



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Assessing Learning

Andreasen, Brian Kjær; Rasmussen, Palle; Rømer, Thomas Aastrup

Publication date:
2001

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Andreasen, B. K., Rasmussen, P., & Rømer, T. A. (2001). *Assessing Learning*. Institut for Uddannelse, Læring og Filosofi, Aalborg Universitet. VCL-serien No. 23

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Assessing Learning

by

**Brian K. Andreasen, Palle Rasmussen
and Thomas Aastrup Rømer**

**Centre for the Interdisciplinary Study of Learning
Aalborg University
2001**

Table of Contents

	Page
Making Sense of Measurement!	3
The Impact of Quality Assessment on Education	17
Apprenticeship and assessment	24

Making Sense of Measurement!

On the comprehension of quality measurements in the educational system

Brian Kjær Andreasen
M.Sc. (Social Science), Ph.D. Student
Centre for the Interdisciplinary Study of Learning
Aalborg University, Denmark
bka@vcl.auc.dk

Abstract:

In this paper I will try to establish an understanding of the different stages and dimensions of evaluation along with the complex educational context in which evaluation activities occur. Within this framework I will discuss the concept of education as a framework for organised development and change and finally discuss the practicability, organisation, logic and practice of evaluation. The present paper is the result of my own reflections on the nature of evaluation rather than a final clarification of evaluation methodical problems it represents a personal wish for clarification of a number of these essential questions. I am especially obliged to Lone Krogh (The Centre for Teaching and Learning, Aalborg University) for having contributed to this paper with constructive criticism.

Introduction

Evaluation with different designs and various purposes form part of the everyday life in every Danish educational institution. In line with an increasing debate on the quality of teaching and learning, evaluation has gained a footing in the everyday life of teachers as well as educational planners and developers. However, quality evaluation in relation to development, planning and implementation of educational activities is by no means a new phenomenon but has existed within an educational pedagogical/didactic context for years. We have, however, throughout the last decade seen an increasing focus on evaluation as tool to control development processes. In my opinion these processes are not easily controlled. In 1998 in the editorial of the Danish Pedagogical Journal the following was written about “The Great Wave of Evaluation”:

“The Great Wave of Evaluation was expected. It marks the shift from the Great Wave of Objectives, in which all employees were engaged in creating forward-looking objectives and plans of action, to the present where the tendency is that the individual institution can implement at random in return for being given grants and admissions for making it through an evaluation successfully.” (Dansk Pædagogisk Tidsskrift, 1998).

This quotation presents a somewhat critical reflection on the instrumental way in which evaluation is often used in relation to assessments of the continuing existence of projects, rather than using evaluation as a tool to identify areas in need of adjustment, in order to meet the more or less well defined objectives of a project.

On a governmental level there seems to be a general agreement that evaluation has positive potential as a tool to control certain educational aspects. The report “Quality in the Education System” (Undervisningsministeriet, 1998) from the summer of 1998 emphasises the need for evaluation in each section of the educational system. The establishment of a governmental institution for evaluation of the educational system is also a significant manifestation of this agreement.

It is my experience that the implementation of an evaluation can be quite time-consuming and that sometimes in the process you ask yourself the question “what is the use of this?” It is my assumption, and therefore also the basic idea of this paper, that evaluation handled with care represents an important learning and control potential to an educational institution. Handled with care and thoroughness evaluation can provide access to important information and perhaps contribute to the establishment of a teaching culture in which quality work forms a natural and integrated part of the ordinary practice of teaching and educating. On the other hand, I also believe that evaluation handled carelessly, as for example short-term contentment surveys, may lead to idle rituals of legitimisation which have no particular influence on the quality of teaching or on the students’ profit.

When using evaluation the applied focus is on the different stages of the learning process in an educational context. In this paper I will attempt to establish an understanding of the complexity and consistency of these stages, and establish a broader understanding of education as a framework for organised development and change¹ - an understanding where the use of evaluation represents

¹ All learning is an expression of change and I understand teaching as an organised change with (more or less) precise statement of aims.

development and change oriented judgements of this organised development. In this connection I will discuss the practicability, organisation, scientific logic and the practice of evaluation².

