

The value of learning

Evaluation and impact of education and training

Third report on vocational
training research in Europe
Executive summary



The value of learning

Evaluation and impact of education and training

Third report on vocational
training research in Europe:
executive summary

Pascaline Descy
Manfred Tessaring

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server (<http://europa.eu.int>).

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

Luxembourg: Office for Official Publications of the European Communities, 2005

ISBN 92-896-0342-9

© European Centre for the Development of Vocational Training, 2005
Reproduction is authorised provided the source is acknowledged.

*Designed by Colibri Ltd. – Greece
Printed in Belgium*

The **European Centre for the Development of Vocational Training** (Cedefop) is the European Union's reference centre for vocational education and training.

We provide information on and analyses of vocational education and training systems, policies, research and practice.

Cedefop was established in 1975 by Council Regulation (EEC) No 337/75.

Europe 123
GR-570 01 Thessaloniki (Pylea)

Postal address:
PO Box 22427
GR-551 02 Thessaloniki

Tel. (30) 23 10 49 01 11
Fax (30) 23 10 49 00 20
E-mail: info@cedefop.eu.int
Homepage: www.cedefop.eu.int
Interactive website: www.trainingvillage.gr

Edited by: Cedefop
Pascaline Descy, Manfred Tessaring
Project managers

Published under the responsibility of:
Johan van Rens, *Director*
Stavros Stavrou, *Deputy Director*

Table of contents

Country abbreviations	3
Introduction	5
1. Understanding evaluation	7
2. Approaches and methods of evaluation and impact research	13
2.1. Programme evaluation	13
2.2. Impact research	15
2.2.1. Education, training and economic growth	16
2.2.2. Social benefits and social infrastructure	18
2.2.3. Impact on company performance	18
2.2.4. Individual benefits and life-course research	19
3. Evaluation of education and training in a changing European context	21
3.1. Learning by comparison	25
3.2. Examples of evaluation of VET systems reform: the Dutch and Danish cases	28
3.3. Combating labour-market exclusion: does training work?	30
4. Impact and benefits of education and training	33
4.1. Determinants of economic growth	33
4.2. Non-material benefits and externalities of education and training	35
4.3. The significance of education and training for company performance	38
4.4. Individual benefits of education and training	39
Bibliographical references	43

List of tables, figures and boxes

Tables

1. Stage of VET programmes and their evaluation	8
2. Comparison of different evaluation methods	14
3. Private and social rates of return (RoR) on education in selected OECD countries by gender 1999-2000 (in percentage points)	40
4. Willingness to pay for education and training, by purpose, population aged 15 and more, EU-15, 2003 (%)	41

Figures

1. Benefits of education and training	17
2. Participation in continuing education and training of the population aged 25-64 years by level of educational attainment, 2003 (%)	22
3. Participation in continuing vocational training courses, EU-15 and new Member States, 1993 and 1999 (% of employees, all enterprises)	22
4. Public expenditures on ALMPs, EU-15, 1999 (% of GDP)	32
5. Effects of education on social cohesion in four countries	36

Boxes

1. Reporting series on VET research	6
2. Lack of internal and external consistency while implementing reforms of VET in selected Phare countries	24
3. The open method of coordination	26
4. EU benchmarks for education and training	27

Country abbreviations

A	Austria
AU	Australia
B	Belgium
CY	Cyprus
CZ	Czech Republic
D	Germany
DK	Denmark
E	Spain
EE	Estonia
EL	Greece
F	France
FIN	Finland
HU	Hungary
I	Italy
IRL	Ireland
L	Luxembourg
LT	Lithuania
LV	Latvia
MT	Malta
NL	Netherlands
NO	Norway
P	Portugal
PL	Poland
S	Sweden
SI	Slovenia
SK	Slovakia
UK	United Kingdom

Introduction

This publication presents selected findings from *The value of learning*, the third in the series of reports on vocational education and training (VET) research, discussing evaluation and impact research on education and training ⁽¹⁾.

The value of learning has been elaborated from contributions by researchers from various fields of education and training research and evaluation, published in three *Background reports* ⁽²⁾. These are complemented by additional research to review the various approaches, methods and results of evaluation and impact research. The report also indicates the implications for policy and discusses best evaluation practices.

Evaluation is spreading across Europe. It is used by governments and stakeholders as a tool for monitoring the implementation and maximising the allocation and use of resources in various social policies, including education and training and active labour-market policies. The European Commission is also evaluating its programmes and requires from the Member States follow-up and outcome evaluation for structural and social funds.

Evaluation is a systematic investigation to determine the significance, worth or benefits of a policy, programme or measure, using relevant social research methods, criteria, standards and indicators. It is also a developmental process that enlightens specific policies, processes and practices for its stakeholders. It contributes to collective learning and to knowledge production. It reduces uncertainties in decision-making, helps to improve design and implementation of social interventions, while ensuring effective allocation of resources. A characteristic of evaluation, compared with other fields of social research, is its direct links to policy- and decision-making.

The conclusions from 'impact research' serve to understand better the relationship between skills and socioeconomic developments and well-being. It investigates the relationship between human capital and various benefits for the individual, the enterprise, the economy or society, irrespective of whether they are the result of specific programmes or interventions (e.g. the impact of the general rise of skills and qualifications on economic and social development).

⁽¹⁾ Descy and Tessaring, 2005.

⁽²⁾ See Descy and Tessaring, 2004, three volumes, edited separately by Cedefop.

Evaluation and impact research have different objectives and partly use different approaches and methods. To discuss these various facets in detail, present case studies and best practices, review the methods used as well as the findings of evaluation and impact research, *The value of learning* is divided into four main parts:

- (a) understanding evaluation;
- (b) approaches and methods of evaluation and impact research;
- (c) evaluation of education and training in a changing European context;
- (d) impact and benefits of education and training.

Selected findings are presented in this summary.

Box 1. Reporting series on VET research

Cedefop has published reports on vocational education and training research since 1998 ^(a). They provide a comprehensive review of current research in initial and continuing VET in Europe, its results and the implications for policy, practice and future research. Attention is also paid to theoretical and methodological foundations and due reference given to economic, sociological, pedagogical and other fields of social research.

By publishing this reporting series, Cedefop improves the transparency of VET research in Europe through fostering cooperation and communication both within the research community itself and between researchers, policy-makers and practitioners.

Each research report is composed of two main publications:

- a background report, containing contributions of renowned researchers in different fields of VET research on selected topics;
- a synthesis report, presenting the main VET research issues, based on the contributions to the background report as well as own research, and drawing conclusions for policy, research and practice.

Both publications are accompanied by an executive summary (since 2001) for a larger public distribution.

Detailed information on the Research reports, including the contributions to the background reports, can be found in Cedefop's European training village www.trainingvillage.gr/etv/Projects_Networks/ResearchLab/

(a) Tessaring, 1998, 1999; Descy and Tessaring, 2001 a and b.

1. Understanding evaluation

Growth in evaluation activities results from both structural and managerial issues. In a context of malfunctions and polarisation of labour markets, in terms of career and earnings and of varying access to education, jobs and technologies, the need for efficient measures to integrate excluded groups has increased. These structural socioeconomic difficulties require multiple goal programmes and policies as well as complex organisational arrangements. At the same time, budgetary pressures have increased the need for improved performance, greater effectiveness and higher quality of public and private programmes and measures. International comparisons have also shed new light on education and training systems and on the preconditions and processes that affect their performance. Finally, the distribution of the European Structural funds, including mandatory evaluation, has also contributed to the development of an evaluation culture in a number of countries.

Because it uses various kinds of data and takes into account benefits, alternatives and consequences, evaluation has become not only a management tool to ensure efficient implementation of programmes and to measure their results but also an important instrument to generate empirical knowledge (Stockmann, 2000: 11) and to transfer results into policy and practice.

A certain logic model underlies the conception of any VET programme, describing the connections between the programme inputs, activities and processes (implementation), outputs, immediate outcomes and long-term impacts. The four stages of VET programmes proposed by Grubb and Ryan (1999: 13-19) ⁽³⁾ exemplify the logical model for VET interventions and programmes (Table 1).

At each of the stages described in Table 1, various forms of evaluation can be used, which relate two main types.

- (a) 'Formative' evaluation collects evidence to redirect and improve interventions in the course of their execution (Plewis and Preston, 2001: 10). Therefore, it is mostly used in the planning, design and

⁽³⁾ Grubb and Ryan (1999) use the term 'human capital development programmes'.

Table 1. **Stage of VET programmes and their evaluation**

Stage of VET programme	Objective and logic of action	Evaluation issues
Stage 1 Implementation	Formulation of the programme objectives and of its characteristics <ul style="list-style-type: none"> • resource allocation (administration, teachers, content, budget, participation, etc.); • legislative backup. 	<i>Ex ante</i> : achieve a consensus among stakeholders on programme objectives and characteristics. <i>Ex post</i> : determine whether a programme was established as intended and whether the allocation of resources was appropriate to the objectives assigned.
Stage 2 Learning process	Increase skills or competences of the participants <ul style="list-style-type: none"> • organise an appropriate teaching and learning process. 	Process: explore what kind of learning takes place, the appropriateness of the teaching methods, the relationship between learning and future outcomes, help the programme actors to achieve the objectives, etc.
Stage 3 Changing human behaviour	The skills acquired will change the behaviour of individuals on the labour market and in jobs.	Process/ <i>Ex post</i> : determine whether the programme has, e.g. laid the ground for more effective work or job-seeking by the unemployed or for an increase in productivity.
Stage 4 Creating long-term employment and non-employment outcomes	The changes in behaviour should result in economic and non-economic benefits for the individual, the enterprise or the organisation, the economy or/and the society.	<i>Ex post/impact</i> : <ul style="list-style-type: none"> • investigate the sustainability of programme effects, e.g. in terms of earnings increase and better job prospects; • take into account measurable outcomes as well as external effects and non-economic benefits (e.g. reduction of criminality, better health).

