorough understanding of the benchmarking approach; a wide-ranging database and multi-method approaches; adequate resources (time, money, personnel); and commitment of involved key actors to organisational learning. They conclude that benchmarking is easier for organisational entities than for policies, whose success depends on other policies and institutions.

Member States have accepted the approach, but there are problems in developing appropriate indicators and secure that data are available to measure these. There are also words of warning against believing that strategies of the best performers can be easily copied under different contexts. We have been aware of this danger during our previous evaluation and policy advisory work.

7.2. Needs for further research

What is presented in this study are the first results of what has turned out to be rather ground-clearing work in VET policy evaluation. Thus, we see our study as a basis for discussion and further research work. Further analysis is needed of the levers and critical factors for achieving genuine change, not only on paper but in the interactions between actors, which implies taking innovations to institutionalisation.

There is a need to develop both formative evaluation procedures and instruments to assist specific systemic reform initiatives in evaluating and monitoring progress towards goal achievement, and to inform future planning. An evaluation model must take into account the creation of new organisational arrangements attempting to bring about simultaneous renewal in more than one institution. Developing guidelines for long-term documentation and evaluation, which define needed data sources and indicators of development and achievement, would serve the continuing planning and development of such ventures and assist in developing comparable data to enhance our understanding of the limits and potential of specific reform efforts, not least the cost-effectiveness of various strategies. Further research is also needed to examine the real costs of education (or VET) reforms.

Given the importance of leadership in the success of these endeavours, additional research is needed to understand how leadership skills develop, and the strategies for developing them. When a study finds leadership to be a critical variable, it presents a dilemma for policy-makers.
Can such leaders be selected and/or trained, or do we have to wait for them to emerge? More research on the specific qualities of educational entrepreneurship is needed, particularly with a focus on how individuals can be trained to lead collaborative efforts.

Finally, more research is required into how benchmarking could be used to improve systems, which indicators should be selected for benchmarking VET systems on an EU or international scale, and what can be considered as best performers concerning chosen indicators.
**List of abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETF</td>
<td>European Training Foundation</td>
</tr>
<tr>
<td>NCEQ</td>
<td>Centre for examination and qualifications</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Cooperation and Development.</td>
</tr>
<tr>
<td>PAS</td>
<td>Pedagogical Academy in Sarajevo</td>
</tr>
<tr>
<td>TIHR</td>
<td>Tavistock Institute of Human Relations</td>
</tr>
<tr>
<td>VET</td>
<td>Vocational and educational training</td>
</tr>
</tbody>
</table>
Bibliography and references


Durand-Drouhin, M; Bertrand, O. *Improving the responsiveness of vocational education and training: flexibility and/or consistency?* Paris: OECD, 1994 [Unpublished draft].


European Commission / SCR (The joint service for the management of Community aid to non-member countries). *Guidelines for: drawing up terms of reference for evaluations, evaluation methodology, criteria and suggested layout for evaluation


Grootings, P. VET reform as institutional change: looking at VET reform through the eyes of New Institutional Economics. Turin: ETF – European Training Foundation, 2002b [Unpublished draft].


OECD. Development cooperation: efforts and policies of the members of the development assistance committee. Paris: OECD, 1997.