A model for understanding education

As mentioned, the concept of evaluation forms at present a central part of the debate on educational policy. In order to understand this centrality I find it necessary to define a general understanding of concepts such as education, requirements, society and systems³.

Broadly, it can be said that the overall objective of education is social development; of knowledge, skills and attitudes for the individual, the community and for society.

"The object of education is social improvement. Education is really needed for the purpose of making better citizens. This is practically the same thing as the higher end, social progress, which we saw to be the condition to increase human happiness. If education cannot accomplish this end, it is worth nothing." (Ward in Lundgren, 1978, p. 2).

In the above quotation Ward indicates the *mutual connection* between the long-term and developmental perspectives of education and the fulfilment of more individual, personal emancipatory needs.

A different and more critical perspective on education is characterised by a fixed focus on the *stratifying effects* of education. Education is, both for the state and the individual, an instrument for the achievement of economic and social status – and thus also a powerful incentive for development. In this connection education can be perceived as an instrument for making changes in the social structure of society. Thus, in Welfare State's educational system you will find broad political objectives about achievement of equal access for all citizens in the society. Such equality guarantees are attached to the educational system as a compensation for underlying social and economic disparities elsewhere in the society and they serve as an example of well-intentioned attempts to change the socio-economic structures of society. But despite these equality guarantees education can, within the framework of this perspective, be characterised as having a "differentiating" effect, e.g. between those who follows an education and those who are not capable (or willing) to do this. Differentiation also occurs in relation to the external actors on the educational scene as education is learning with the objective of providing necessary knowledge and qualifications with a view to a following working life. In this context education is a question of qualifications and interests but scarcely the same qualifications and interests that an individual and emancipatory educational perspective would emphasise. The phenomenon of education is thus perceived differently dependent on which perspective one chooses to apply.

Education is also a socialising factor, not only in relation to the following working life, but also as the catalyst, which educates us to live our lives in a certain culture and within a certain societal context. Mannheim (1962, p. 148) brought attention to this relation between education and society:

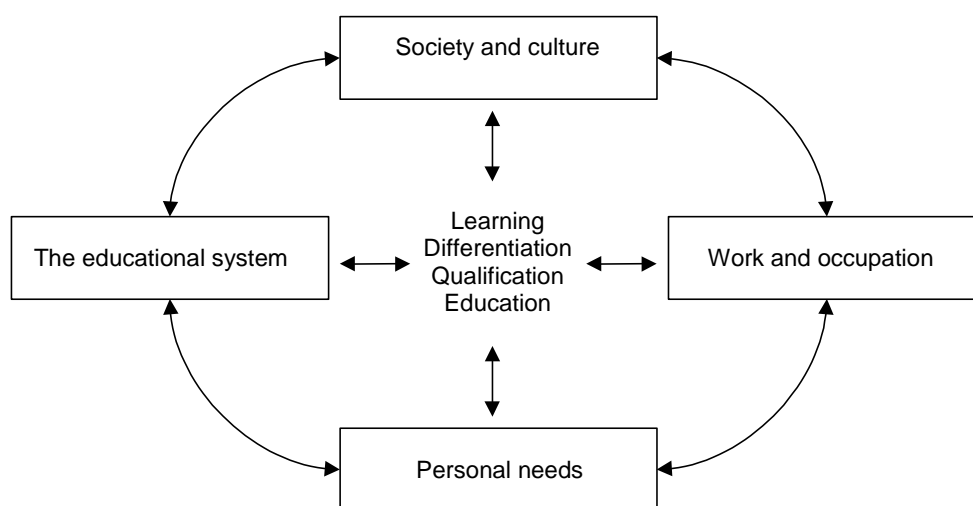
² It is by no means my intention to try to create an overview of the many different methods and applications that exist within this area.

³ Since human development and activity in general takes place within the society, no development can avoid societal influence. The science of the human development must therefore be the science of the human being within the society. Thus, when the focus of this paper is on education as a form of organised development, it must be understood as the continuously changing interplay between human and society.