Source: Authors, based on Grubb and Ryan, 1999: 13-19.

implementation phase of a programme and uses mainly qualitative methods. The results of formative evaluation allow for improving procedures, learning to do things better, overcoming resistance, etc. (Stern, 2004).

- (b) ‘Summative’ evaluation uses mostly quantitative methods to judge the outcomes, effects and impact of an intervention. It is usually carried out *ex post*, after a programme has been completed. By providing respective feedback loops for the future of the intervention or for

follow-up projects, summative evaluations also have a formative character. According to Plewis and Preston (2001: 10) summative evaluations highlight different policy options, e.g. whether the intervention should be continued or not, whether it should be replaced by something different or better, and whether its success merits extension to a wider population.

Scriven (1991) illustrates these two types as follows: ‘when the cook tastes the soup, that’s formative evaluation; when the guest tastes it, that’s summative evaluation.’ Evaluations should include both formative and summative aspects depending on their adequacy for the various stages.

These types of evaluations reflect partly different basic philosophies. According to positivist thinking, the most traditional school of evaluation, objective knowledge exists. It can be obtained through observation and causal relations can be deduced. Adherents to the positivist view of the world are more likely to carry out *ex post* summative evaluation, trying to explain what works and what does not, i.e. to measure the outcomes, the benefits and impacts of policies and interventions. They tend to rely mostly on quantitative methods. ‘Scientific realists’ continue positivist thinking but are interested in opening the ‘black box’, i.e. understanding underlying mechanisms, processes and contexts which make programmes work or not. They tend to carry out *ex post* and implementation evaluation using both quantitative and qualitative methods of investigation to penetrate beneath the surface of observable inputs and outputs of a programme. ‘Action-theory’ evaluators are interested in the political rather than in the scientific use of evaluation. They take an interventionist role to help those participating in a programme or project to develop it according to their needs. Finally, ‘constructivists’ believe that reality is socially constructed. Constructivist evaluations involve all stakeholders in a development process, confronting their views to achieve a consensus, feedback to all, clarification of areas of agreement and disagreement in order to design interventions and policies.

To understand evaluation it is also useful to consider the level of evaluation activity in the various European countries and the factors that led to its development. Following Furubo and Sandahl (2002), European countries can be grouped in clusters according to their history and regular use of evaluation:

- (a) first wave countries (Germany, Sweden, the UK and the US) introduced and developed evaluation between 1960 and 1975;
- (b) second wave countries (Denmark, France, the Netherlands, Norway and Switzerland) developed evaluation as a governance practice in 1975 to 1990;

- (c) third wave countries introduced evaluation from the 1990s onwards, for example in response to external pressure from supranational organisations making it a condition for receiving financial support.

These clusters partly reflect the evaluation cultures as well as the maturity and recognition of evaluation activities in different countries. There are countries in Europe where evaluations are widespread and accepted, where they happen as a regular feature of administrative practices. The community of practice of evaluators is broad in these countries, with courses for evaluators offered at university level and special chairs devoted to evaluation. In contrast, in other European countries evaluations are carried out incidentally and only because they are required to receive external funding. Historical factors and the overall political and administrative culture determine the degree of institutionalisation of evaluation at national level.

Evaluators are subject to pressures. First there are time pressures because results need to be delivered quickly for decision-making. Second, political pressures arise because commissioners are often stakeholders of the domain being evaluated and might be tempted to influence results. Third are financial pressures as evaluation budgets are often only a small part of an intervention's budget. Despite these pressures, evaluators are expected to be as rigorous as social science researchers, while achieving results that are meaningful and can be of direct use for policy-makers. As a result, evaluators need to formulate standards to establish a reference point for their practice. Standards thus serve many purposes: as a framework for judging evaluations; as a protection that evaluators can evoke when being put under pressure; as a way of establishing a shared understanding between evaluators and clients as to what an evaluation is, and how it should be conducted; and as a tool for critical self-reflection.

Four clusters of standards formulated by the American joint committee on standards for educational evaluation have been applied in Europe as well:

- (a) utility standards; the evaluation should deliver the information that the client needs, and at the stipulated time and in an appropriate form;
- (b) feasibility standards; the scope of the evaluation should have a realistic relationship to factual needs and economic aspects, and be conducted in a diplomatic and prudent way;
- (c) propriety standards; the evaluation should respect the personal rights of the people who are involved in the evaluation as well as those affected by the results;
- (d) accuracy standards; the evaluation should produce valid and adequate findings by means of methodologically correct procedures (Beywl and Speer, 2004).

Adhering rigorously to all four clusters of standards is often not possible, and it is up to the evaluator to strike an appropriate balance between the intrinsic demands of social research methods and the extrinsic demands coming from the environment and the users of the evaluation.

2. Approaches and methods of evaluation and impact research

2.1. Programme evaluation

'Programmes' are measures and interventions of a limited duration, aiming to solve specific problems, and targeted at specific groups. Programmes can be implemented at local, regional, national or international level. Typical examples are reintegration measures developed in the framework of active labour-market policies to combat unemployment and other forms of exclusion. Qualitative and quantitative methods are used to investigate the implementation and results of programmes.

Qualitative methods deepen the understanding of process and outcomes of a programme. Though they can be used on their own, qualitative evaluation methods often usefully complement quantitative methods. While the latter are primarily used to generalise results and make them comparable across studies by using standardised variables and common dimensions, the former allow understanding of the mechanisms that lead to the observed programme outcomes. They also open the door to understanding the participants' experience in the programme. In short, they capture the non-measurable effects. To do so, qualitative approaches mainly use small rich-information samples, sometimes even single cases, selected purposefully to permit in-depth enquiry into, and understanding of, a phenomenon (Patton, 2002: 45 f.).

Quantitative methods are used to measure outcomes and impacts of a specified variable, in this case training programmes, for participants and to compare them with those same outcome categories for people who did not participate. They provide information on accountability, efficiency and effectiveness of an intervention. The application of quantitative methods, however, encompasses several methodological difficulties to measure the 'true' effect of an intervention:

- (a) distinguishing the programme effects from the effect of other factors;
- (b) determining what would be the hypothetical outcomes for the same people if they had not participated in the programme;
- (c) taking into account short-, medium-, and long-term outcomes;
- (d) avoiding selectivity and heterogeneity biases while ensuring validity.

Table 2. Comparison of different evaluation methods

Method	Description	Advantages (+) and disadvantages (-)
Social experiment	Compares the outcome of a programme for participants and non-participants, selected randomly before the programme takes place.	<ul style="list-style-type: none"> + full control, unbiased selection and estimation of 'true' outcomes + unobservable characteristics become irrelevant - exclusion of eligible persons from programmes is problematic for ethical and legal reasons - uncontrollable interaction and moves between treatment and control group that bias results
Matching (quasi-experiments)	Compares the outcome of participants in the period after the programme, with the outcome of matched (statistical twins) non-participants in the same period.	<ul style="list-style-type: none"> + no selection bias caused by unobservable characteristics - needs data for participants and non-participants before and after the programme (e.g. labour-market histories) - unobservable characteristics might bias results
Before-after comparison	Compares the outcome for participants before and after the programme took place.	<ul style="list-style-type: none"> + easy to implement + low data requirements (no information for non-participants needed) - general economic changes before/ until/ after the programme period might be falsely attributed to the programme - Ashenfelter's dip ^(a) (change of behaviour in the period preceding the programme)
Difference-in-differences	Compares the change of before-after outcomes for participants with the before-after change of outcomes for non-participants.	<ul style="list-style-type: none"> + takes account of selection on unobservable characteristics - needs data for participants and non-participants before and after the programme took place - Ashenfelter's dip ^(a)
Cross-section	Compares the outcome for participants in the period after the programme with the outcome for non-participants in the same period.	<ul style="list-style-type: none"> + economy-wide changes are not attributed to the programme - needs data for participants and non-participants for the period after the programme

^(a) 'Ashenfelter's dip' occurs when anticipation of participation in a programme affect the outcome category (e.g. if individuals reduce their job search behaviour in order to become eligible to participate in a programme)

Source: Hujer et al., 2004a: Section 2.2.7.

To overcome these methodological difficulties, quantitative evaluation methods apply various sophisticated evaluation designs (Table 2) and statistical models.

There are mutual critiques and misunderstandings between the respective supporters of quantitative and qualitative methods. The primary critiques of quantitative evaluations are: oversimplifying the complexities of real-world programmes and participants' experiences; missing major factors of importance that are not easily quantified or observable; and failing to portray the programme and its impact as a whole.

In contrast, qualitative inquiry is accused of being subjective and unscientific; qualitative 'soft' data can be dismissed as mere anecdote. Evaluators and researchers can be locked in their philosophies, paradigms, schools of research and preferences. However, when evaluating programmes and policies, pragmatism is required.

The logic of qualitative enquiry is more one of discovery, trying to uncover processes and particular dynamics of cases, being open to realities as they appear to the participants or other stakeholders. Quantitative methods adopt more a logic of verification and research to discover regularities or even 'laws' or to make generalisations, trying to understand what works and what does not work and thus to ensure accountability. However, both work towards designing future interventions and improving current ones. They offer different perspectives on the processes and outcomes of the programme under evaluation. Their use is also partly defined by the stages and aims of evaluation. Therefore, both quantitative and qualitative methods should be used, depending on the stage and objectives of an evaluation exercise.

2.2. Impact research

Impact research is different from evaluation insofar as it does not focus on assessing the results of specific programmes but investigates the impact – both material and non-material – of education, training and skills on economic and social development as well as on company and individual performance. By reviewing most common theories, objects, approaches and methods of impact research, this chapter provides the theoretical and methodological background necessary to understand the results of impact research that are presented in Part 4.

The benefits of education, training and skills, or of human capital, extend beyond economic or monetary aspects – though these are important issues to justify public and private expenditure in education and training – to include non-material benefits as well. Moreover, the presence of external effects which are to the benefit of other people, companies or society in general are – if positive – an important additional justification of public educational investments (Figure 2).