“One of the most powerful social stimuli is the achievement of status in society. It helps to enhance self-esteem through the gaining of power and the recognition of authority with the privileges attendant upon it. A career is on one hand a social phenomenon and on the other it is a pattern around which many people organize their lives. Money is often a sign of success in this way; school, degree of education, profession – these are all criteria which help to give status in various forms, and here again the educator should be aware of the various resources in society for according prestige and leadership.”

Education must be understood as an interplay between these perspectives, interests and contrasts. But the essential problem is whether education serves a social and individually developing purpose, or whether it contributes to maintain a certain social order, and by that means in reality serves an entirely different purpose and very different interests. In the first case education is a catalyst for development of the individual and perhaps the community, while education in the second case mainly serves the purpose of maintaining the established socio-economic structures.

Figure 1: Actors of Change⁴



The consequence of the above is that the concept of evaluation in an educational relation may be perceived differently and serve different purposes, dependent on which educational perspective one chooses to apply. Deliberate use of evaluation tools can from a positive point of view contribute to an increased focus on the differentiating mechanism, which marginalises certain groups in society and evaluation can therefore be seen as an attempt to minimise the effects of this process. At the same time, this could contribute to an improvement of a qualification process, which is aimed at the following working life.

Somewhat simplified, changes in the educational system can be attributed to at least two different types of requirements; *external requirements* where the justification of the changes is to be found in the societal requirements of education. Changes in e.g. the nature of production will imply changes in the requirements for the qualifications that the educational system is expected to generate. Motives for change can, however, also be found *internally* in the educational system where the

⁴ The figure is based on presentation by Rasmussen, Palle, Centre for the Interdisciplinary Study of Learning, Aalborg University, 1998.

development of knowledge on learning and pedagogic in general has produced new practices. (Ward in Lundgren, 1978, p. 3).

From my point of view however, it seems much more interesting to focus on the interplay between these external and internal motives for change. It is my firm belief that this interplay predominantly appears as a “one way process”, and that the changes in the practice of the educational system are mainly to be attributed to the external requirements. I will, however, not reject the possibility that development can emanate from institutional relations and thus be generated internally, but I will adhere to externalities as the *main* catalyst of change in an educational context. By way of example I may mention the changes which are currently seen in education; attempts to break down traditional boundaries, new educational initiatives and the establishment of new educational structures.

Educational evaluation and the learning processes

With the above reflections in mind, one may claim that concerning most types of education the objective is positive change in the participants’ general or specific level of qualification and competence. It may be reasonably expected of an education that its objective is an improvement of the competencies, abilities, skills or knowledge, which the participants may have had at the beginning of the education⁵. With this perspective in mind, let us contemplate:

- a) **Learning** as a form of change
- b) **Education** as a form of organised change
- c) **Evaluation** as an assessment of the organisation and its influence on the nature of the change

On the basis of this understanding I will divide the organised change (education) into three (analytically) separated phases⁶:

The first phase is characterised by planning and formulation of objectives, estimation of resources, possibilities for attaining the objectives, etc. The second phase relates to the course of events itself, and it is the main objective of the evaluation. Finally there is the third phase, the actual evaluation, which ideally is characterised by gathering the experiences from the two previous phases and evaluating them in relation to each other.

Within this division evaluation takes on the form of a retrospective control tool, which is mainly focused on processes of change that have already taken place – and one may therefore claim that the evaluation expresses a retrospective logic of action.

A change can have objectives at different levels. Certain societal objectives of a given education can thus in general be identified. Likewise, specific institutional objects of a given education may exist (e.g. an increased focus on the institution’s part on environmental issues can have an effect on the institution in the form of prestige). Finally, the individual student will ordinarily have individual

⁵ By the expression general level of qualification I mean not only a positive development of already existing qualifications but also the development of new ones.

⁶ An important phase, which is not included in the above division, is the phase of operationalisation. This phase concerns the actual statement of objectives, and should be clarified before the evaluation is commenced. The operationalisation of the evaluation may be said to concern to what extent the method of the evaluation corresponds to the objective (control or development). The logical order is thus: objective¹ – operation – evaluation – objective². Objective² is ideally a reflection of objective¹.

and personal objectives when taking an education. The personal objectives may, as illustrated in the following diagram, be understood as an individual utility maximisation in relation to a following working life, or it may have a more emancipatory or developmentally oriented perspective.