2.2.1. Education, training and economic growth

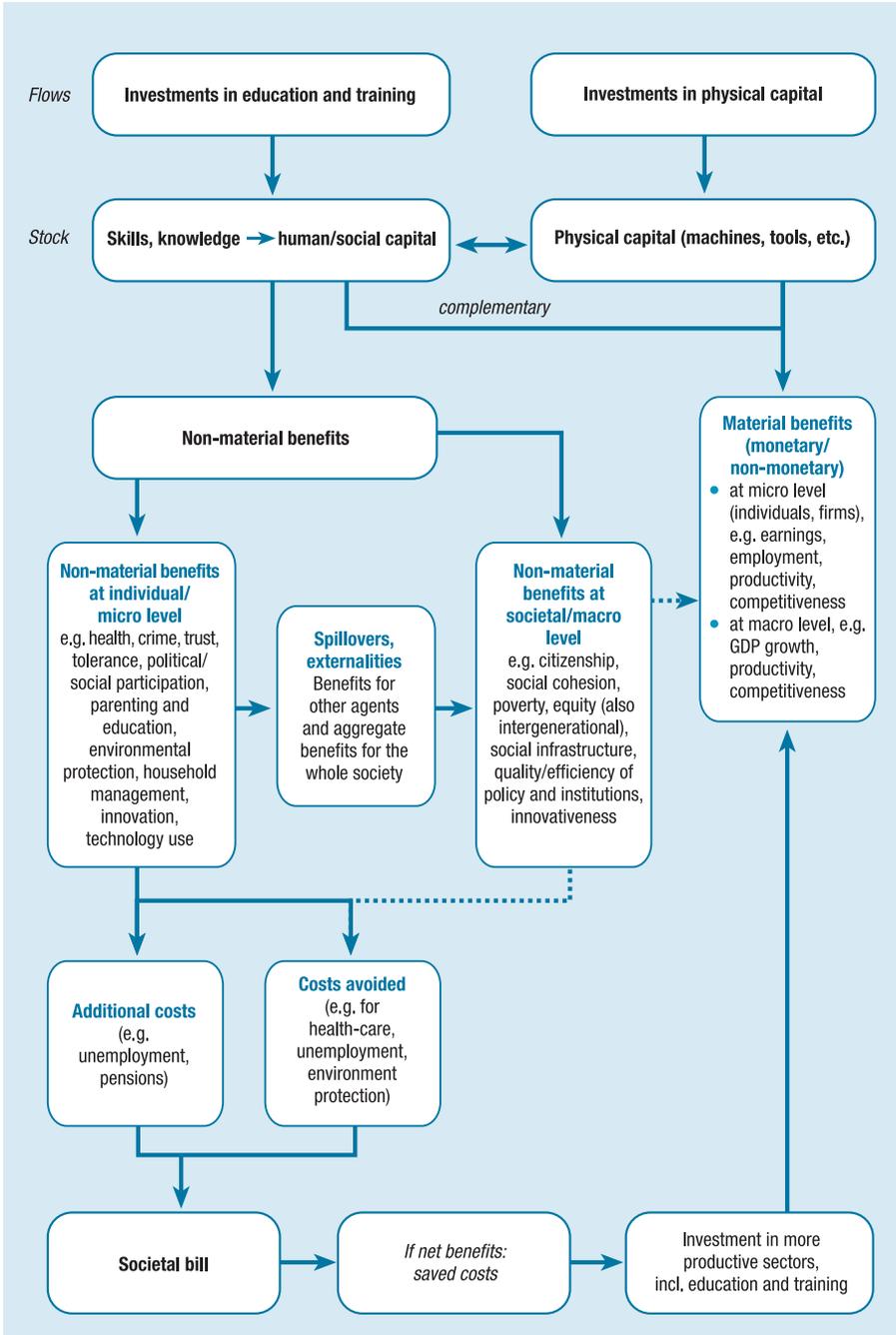
In new theories of endogenous growth, human and social capital and research and development (R&D) are seen as key elements in creating, using and disseminating new knowledge and ideas. They consider education, training and skills the main determinants of growth in the long-term⁽⁴⁾. Thus, education, training and learning should be at the heart of State or company policies to foster growth and productivity.

Another important issue discussed in research on endogenous growth is whether the most important determinant of growth and productivity is accumulating human capital (i.e. investments and participation in education and training), or human capital stock (i.e. the skill levels of workers). This distinction reflects two major strands of endogenous growth approach. In the framework of human capital accumulation, a subsidy to education or an educational programme which raises the level of human capital will have a limited once-and-for-all effect on GDP growth. In the second approach, the growth rate of GDP will be increased forever, via dissemination of knowledge and technology. The first approach would imply policy actions to upgrade human capital in an economy, whereas the second would also require policies to foster R&D and technology. From the current state of research, one cannot conclude what policy should be preferred and when.

Researching the link between human capital and growth is limited by the poor quality of data – particularly in cross-country comparisons – and by the fact that human capital is mostly measured by formal education and training and qualifications. Individuals (and firms) are likely to underinvest in education and training because they do not value or reap the wider benefits. Improving transparency on the benefits for all agents in an economy is likely to foster education and training investments and thus economic growth.

⁽⁴⁾ 'Whatever education does for growth, it is unlikely to do quickly' (Temple, 2000: 41).

Figure 1. Benefits of education and training



2.2.2. Social benefits and social infrastructure

Empirical research on the wider social benefits of education and training at macro or micro level has developed considerably in recent years. It has been supported by the availability of large world-wide surveys (such as the World values surveys) which include a number of interesting indicators for social benefits, such as trust, tolerance and criminal behaviour.

External effects, spilling over from skilled workers or R&D activities to other people, companies or society can be direct or indirect outcomes of education and training. In cases such as criminality reduction, social cohesion, trust, etc., the effects of education and training are indirect, via, for example, reducing poverty or increasing the efficiency of public services. In this context, a 'good' social infrastructure (and, closely related, social capital), i.e. a context in which human capital can grow, exerts an important positive influence on economic growth and efficiency (Coleman, 1990, cited by Healy, 2000).

There are several ways of measuring social infrastructure and social capital and investigating the full social returns on education and training. They use quantitative and qualitative techniques to measure, model and evaluate these concepts and the links between education, training and social outcomes.

2.2.3. Impact on company performance

Economic growth denotes the growth in production of all goods and services in a given country and period of time. These goods and services are mainly produced by companies. However, in contrast to the large body of research on macroeconomic determinants of growth, the factors that influence companies' performance and production growth – and in particular the role of education, training and skills – are significantly under-researched.

The main methods of carrying out company surveys, i.e. cross-section and longitudinal surveys, have relative advantages and disadvantages. Cross-section surveys are less expensive and can include more information than longitudinal surveys or enterprise panels. The latter, however, are able to take account of heterogeneity of companies and their development over time. Cause-effect relationships, for example between training investments and outcomes, and changing behaviour of firms in terms of adapting to new challenges, can only be explored by longitudinal data.

Existing studies struggle with missing or inappropriate data, various – and in many cases not comparable – definitions of training, costs and performance measures and restricted information on the development of companies over time. The weakness of data and methods used in many studies understate the role of training in increasing productivity and competitiveness and in generating profits.

2.2.4. Individual benefits and life-course research

Investigating individual benefits of education and training was, until recently, a domain of microeconomic and econometric research, relying on human capital theories. In most cases, the benefits of education and training studied are material, in particular monetary returns (earnings), rates of return and some forms of non-monetary returns such as unemployment probability, occupational career, etc.

A recent strand in education and training research at micro level is life-course and biographical research. Life-course research focuses on decisions individuals have to make at specific transition points in their life history and their influence on subsequent pathways. Biographical research focuses on the subjective perception of events in life which have an important impact on social stratification processes.

Most research studies on the benefits of education and training in a life-course perspective focus on participation in education and training and the labour market (including unemployment spells) and on occupational career and status changes. However, many factors intervene during the life-course of an individual and render the attribution of single effects of education and training difficult after some years. This is why many questions can only be answered using longitudinal data, or interpreting biographical perceptions of individuals. These include issues of equal opportunities, transition behaviour, abilities and performance, which are not stable over time but change in relation to external factors such as the labour-market situation, provision and delivery of education and training, the economy, and changes in social and political environments.

Another issue studied in life-course research is the influence of social differences, in particular gender, social origin and ethnicity. The added-value of life-course research in explaining social differences comes from the fact that inequalities, for example between gender, social classes and ethnic groups, change over time. Research also attempts to answer the question of whether inequalities continue beyond the life-course of individuals, by the reproduction of inequalities in successive generations.

A related strand of research is the cohort approach, i.e. observing the pathways of people with equal starting positions (e.g. year of birth or year of leaving compulsory school) over a longer period of time. By comparing the participation of different cohorts in education, training, employment and unemployment and the changes over time, social change becomes visible and quantifiable. Research on intergenerational mobility and on the factors that influence participation behaviour are important subjects for cohort approaches.

3. Evaluation of education and training in a changing European context

Part 3 discusses practically how evaluation can support the implementation and assessment of long-term education and training system reforms and of more targeted and short-term interventions such as active labour-market policies and programmes. But, before discussing evaluation practices, it is important to sketch the context in which education and training programmes and policies are implemented and the role they are expected to fulfil. European labour markets are faced with changes:

- (a) higher level skills are needed in modern global economies, especially in knowledge intensive-sectors, but at the same time labour markets are becoming more polarised, with more precarious and repetitive jobs for the less educated;
- (b) the population of the EU is ageing and this trend will be accelerated by the accession of the 10 new countries;
- (c) the initial education level of the EU population is improving but a substantial proportion is still characterised by a low level of formal education;
- (d) unemployment is a concern for EU economies, especially after enlargement; further efforts are required to fight against it through targeted policies and programmes.

Therefore, skill renewal through lifelong learning is becoming increasingly desirable in EU countries (Figures 2 and 3 present a picture of the current situation).

Overall, these changing contexts require policy interventions. They can take the form of either short-term targeted programmes, perhaps aimed at quick reintegration of people into the labour market, or more profound and long-term system reform ⁽⁶⁾. These interventions also aim to address various

⁽⁶⁾ We do not discuss here the various reforms in detail but focus on how the reforming process can be steered, using evaluation, and how results and impact of reforms can be evaluated. The reader interested in detailed reviews of various reforms of education and training systems in Europe may refer to Descy and Tessaring (2001a, various chapters).

Figure 2. Participation in continuing education and training (a) of the population aged 25-64 years by level of educational attainment, 2003 (%) (b)

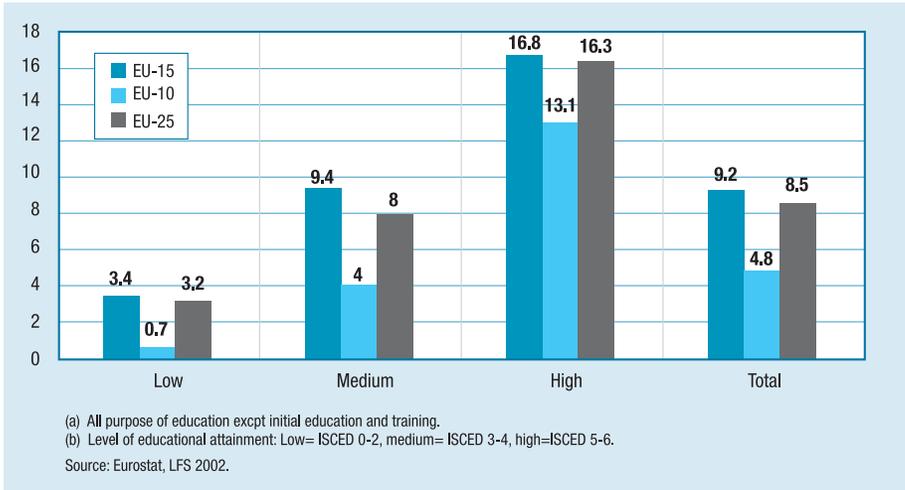
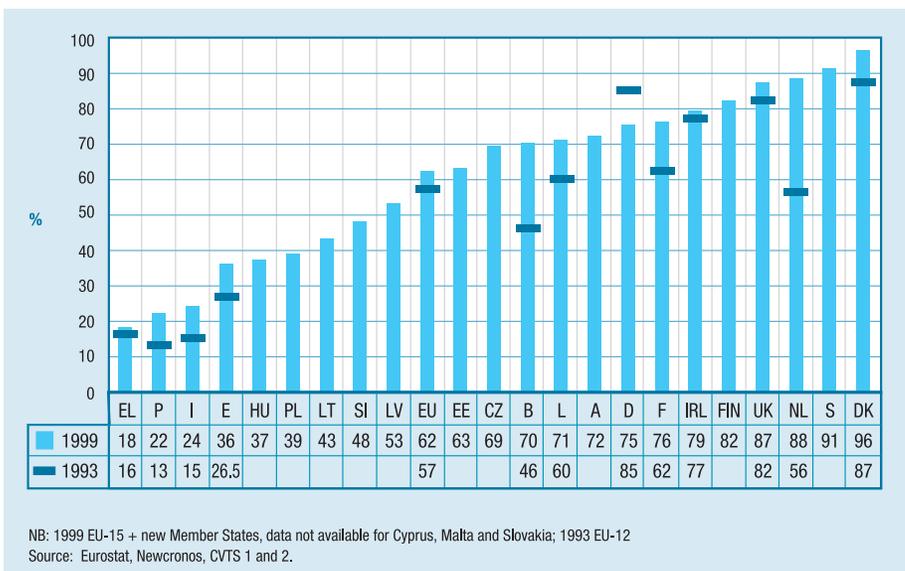


Figure 3. Participation in continuing vocational training courses, EU-15 and new Member States, 1993 and 1999 (% of employees, all enterprises)



social issues. Evaluation is a tool both to accompany policy implementation and to assess whether the problems in question have been addressed.

Education and training are complex systems that do not exist in isolation but have social and economic roles. Reforming them is a process that requires debate and compromise on what is to be changed and for what reasons, not only between those in the systems but also other stakeholders, especially social partners. It is essential to envisage and to evaluate the internal consistency of a new policy with other elements of the system (e.g. one cannot implement changes in curricula and in learning methods without involving – and properly training – teachers). But reforms have also to take into account the needs – and be coherent with the modes of functioning – of the production system and of society more generally (external consistency of a policy) (Box 2).

Traditional evaluations of policies (or programmes) focus strictly on policy as a separate entity, treating it as if it was self-contained and independent from the historical, structural and institutional context. This reinforces the tendency to design policies that are limited in scope and do not take into account the existing institutions and interventions as well as the modes of functioning of economic and social systems. We favour instead a systemic approach, leading to more coherent approaches to education and training, and VET reforms. However, while a fairly robust evaluation methodology exists for programme evaluation, this is not the case for system evaluation.

Viertel et al. (2004, Section 7.1) conclude: ‘there is no holy grail in terms of conceptualisation or methodology related to VET policy evaluations. The engineer’s toolbox is of limited use. [...] 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.’

Box 2. Lack of internal and external consistency while implementing reforms of VET in selected Phare countries

One of the main objectives of the Phare programme in Central and Eastern European countries is to support the process of reforming VET to make it more responsive to the demands of the market economy. VET has to be made more flexible to respond to changing demands as well as to individual needs and preferences. This should ideally be achieved while ensuring a certain consistency of VET provision and qualifications, thus avoiding a lack of national standards and fragmentation of the system.

In Bosnia-Herzegovina, the reform of VET included changes in VET equipment, curricula and methodology. However, vocational teacher training was not originally a project component. The reform was therefore not fed into pre-service (and in-service) training of teachers. Inspector training was neglected as well. This created a serious barrier to broad acceptance of the new policy. This situation was corrected later and a budget reserve of EUR 200 000 allocated to run a programme for teacher and management training.

The curriculum design adopted in Estonia, modular-based and employer-led, is changing the logic of the VET system. Skill needs and educational goals have been identified and occupational profiles translated into curricula, while employers and their representatives are influencing VET provision. However, the infrastructure supposed to support curriculum implementation is very fragile. Clearly the National Centre for Examination and Qualifications is methodologically and technically too weak. This Centre, whose role is to award and accredit qualification, is the cornerstone of the new VET system. This puts the momentum of the innovation process and the fragile interaction between the employment and the education system at risk.

In Romania, a pilot Phare VET reform was implemented in about 10 % of the VET schools. After that phase the government decided to carry out this reform in all schools. However, no new equipment or training was provided to accompany this. The national inspectorate, only half familiar with the reform, was asked to support the process. The careful necessary planning and the substantial resources needed to procure new equipment and to train both teachers and school managers were not present.

Source: Viertel et al., 2004, Section 4.4.

3.1. Learning by comparison

The common transformations observed, and the laboratory of reforms and new policies generated, demand policy learning both within and between countries. Mutual learning should take place and debates should be triggered by investigating others' 'solutions'. This is one way to develop learning economies. Two particular cases may be contrasted in this respect:

- (a) the OECD policy review methods, where interdisciplinary teams of researchers undertake country visits, generating national and comparative evaluations as well as recommendations for future policy design (Pont and Werquin, 2004);
- (b) the application of benchmarks in the EU, setting common targets for education and training policy in the framework of the Lisbon strategy.

OECD thematic policy reviews can be seen as a form of a peer review in international comparisons. They examine under what conditions a policy may provide successful results to offer lessons for the future and for the countries being compared. They try to identify good practices and to measure their transferability. A thematic review relies on combining a self-evaluation carried out by the country under review (in the form of a background report) and external evaluation (carried out during a country visit) from a team of experts. All the steps followed during a thematic review are designed to optimise the depth and the quality of the process, which is enriched by constant exchanges, during the visit and after, between the people in charge of the project in each country and the experts.

The strength of a thematic review is that it is an independent evaluation led by external experts which provides reference points for national policies in an international context. It also obliges institutions, ministries, social partners and providers to combine their efforts to offer the most appropriate overview of the system at all stages of activities. Therefore, it becomes an opportunity for all actors to come together, to be participative in their own evaluation and to learn about their own system and initiatives. It generates collaboration between the various actors and mutual learning (even sometimes between policy-makers and practitioners).

Within the EU, benchmarking is used as a tool to reach the goals set at the Lisbon Summit in 2000. It has been used in employment policy ⁽⁶⁾ and it is progressively being applied to education and training (see EC, 2001;

⁽⁶⁾ The use of benchmarks in social policy at EU level was initiated in the context of the process started in Luxembourg in 1997 and called the European Employment Strategy. Employment benchmarks are now an integrated part of the Lisbon Strategy.

Box 3. The open method of coordination

While respecting the breakdown of responsibilities envisaged in various EU treaties, the open method of coordination is a way of spreading knowledge of best practices and achieving a greater convergence towards the main EU goals. It is a new form of cooperation for the Member States based on a fully decentralised approach using variable forms of partnerships and designed to help them to progressively develop their own policies. It is based essentially on:

- identifying and defining jointly the objectives to be reached (the benchmarks, see Box 4);
- commonly-defined yardsticks (statistics, indicators) enabling Member States to know where they stand and to assess progress towards the set objectives;
- comparative cooperation tools to stimulate exchange and dissemination of good practices.

Source: http://europa.eu.int/comm/education/policies/pol/policy_en.html#methode; detailed work programme on the follow-up of the objectives of education and training systems in Europe (Council of the European Union, 2002).

EC, 2002). In this respect, the Commission and the Council have adopted a common work programme and defined the method of open coordination. Box 3 explains this method and Box 4 presents EU benchmarks in education and training.