The levels of change

Societal objectives	Qualification needs, differentiation or socialisation mechanisms, control or developmental objectives.
Institutional objectives	Prestige, control or developmental objectives.
Personal objectives	Prestige, emancipatory needs, social development.

Evaluation relates to one or more of these levels of change and includes one or more of the analytical phases described previously. Ideally, it is however, the objective of an evaluation to bridge the stated objectives and the planning, which has taken place prior to the actual course of events.

The point is, however, that evaluation much too often concerns the fact that a change *has* taken place and attempt to isolate the nature of this change. It is only on rare occasions that the evaluation phase refers to the first phase (the planning and formulation of objectives), and the actual possibilities for control and development, which lies in the evaluation phase, are replaced by an unreflective “go-stop” logic, which at best can serve as a short-term justification tool⁷, but hardly as a long-term tool of control and analysis. Herein lies also the prospective aspect of evaluation, as where evaluation in many cases is said to express a retrospective logic of action, evaluation in this context can have a more quality ensuring or quality improving objective, in which focus (on the basis of prior change) is directed towards future changes.

In practice, most evaluation of educational development will express weighted proportions. As regards their content, they will reflect balanced weightings of prospective and retrospective logics – and (to a variable extent) include the different levels of objectives. In the following paragraphs I will, among other things, take a closer look at these aspects.

The logic of evaluation and science

An evaluation is, in the broadest sense of the word, a *comparison of characteristics with criteria*. It is a comparison, which has reference to drawing possible conclusions or consequences. An evaluation can be perceived as:

“... the act of interpreting the value (merit, worth, significance) of some activity, decision, program, policy, idea, and so forth. It is an inescapable part of daily life.” (Schwandt, 1999, p. 553).

⁷ One should not underestimate the fact that evaluation often forms the basis of an important societal and institutional resource allocation, and that evaluation can thus also have a legitimising aim in relation to an actual educational practice.

It is my belief that the traditional discursively dominated perception of evaluation and the knowledge an evaluation engenders points in the direction of an objective scientific view. By objective I understand the rationalistic scientific perception of knowledge and production of knowledge, which is based on principles of distance, impartiality and objectivity. This is a discursive interpretation of knowledge, which clearly distinguishes between real and true expert knowledge and subjective layman knowledge that inherently is interpreted as distorted by interests, feelings, personal preferences and untenable assumptions. (Schwandt, 1999, p. 554).

"Expert knowledge is often regarded not simply as a different kind of knowledge but as a superior knowledge." (Schwandt, 1998, p. 554).

The scientific observer/evaluator – independent, objective and at work in his laboratory or in the field equipped with recording instruments – is a cliché typical of our time and culture (Carini, 1979). The predominant view of the observer is indeed that his function is virtually synonymous with distance, independence, objectivity and neutrality. When education is to be evaluated the situation is characterised by the same logic. In order to bring about objective observations it is perceived as a necessity that the observer is an outsider. In the actual educational situation the teacher himself is *too* involved, *too* close to the situation and *too* involved in his own practice to be able to make objective observations.

In this scientific logic the observer is a constant⁸. In principle any observer can be replaced with another without changing the nature of the observation in doing so. In order to live up to scientific requirements about objectivity the observer often assumes a *passive* role and in doing so he prevents himself from direct *participation*; a participation that in the nature of things would compromise the objective scientific status of the observations. Ideally, the purpose of this type of observer role is only to *record* behaviour, objects, incidents and events – generally in relation to standardised procedures. Observations can be coded numerically in order to process and analyse them later – without the risk of compromising data by means of personal involvement and incorrect interpretations contingent on the situation in which the observations were recorded. (Carini, 1979).

Opposite this objectivistic perception of evaluation, knowledge and production of knowledge I will present a more critical perception, which is focused on objectivity, the nature of knowledge and the rationality of the evaluation.