In the EU, the main obstacles for benchmarking in education and training are developing appropriate indicators and the availability and comparability of data. Furthermore, various concerns may be expressed about the approach taken:

- (a) using the EU average as a benchmark implies that if the best performing countries make further progress, this will 'relieve' others, less advanced, from the 'burden' of investing in education and training over-proportionally;
- (b) the targets set refer to a non-weighted EU average, i.e. it does not take into account the size of the population in each country. Therefore, progress in a large country affects the average more than a progress in a small country;
- (c) there could be some saturation level at which progress cannot be made any longer. In all societies, there will be always a group of people who are not able or not willing to study further (Tessaring, 2003);
- (d) the various benchmarks set are treated as independent but are, in fact, interrelated. For example, reducing early school drop-out and increasing the number of young people who leave education with at least the upper secondary level are two complementary objectives;
- (e) countries' relative necessary progress is not considered. The rationale behind benchmarks is 'low achievers have to catch up' and in some

Box 4. EU benchmarks for education and training

To achieve the Lisbon European Council's strategic goal to make the EU by 2010 'the most competitive and dynamic knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion', the heads of States and Governments agreed, in 2001 in Stockholm, on some common objectives for education and training in the EU, within the overarching framework of lifelong learning:

- improving the quality and effectiveness of education and training systems in the EU;
- facilitating access of all to education and training systems;
- opening up education and training systems to the wider world.

Subsequently, the 2002 Barcelona European Council set a new goal 'to make Europe's education and training systems a world quality reference' by 2010.

The work programme for achieving these objectives foresees that progress will be monitored against agreed indicators expressed as average level performance of the EU-15 Member States and of the three best performing ones.

Consequently, the Education Council agreed in May 2003 on the following European benchmarks:

- 'By 2010, all Member States should at least reduce the rate of early school leavers, with reference to the rate recorded in the year 2000, sufficiently to achieve an EU-average of 10 % or less.
- By 2010, the total number of graduates in mathematics science and technology in the EU should increase by at least 15 % while at the same time the level of gender imbalance should decrease.
- By 2010, at least 85 % of 22 year olds in the EU should have completed upper secondary education.
- By 2010, the percentage of 15 year olds with low achievement in reading literacy in the EU should have decreased by at least 20 % compared to the year 2000.
- By 2010, the EU-average level of participation in lifelong learning should be at least 12.5 % of the adult working age population (25-64 age group).'

In addition, 'the Commission invites Member States to continue to contribute to the achievement of the Lisbon objectives of substantial annual increases in per capita investments in human resources, and, in this respect, to set transparent benchmarks [...].'

Source: EC, 2002b; Press release 5/05/2003, available from Internet: <http://europa.eu.int/rapid/pressReleasesAction.do?reference=IP/03/620&format=HTML&aged=0&language=EN&guiLanguage=en> [cited 20.8.2004].

cases there is a long way to go. Furthermore, is there really no progress to be made in best performing countries?

- (f) another assumption hidden behind these benchmarks is the idea that 'since some education makes some of us rich, more would make more

of us richer, or “if two aspirins are good” it means that “five aspirins are better”” (Wolf, 2002; p. 28). This assumption is not necessarily true; substitution mechanisms and overqualification, for example, tend to prove it;

- (g) finally, the efforts to reach each education and training targets in each country should be converted into financial figures and additional support to countries where substantial financial investment would be required, should be considered (Tessaring, 2003).

Country benchmarking is rarely more than just an international comparison exercise and cannot be used for one country to learn directly from another. Nevertheless it has the merit of opening up political and national debate. In the EU, it provides common targets and frameworks of reference but needs to be complemented by a considered and tailor-made policy for each context. Perhaps the solution is to develop tailor-made benchmarking, i.e. to define a common framework for the Member States but then to reflect on tailor-made targets, areas for improvement and desirable results for each country given its ‘starting point’.

3.2. Examples of evaluation of VET systems reform: the Dutch and Danish cases

In Denmark and in the Netherlands, both the reforms of VET systems and their evaluation were quite different. In Denmark, evaluation was applied in a formative perspective, accompanying the process of reforms while in the Netherlands evaluation was carried out *ex post*, four years after reform implementation.

The reform (VET Reform 2000) of the Danish VET system was initiated, designed and implemented in ways that consciously aimed at fostering a learning process – a ‘reflective gathering of experience’ – between policy formulation and implementation, using interactive/formative evaluation. Evaluation was used twice, as a corrective and improvement mechanism: after the pilot stage of the reform and immediately after the expansion of the reform to the entire system. Evaluation is then part of the learning and changing process; the evaluators are involved, with system stakeholders, in the reform. This case illustrates how evaluation can be used to accompany reform, involving different actors not only as target groups for the evaluation but also in the evaluation design and, in doing so, developing a mechanism for (early) feedback and adjustments. The strength of the evaluation of the

Danish reform was not only the material produced and published but the willingness and ability to present and debate findings and policy recommendations with different stakeholders. Used in that way, evaluation not only supports the achievement of the initial aims of the reform, it also permits learning from the actors involved at different levels, from policy makers to students or teachers. This kind of evaluation is formative: it seeks to influence and promote change rather than providing ‘proofs’ about the effectiveness – or non-effectiveness – of new policies.

In the Netherlands, a new law on vocational and adult education was launched in 1996 to increase the efficiency and responsiveness of the VET system: the WEB (7). Built into the WEB was an *ex post* evaluation. In 2000, various evaluation teams started to work on different aspects, leading ultimately to a single report. The approach followed was summative; i.e. the research teams were supposed to take an *ex post* ‘objective’ picture of the state of the art on their respective subject. This involvement of various research teams, with different specialities and backgrounds and different evaluation topics is an interesting feature of this evaluation. However, in a system like the Dutch one, where all political and social actors were involved in the formulation and implementation of the WEB, the summative approach made it difficult to identify failures, possible reasons and remedies *a posteriori*. This limited the value of evaluation results. Absorbing the results of a reform evaluation means adjusting and correcting. The fact that the evaluation was conducted five years after large-/system-scale implementation generated some resistance to change (again) from the actors of the system because new mechanisms were already in place and systemic changes had already occurred. Summative evaluations are most useful when used after a pilot phase (before a new policy is implemented full-scale) and in collaboration with the main reform partners (institutions, teachers and trainers, individuals/participants). They provide an opportunity to identify success, failures and areas for improvement and can feed into reforms and policy formulation. Their function is then formative.

According to Lundvall (2000, cited by Nieuwenhuis and Shapiro, 2004: Chapter 3) knowledge construction and obsolescence is taking place at an accelerated pace. Therefore, it is not so important any longer to have knowledge *per se*. The key to success is learning (and forgetting). In the evaluation context, this means that new purposes for evaluation are

(7) WEB = *Wet Educatie en Beroepsopleiding* (Law on [adult] education and VET).

emerging. Policy evaluation should also become an instrument of critical reflection and learning, in addition to measuring results and performance. Thus, it should form an integral part of education and training policy. Assessing the effectiveness and the utility of actions and policies, and accompanying their implementation, results should guide decision-makers in choosing the most appropriate actions for specific objectives: to gain consensus; to maintain involvement from all actors, from learners to investors; and to make policies viable in the long-run (Pont and Werquin, 2004).

3.3. Combating labour-market exclusion: does training work?

The rise of unemployment and the tightening of public budgets generate growing concern about the effectiveness of active labour-market policies (ALMP) designed to reintegrate people quickly into the labour market. Evaluations are, therefore, increasingly carried out to assess whether this objective has been reached and at what cost.

Training is not the unique form of ALMP but, in comparison to other ALMP, it receives the most important share of public expenditure (Figure 4). It must be considered whether it achieves the expected results in comparison to other types, such as job search assistance or subsidies to employment. The conclusion of the review of existing evaluations is that the effect of training on employment and earnings tends to be modest and that it is, at the same time, a very costly type of ALMP. In addition, it is not effective for all target groups, across all regions or all episodes of the business cycle. The conclusion might be that training does not work well enough and that other types of ALMP should be preferred. We argue that we need further evidence before drawing these kinds of conclusions.

The traditional evaluations of ALMPs 'compare' explicit programme goals with measurable outcomes (mainly in terms of employment probability and earning) and look at cost-benefits. They allow estimation of the relative effectiveness and/or efficiency of different interventions. Most of them are quasi-experimental, with emphasis on the econometric elaboration of programme outcomes. Because traditional evaluations do not open the black box, their results are generally limited in terms of indicating possibilities for change and improvement of programme design and

implementation. Instead, they lead to a decision either to stop or to continue programmes. They concentrate on what works rather than trying to answer why it works or not. In order to design new and more efficient policies, policy-makers need to know ‘what worked?’ (or not) as well as ‘why?’ or ‘under what circumstances?’ Therefore, the focus of ALMP evaluation should be broadened to provide additional evidence. It should ⁽⁸⁾:

- (a) address the determinants of various outcomes, compared across subgroups of the labour market and explore participants’ motivational factors;
- (b) be carried out over longer periods than now is the case, to see if the short-term effects on earnings and employment prevail in the long-run and if there are some delayed effects ⁽⁹⁾;
- (c) prefer increased employment probability or reduced unemployment duration as outcome variables. Earnings and wages might be questionable outcome variables in the European context owing to the fact that welfare state and minimum wage regulations result in distortions between employment status and earnings;
- (d) extend the scope of effects (or side-effects) for consideration of non-material ones (increased self-worth, better health, wider social and behavioural gains, further learning experiences, etc.) although they do not constitute a primary objective of ALMP;
- (e) be foreseen from the outset when an intervention is designed so that the necessary information is gathered in a continuous process;
- (f) attempt to assess better the effect of the structure, content and design of training courses as well as whether they can be adapted to different circumstances, such as changing labour-market needs.

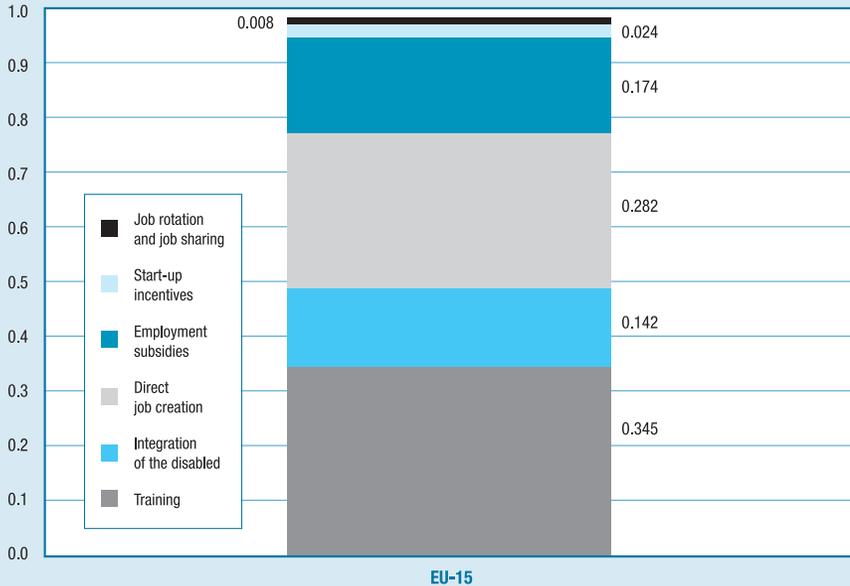
In addition, also in the case of ALMP, a more systemic approach to evaluation should be envisaged to broaden the current perception of effectiveness by focusing on the way interventions are articulated with other VET programmes, employers’ hiring practices, and with existing institutions. Other systemic issues are whether a programme can be expanded, replicated and the cumulative impact of interventions (Schmid et al., 1996: 2).

⁽⁸⁾ Based on the following literature: Fay (1996, cited in Walsh and Parsons, 2004), Martin and Grubb (2001: 16), Walsh and Parsons (2004), Hujer et al. (2004a), Grubb and Ryan (1999).

⁽⁹⁾ This aspect is crucial because programmes have been established to address systemic and enduring problems such as unemployment, social exclusion, transition from school to work, etc. These are the kinds of problems that require long-term solutions. It must, therefore, be established if the short-term results of a programme (what is generally captured by evaluation) are sustainable.

Evidence-led policy-making requires information on the circumstances, both exogenous and endogenous, that lead to programme quality, positive outcomes and cost-effectiveness. Policy evaluation should define the criteria and provide supporting empirical evidence on which policy types and which policy mix promise superior solutions to a society's problem (Schmid et al., 1996: 12).

Figure 4. **Public expenditures on ALMPs, EU-15, 1999 (% of GDP)**



Source: Eurostat, LMP data collection, Statistics in focus, EC, 12/2002.

4. Impact and benefits of education and training

4.1. Determinants of economic growth

Stimulating economic growth and employment is one of the principal objectives of public investments in education and training; in addition, increases in growth and employment are likely to facilitate achieving social objectives and avoiding social inequalities. Growth in gross domestic product (GDP) comes from the combination of physical capital, labour and human capital and the investments made in them.

A number of empirical studies carried out in the past 10 to 15 years identify skills and related investment in education and training as the key determinants of economic prosperity, which is an important precondition for social cohesion and stability. OECD countries have invested heavily in education and training and this has had direct effect on growth rates of GDP per person employed and on labour productivity. Wilson and Briscoe (2004) conclude on this issue: 'Overall, these growth models demonstrate that higher educational investments have had a significant impact on national economic growth. Broadly, the weight of evidence suggests that a 1 % increase in school enrolment rates has led to an increase in GDP per capita growth of between 1 and 3 %. An additional year of secondary education, which increases the stock of human capital, rather than simply the flow into education, has led to a more than 1 % increase in economic growth each year.'

On the relative contribution of various levels of education, it seems that primary and secondary skills contribute to growth in the poorest and in intermediate developing countries respectively while tertiary skills are the most important for growth in OECD countries. Calculating the implications for growth of achieving the EU target for human resources and a knowledge society has, therefore, to take into account the different levels of advancement across European countries. Research on the contribution of vocational training versus general education is practically non-existent and no conclusions can be drawn so far in this respect.

Another issue raised by research is that the quality of human capital is crucial for economic growth, as it is for individual and company performance. It seems to be even more important a factor than the quantity of human capital as measured by years of schooling or level of education. Raising the quality of education should, therefore, be at the centre of human capital policies. Empirical work points towards some concrete steps to accomplish this objective, but considerable uncertainty remains and more research is necessary to identify the determinants of school performance and student achievement. It is clear, however, that the objective of raising the average quality of human capital does (or should) not stand in contradiction with the objective of enhancing social cohesion.

The observation that research findings display considerable cross-country and regional differences in terms of economic growth and human capital of populations gives rise to a basic question: what makes some countries or regions accumulate more human capital than others or what makes them more efficient than others in the use of such inputs? Social capital and infrastructure can be seen as both independent framework conditions for economic success and the outcomes of increased education and training investments. Higher levels of human and social capital, and of social infrastructure, are associated with better environment, higher levels of health, trust and greater social cohesion. All these effects have been found to feed back into faster economic growth. However, studies are rather disparate in empirically proving this link, not least due to the difficult operationalisation of many of the ingredients of social capital.

In conclusion, measures aimed at increasing the quantity and quality of human capital should be an important part of any growth-promoting policy package. Implementing adequate human capital policies 'appears especially important for those regions of the EU that are lagging behind in productivity and income per capita. It is important to recognise, however, that successful action requires a clear picture of the quantity and quality of regional human capital stocks in order to understand local needs and to identify those policies that are likely to be most effective. For example, it would be important to extend to the regional level recent studies that have tried to assess the skill levels of younger cohorts and of the workforce at large, and to support further research into the determinants of the performance of educational systems. These studies can be a useful input for the formulation of a systematic human resources policy that should be an important part of the EU's ongoing effort to increase regional cohesion.' (de la Fuente and Ciccone; 2002: 8).

In addition to its direct effect on growth, human capital may yield additional indirect benefits if it stimulates the accumulation of other productive inputs such as physical capital, technology, health, crime reduction and social cohesion. These, in turn, may foster economic growth. Human capital is a rather attractive investment alternative – for the State as well as for individuals and companies – if one considers also non-market returns and wider social benefits in terms of health, crime reduction, trust, citizenship and social cohesion.

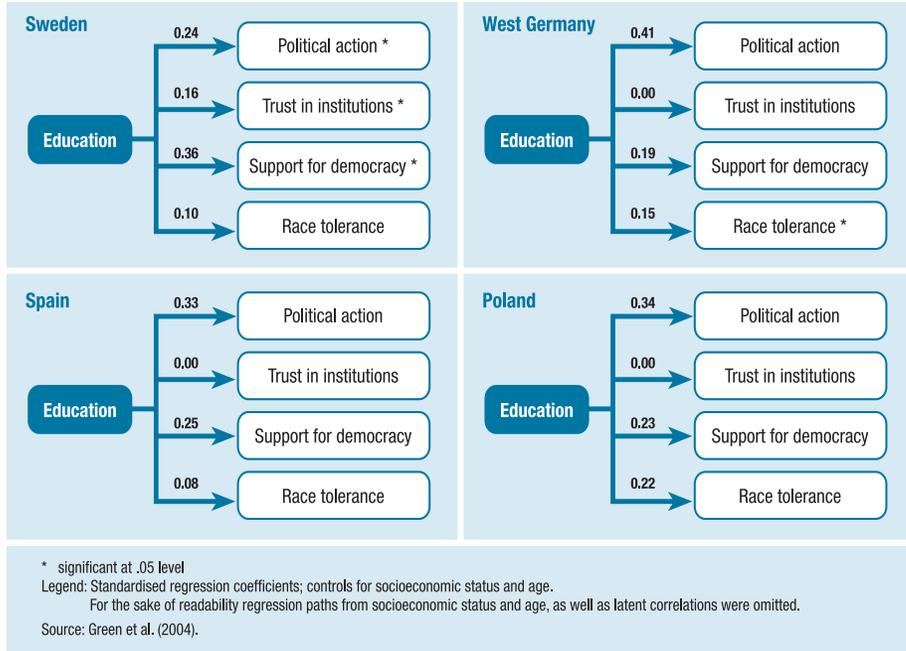
4.2. Non-material benefits and externalities of education and training

The effects of education, training and skills on macrosocial outcomes including crime, social cohesion, citizenship, civic and political participation can be analysed through both macrosocial aggregates and microsocial data. Across EU countries, some generalisations can be made concerning the macrosocial benefits of education, training and skills (Green et al. 2004).

For some macrosocial benefits there are common antecedents, i.e. other influences which themselves are closely associated with education and skills. For example, criminal activity and low tolerance could be associated with unemployment, poverty and alienation, which themselves are closely correlated with education, training and skills. This is confirmed, for example, by research on football hooliganism (Dunning, 2000), juvenile delinquency and hate crime (Watts, 2001) for which unemployment and alienation, both related to education, are antecedents. Equally, income and educational inequality are structural antecedents of crime (Kelly, 2000; Lee, 2000). Green et al. (2004) find clear relationships between education inequality, income inequality and general trust, crime and feelings of community safety.

The effects of education can also be studied using microsocial data within a cross-country comparative context. Figure 5 shows, for the countries selected, the effect of education on the four microsocial indicators of social cohesion: political action, institutional trust, support for democracy and race tolerance. The effect of education was controlled for socioeconomic status and age. Each arrow shows the effect size (regression coefficients). Significance tests were used to show where the effect sizes differed significantly between countries.

Figure 5. **Effects of education on social cohesion in four countries**



The perception of macrosocial benefits such as trust, crime and tolerance depends greatly on societal norms and inequalities in a given country; these are hard to change. In the longer term, this change could be fostered by education or training and other institutions involved in state formation (Eisner, 2001). Thus, education and training have an important role to play in the formation of values and the removal of inequalities.

Green et al. (2004) illustrate this point by taking two examples: the Nordic countries and the UK. Nordic countries are, in general, characterised by high trust, low crime but moderate levels of civic participation. In the Danish case, this goes along with high levels of lifestyle permissiveness but rather low levels of tolerance towards foreigners. The high levels of trust may be associated with the strong welfare states and historically high levels of ethnic/cultural homogeneity (Knack and Keefer, 1997). However, this may also be a detriment to social cohesion in increasing ethnic heterogeneity. High levels of trust may also relate to relatively high levels of income equality. In this context, education equality may promote trust and lower crime through its impact on income equality. In Sweden, the strong effect of education on trust in institutions and in democracy may be attributable to

social solidarity principles imparted in curricula and in the universal non-selective nature of the primary and secondary school systems. However, evidence that education has a weak effect on civic association in this country may 'reflect the fact that civic participation in Sweden is highly institutionalised, not least with the prominent role of trades unions within the social partnership system.' (Green et al., 2004: Section 8).

As a counter-example, the UK has high levels of crime and – compared to the other countries analysed – low levels of both trust and also of tolerance (Halman, 1994). Higher crime and lower trust may reflect the high income inequality (amongst the highest in the EU), and higher levels of intolerance may partly be due to the high immigration over the past 40 years (Halman, 1994) ⁽¹⁰⁾. Education may play a part in generating lower levels of trust and higher crime through its impact on income inequality. A highly marketised and competitive system as the UK with high levels of inequality in outcomes between schools and regions, and consequently educational inequalities, may generate income inequality and lower trust; this is the opposite of the Nordic emphasis on social solidarity in the school curriculum.

In conclusion, in a cross-country comparative viewpoint, there are specific historical conditions which influence the relationship between education, training, trust, tolerance and social cohesion more generally. Following Green et al. (2004), education and training may have important effects on many of these outcomes, provided that transmitting values such as support for democracy and race tolerance are an important part of the school curricula. However, these effects are mostly indirect and conditional on other – often more powerful – contextual determinants. Much of the work of explaining complex interactions will require more in-depth comparative analysis.

Raising educational, skills and training levels is neither a necessary nor sufficient condition for promoting macrosocial benefits. However, improving the distribution of educational outcomes may be one way in which education and training can make some contribution to more general economic and social redistribution (Green et al., 2004: Section 8).

⁽¹⁰⁾ However, this statement should be qualified as other countries, for example Germany, had the same or even higher level of immigration during the past decades.

4.3. The significance of education and training for company performance

The majority of recent research findings confirm that investments in training generate substantial gains for firms. Furthermore, there is increasing evidence that it is not a matter of whether the nature of training is general or specific, i.e. only useful for the training firm, but more a question of how to stay ahead of competitors. Increasingly firms are financing all types of training, general as well as specific. Positive training outcomes are most evident in several studies that connect training investments with changes in productivity, profitability and stock market performance. The majority of these studies also indicate the direction of these relationships, i.e. that training generates performance and not the other way around. Research on firms' investment in human capital and its impact on company performance confirms that the firms' human capital is a major determinant of performance and complementary to technological capital (Ballot, 2003). This applies also to small- and medium-sized enterprises (SMEs), although research on the links between training and economic performance of SMEs is not well advanced.

Supporting human resource development practices and analysing training needs are seen as important elements in explaining and ensuring the provision of training and training outcomes. Similarly, human resource management practices together with training are associated with firm performance and closely related to a firms' innovative capacity.

Although comparative international research on human resource management shows that the skills of individual workers are often a critical factor in an organisation's competitive advantage, employers in many countries have a rather negative attitude to investment in skills and training. Others, including capital investors, may underestimate the benefits of training investment for a company. The finding of Bassi et al. (2001), for example, suggests that most investors do not know about the payoff from investments in training, for example on productivity and stock market performance. Thus the lack of information about training in company reports leads to under-investment in profitable training projects which have a positive net present value.

Unless the institutional and legal infrastructure encourages employers to invest in training, many of them will rely on others to make this investment. Therefore, governments should provide more incentives for companies, for example through tax allowances, direct grants, etc. Such training support can provide strong spill-over effects for society.

4.4. Individual benefits of education and training

Research studies on individual returns on education and training suggest considerable individual benefits of education, training and skills, both monetary and non-monetary (Table 3). The high rates of return (RoR; between 6 and 17 %) on investment in education and training clearly shows that it constitutes a more profitable investment for individuals than others such as investments in physical capital; savings, for example, lead to returns higher than 5 % only in exceptional cases. In addition, RoR have not changed very much during the past decade, despite a dramatic education expansion. Some national data even display a widening gap between earnings of higher and lower skilled people (thus increasing rates of return). However, one must realise that this widening gap might not be the result of increased returns on education but of increased inequality in the labour market.

In addition, education and training yield considerable non-material individual benefits in terms of better health, parenting, crime reduction and social inclusion. However – as was the case at the macro level – education and training influence such benefits indirectly.

Despite these positive returns, only around half of EU-15 citizens say they are prepared to pay partly or totally for the cost of training. However, this depends on the purpose of the training concerned. People seem more inclined to pay for training when personal returns are involved, for example when the purpose is to improve their private life, to learn a language or to obtain a recognised certificate (Cedefop, 2003; Table 4) ⁽¹¹⁾.

Research also displays considerable differences between countries in average monetary returns, employment and unemployment rates and non-material benefits ⁽¹²⁾. The obvious suggestion for countries with lower labour-market participation and higher unemployment would be to increase and better target their investment in education and training. This is justified by research results which indicate substantial individual and social benefits from policies aimed at lowering the number of early school leavers and providing socially deprived groups with possibilities and incentives of continuing education and training (Asplund, 2003).

⁽¹¹⁾ This table derives from the recent lifelong learning questionnaire included in standard Eurobarometer (wave 59.0) in 2003 (Cedefop, 2003, Chisholm et al., 2004).

⁽¹²⁾ Human capital rates of return, however, ignore grown structures of a country which may have generated inequalities in the access to, and outcomes of, education and training. Furthermore, these models mostly assume a balanced situation which can not be found in reality.

Table 3. **Private and social rates of return (RoR) on education in selected OECD countries by gender 1999-2000 (in percentage points)**

	RoR on upper secondary compared to lower secondary education ^(a)				RoR on tertiary education compared to upper secondary education ^(a)			
	Males		Females		Males		Females	
	Private RoR	Social RoR	Private RoR	Social RoR	Private RoR	Social RoR	Private RoR	Social RoR
Canada	13.6	m	12.7	m	8.1	6.8	9.4	7.9
Denmark	11.3	9.3	10.5	8.7	13.9	6.3	10.1	4.2
France	14.8	9.6	19.2	10.6	12.2	13.2	11.7	13.1
Germany	10.8	10.2	6.9	6.0	9.0	6.5	8.3	6.9
Italy ^(b)	11.2	8.4	m	m	6.5	7.0	m	m
Japan	6.4	5.0	8.5	6.4	7.5	6.7	6.7	5.7
Netherlands ^(c)	7.9	6.2	8.4	7.8	12.0	10.0	12.3	6.3
Sweden ^(d)	6.4	5.2	m	m	11.4	7.5	10.8	5.7
United Kingdom	15.1	12.9	m	m	17.3	15.2	15.2	13.6
United States	16.4	13.2	11.8	9.6	14.9	13.7	14.7	12.3

^(a) comprehensive rates of return;

^(b) data for males derive from 1998 post-tax earnings data;

^(c) 1994;

^(d) for women, earnings differential between upper secondary and lower secondary levels are not large enough to permit a positive RoR calculation.

m = missing data.

Source: OECD (2003: Tables A14.3 and A14.4); see this source also for methodological explanations.

A relatively new area of education and training research is to look at the benefits for individuals during their life-course or analysing their biographies. Compared with cross-section analyses which are based on data for one point in time, longitudinal investigations make visible contexts and dependencies of human capital formation over time. Human capital formation and allocation is not static but a process in the individual life-course and historical change. In Europe, longitudinal research is concentrated in Germany, France, the Netherlands, the UK, and some Nordic Member States, whereas there is still a weak database for life-course investigations in southern and eastern European countries.

Life-course and biographical studies demonstrate the relevance of education and training for the whole individual life. They show that participation in education and training and upgrading skills have increasing

and cumulative effects for occupational career and personal development ⁽¹³⁾. However, empirical life-course research also shows the still existing selectivity of access to continuing vocational training which depends on previous educational level, gender and family background and other social factors. All these contribute to the persistence and accumulation of social discrimination throughout the life course. This implies that a mere increase in continuing training opportunities would not be an appropriate solution for improving individual labour-market participation. Instead, this might also lead to undesirable side effects such as the further polarisation of education and occupation opportunities.

Table 4. **Willingness to pay for education and training, by purpose, population aged 15 and more, EU-15, 2003 (%)**

To...	keep present job	have a better private life	get a promotion	learn a new language	set up one's own business	get new knowledge for a hobby	open up job and career opportunities	get a recognised certificate	get a pay rise	prepare for retirement	get new knowledge in one's work field	get back into the labour market
would contribute financially	37.7	51.4	38.7	46.7	45.2	46.3	48.2	48.1	39.7	34.8	43.6	39.7
would pay all of the cost	12.9	21.8	11.7	18.6	23.0	21.5	16.4	18.1	14.8	11.7	12.9	14.8
would pay some of the cost	24.8	29.6	27.0	28.0	22.2	24.8	31.8	30.0	24.9	23.1	30.6	24.9
would pay none of the cost	46.7	39.1	48.7	44.9	41.2	44.4	40.5	42.2	47.4	50.9	46.0	45.5
do not know	2.7	9.5	12.6	8.4	13.6	9.3	11.3	9.7	12.9	14.3	10.4	14.8

Source: Lifelong learning questionnaire included in standard Eurobarometer, wave 59.0, 2003.

Life-course research also shows the 'wider' or non-material benefits of education and training, especially regarding health, life expectancy, personal development and family formation. Increased awareness of these cross impacts between different social spheres – education/training, health, family, crime, etc. – would be a powerful political tool for integrating different 'competing' policy areas (and budgets).

⁽¹³⁾ See for example Bukodi and Roberts (2002); Noguera et al. (2002); Becker and Schömann (1999).

Regarding the cumulative effects of initial and continuing VET throughout the life-course, some important policy recommendations are:

- (a) to remove the obstacles to participation in education and training of socially deprived groups and to foster actively the participation of the lower qualified in continuing education programmes;
- (b) to ease transitions between education and training, initial education/ training and higher education and between education/training and work;
- (c) to facilitate vertical and horizontal mobility in a system of lifelong learning.

Assessing the impact of education and training over time is highly relevant in view of the rapid economic advancement of most European countries in recent years. This, however, requires increased efforts to set up comparable longitudinal data sets across all European countries.

Bibliographical references

- Asplund, R. *Public funding and returns on education*. Ely: PJB Associates, 2003 (New perspectives on learning – Briefing paper 29).
- Ballot, G. *Firms investment in human capital: sponsoring and effect on performance*. Paper submitted to the European conference on 'The future of work: challenges for the European Employment Strategy', held in Athens, 13 and 14 February 2003.
- Bassi, L.J. et al. *Human capital investments and firm performance*. Washington: Human Capital Dynamics, 2001 (Working paper).
- Baumgartl, B.; Strietska-Illina, O.; Schaumberger, G. Consultancy for free? Evaluation practice and culture in the European Union and eastern Europe. Findings from selected EU programmes. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- Beywl, W.; Speer, S. Developing standards to evaluate vocational education and training programmes. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).
- Cedefop. *Lifelong learning: citizens' views*. Luxembourg: EUR-OP, 2003.
- Chisholm L.; Larson A.; Mossoux A-F. *Lifelong learning: citizens' views in close-up*. Luxembourg: EUR-OP, 2004.
- Coleman, J. *Foundations of social theory*. Cambridge, Mass: Harvard University Presse, 1990.
- Coles, M. Evaluating the impact of reforms of vocational education and training: examples of practice. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- Council of the European Union. *Detailed work programme on the follow-up of the objectives of education and training systems in Europe*. *Official Journal of the European Communities*, 14 June 2002, C142, p. 1-22.
- de la Fuente, A.; Ciccone, A. *Human capital in a global and knowledge-based economy*. Final report. 2002.

- Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004c (Cedefop Reference series, 57).
- Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004a (Cedefop Reference series, 54).
- Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004b (Cedefop Reference series, 58).
- Descy, P.; Tessaring, M. (eds) *Training in Europe. Second report on vocational training in Europe 2000: background report*. Luxembourg: EUR-OP, 2001b (Cedefop Reference series, 3008).
- Descy, P.; Tessaring, M. *The value of learning – Evaluation and impact of education and training. Third report on vocational training research in Europe: synthesis report*. Luxembourg: EUR-OP, 2005 (Cedefop Reference series, 61).
- Descy, P.; Tessaring, M. *Training and learning for competence. Second report on vocational education and training research in Europe: synthesis report*. Luxembourg: EUR-OP, 2001a (Cedefop Reference series, 6).
- Dunning, E. Towards a sociological understanding of football hooliganism as a world phenomenon. *European Journal on Criminal Policy and Research*, June 2000, Vol. 8, No 2, p. 141-162.
- Eisner, M. Modernisation, self-control and lethal violence: the long-term dynamics of European homicide rates in theoretical perspective. *British Journal of Criminology*, 2001, Vol. 41, No 4, p. 618-638.
- European Commission. *Communication from the Commission – European benchmarks in education and training: follow-up to the Lisbon European Council*. Luxembourg: EUR-OP, 2002 (COM(2002) 629 final).
- European Commission. *The concrete future objectives of the education systems*. Report from the Commission. Luxembourg: EUR-OP, 2001 (COM(2001) 59 final).
- Eurostat. *Public expenditure on labour market policies in 1999 varied greatly among Member States*. Luxembourg: EUR-OP, 2002 (Statistics in focus: population and social conditions, 12/2002).
- Fay, R.G. *Enhancing the effectiveness of active labour market policies: evidence from programme evaluations in OECD countries*. Paris: OECD, 1996 (Labour market and social policy occasional papers, No 18).
- Furubo, J.E.; Rist, R.C.; Sandahl, R. (eds) *International atlas of evaluation*. London: Transaction Publishers, 2002.

- Green, A.; Preston, J.; Malmberg, L.-E. Non-material benefits of education, training and skills at a macro level. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).
- Grubb, W.N.; Ryan, P. *The roles of evaluation for vocational education and training: plain talk in the field of dreams*. Geneva: ILO – International Labour Organisation, 1999.
- Halman, L. Variations in tolerance levels in Europe: evidence from the Eurobarometers and European values study. *European Journal on Criminal Policy and Research*, 1994, 2-3, p. 15-38.
- Hansson, B.; Johanson, U.; Leitner, K.H. The impact of human capital and human capital investments on company performance: evidence from literature and European survey results. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).
- Healy, T. Investing in human capital – the OECD view. In Weiß, M.; Weishaupt, H. (eds) *Bildungsökonomie und neue Steuerung*. Frankfurt/ M.: Peter Lang, 2000, p. 19-29.
- Heise, M.; Meyer, W. The benefits of education, training and skills from an individual life-course perspective with a particular focus on life-course and biographical research. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).
- Hellwig, W. et al. Evaluation of EU and international programmes and initiatives promoting mobility: selected case studies. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- Hujer, R.; Caliendo, M.; Radić, D. Methods and limitations of evaluation and impact research. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004a (Cedefop Reference series, 58).
- Hujer, R.; Caliendo, M.; Zeiss, C. Macroeconometric evaluation of active labour-market policy – a case study for Germany. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on*

- vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004b (Cedefop Reference series, 54).
- Izushi, H.; Huggins, R. Empirical analysis of human capital development and economic growth in European regions. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).
- Kelly, M. Inequality and crime. *Review of Economics and Statistics*, 2000, Vol. 82, No 4, p. 530-539.
- Knack, S.; Keefer, P. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics*, 1997, Vol. 112, No 4, p. 1251-1288.
- Lee, M.R. Concentrated poverty, race and homicide. *Sociological Quarterly*, 2000, Vol. 41, Issue 2, p. 189-206.
- Martin, J.P.; Grubb, D. *What works and for whom: a review of OECD countries' experiences with active labour market policies*. Uppsala: IFAU – Institute for labour market policy evaluation, 2001 (Working paper series, No 2001:14).
- Molsosa, J. Evaluation activities in the European Commission. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- Nieuwenhuis, L.; Shapiro, H. Evaluating systems reform in vocational education and training: learning from Danish and Dutch cases. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- OECD. *Education at a glance*. Paris: OECD, 2003.
- Patton, M.Q. *Qualitative research and evaluation methods*. Third edition. London: Sage, 2002.
- Plewis, I.; Preston, J. *Evaluating the benefits of lifelong learning: a framework*. London: Institute of Education, 2001 (The wider benefits of learning papers, No 2).
- Pont, B.; Werquin, P. Look, listen and learn: an international evaluation of adult learning. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).
- Rychen, D.S. An overarching conceptual framework for assessing key competences in an international context. Lessons from an inter-

- disciplinary and policy-oriented approach. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).
- Schmid, G.; O'Reilly, J.; Schömann, K. (eds) *International handbook of labour market policy and evaluation*. London: Edward Elgar, 1996.
- Scriven, M. *Evaluation thesaurus*. Fourth edition. London: Sage, 1991.
- Stern, E. Philosophies and types of evaluation research. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).
- Stockmann, R. Evaluation in Deutschland. In Stockmann, R. (ed.) *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Opladen: Leske & Budrich, 2000, p. 11–40 (Sozialwissenschaftliche Evaluationsforschung No 01).
- Straka, G.A. Measurement and evaluation of competence. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).
- Struyven, L.; Steurs, G. Quasi-market reforms in employment and training services: first experiences and evaluation results. In Descy, P.; Tessaring, M. (eds) *Evaluation of systems and programmes. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 57).
- Temple, J. *Education and economic growth*. Paper presented at the HM Treasury seminar on economic growth and government policy, 12 October 2000.
- Tessaring, M. (ed.) *Vocational education and training – the European research field. Background report*. Luxembourg: EUR-OP, 1998 (Cedefop Reference document).
- Tessaring, M. *Achieving the Lisbon goals: implications for human resources investment in education and training and economic development*. Thessaloniki: Cedefop, 2003 (Working paper).
- Tessaring, M. *Training for a changing society: a report on current vocational education and training research in Europe*. Second edition. Luxembourg: EUR-OP, 1999 (Cedefop Reference series).
- Viertel, E. et al. From project to policy evaluation in vocational education and training – possible concepts and tools. Evidence from countries in transition. In Descy, P.; Tessaring, M. (eds) *The foundations of evaluation and impact research. Third report on vocational training research in*

Europe: background report. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 58).

Walsh, K.; Parsons, D.J. Active policies and measures: impact on integration and reintegration in the labour market and social life. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).

Watts, M. Aggressive youth cultures and hate crime: skinheads and xenophobic youth in Germany. *American Behavioral Scientist*, 2001, Vol. 45, No 4, p. 600-615.

Wilson, R.A.; Briscoe, G. The impact of human capital on economic growth: a review. In Descy, P.; Tessaring, M. (eds) *Impact of education and training. Third report on vocational training research in Europe: background report*. Luxembourg: EUR-OP, 2004 (Cedefop Reference series, 54).

Wolf, A. *Does education matter? Myths about education and economic growth*. London: Penguin, 2002.

Cedefop (European Centre for the Development of Vocational Training)

The value of learning

Evaluation and impact of education and training

Third report on vocational
training research in Europe:
executive summary

Pascaline Descy
Manfred Tessaring

Luxembourg:
Office for Official Publications of the European Communities

2005 – VI, 48 pp. – 17.5 x 25 cm

ISBN 92-896-0342-9

Cat. No: TI-64-04-272-EN-C

Free of charge – 4042 EN –

The value of learning Evaluation and impact of education and training

**Third report on vocational
training research in Europe**
Executive summary



European Centre for the
Development of Vocational Training

Europe 123, GR-570 01 Thessaloniki (Pylea)
Postal address: PO Box 22427, GR-551 02 Thessaloniki
Tel. (30) 23 10 49 01 11, Fax (30) 23 10 49 00 20
E-mail: info@cedefop.eu.int
Homepage: www.cedefop.eu.int
Interactive website: www.trainingvillage.gr

Free of charge – On request from Cedefop

4042 EN