"Is evaluation knowledge ever really interested 'disinterested', 'value-neutral', and unencumbered by ideology or by the perspectives of a given evaluator? Should the evaluator be an impartial judge, a change agent, an advocate (for program change, for justice, for empowerment)? Should an evaluator take the side of the less powerful stakeholders in an evaluation? Should the evaluator be more like a co-researcher or co-participant in an evaluation study and less like an outside expert?". (Schwandt, 1999, p. 555).

I see this critical perception of evaluation as an approach, which emanates from the philosophical hermeneutics and which derives its major arguments and perspectives from the traditions and discourse of qualitative science. At this end of the scale we find e.g. the participatory observation. This form of observation and its underlying scientific logic is different from the objective scientific

⁸ It is also assumed that the object of observation is a constant – and that the nature of it is thus in principle independent of the observer or the observational situation.

form of observation. Participatory observation gives the observer a more active function. The participatory observer forms part of the process, which he is to observe, and he will willingly give up control with the observed variable that the objective scientific observer is trying to maintain. Focus is more on the *production and application* of knowledge than it is on the production of knowledge. And I believe that the application aspect of an evaluation should not be overlooked. If evaluation is a comparison of characteristics with criteria and the objective is to draw possible conclusions or consequences, who decides the criteria? – and what are the conclusions to be used for? Is evaluation and the following increase of knowledge a tool to change the institution and its development activities, or is it merely a method of obtaining detailed information about different aspects of practices, e.g. organised processes of change? In the following paragraph I will discuss these questions more intimately.

Knowledge as “a tool for” or “information about”

It is often assumed that the knowledge produced by an evaluation is directly transmittable and applicable. This instrumental approach to the results of an evaluation presupposes a direct agreement between on one hand the knowledge uncovered by the evaluation and on the other hand reality. Knowledge is *directly* applicable in relation to the actors of the process and can compel improvement within the system. This perception (and application) is practised in situations where the results of an evaluation are reserved for management, and where the evaluation data are used as a tool to control the implementation of managerial strategies.

Alternatively, knowledge can be perceived as information. That is, information, which enhance the actor's and the evaluated individual's *understanding* of strengths and weaknesses (Schwandt, 1999, p. 554). The applicability of the evaluation is in this perspective (contrary to the instrumental perception's prerequisite about direct applicability) to be perceived as an *indirect* applicability, as the primary object of the evaluation is information and *not necessarily* change even though the latter is often a natural consequence of an evaluation.

In the traditional scientific terminology it is thus possible to distinguish between the positive and normative dimensions of an evaluation:

- The positive evaluation, in which the evaluator is characterised by key words such as distance, impartiality and objectivity, and where the scientific ideal is positivistic, objective and value neutral production of knowledge. The evaluation is to produce valid knowledge – irrespective of the following application.
- The normative (and critically self-referential) “meta evaluation”, in which the evaluator to a higher degree is a participant than an observer and where the scientific ideal is critical, subjective and self-reflecting. The evaluator questions the purpose of the evaluation and his role at this level is thus self-reflecting, critical and highly normative.

The logic of evaluation resembles the logic of science. I believe, however, that it is necessary to be very conscious of the nature of the knowledge, which is produced on the basis of this logic. As an evaluator you are not only responsible for the production of valid knowledge but also for a *sensible application* of the produced information and knowledge. This is an aspect, which demands a critical self-reflective attitude of any evaluator.

Evaluation – between development and control

The most important aspect of an evaluation is its objective. You may distinguish between two different forms of evaluation: Formative and summational evaluation.

The formative evaluation assesses ongoing educational activities or projects with a view to continuous securing – or improvements of quality according to the main objective, while summational evaluation to a higher degree is characterised by control in relation to the effects of educational activities or the effectiveness of a project.

The concepts formative and summative are of course analytical abstractions and in practice the evaluation will and should include formative as well as summative elements.

An evaluation does, however, in general have several dimensions, which one should be aware of. Beyond the formative/summative dimension I find it necessary to draw attention to further three:

The dimensions of an evaluation

1. Formative vs. Summative	Development or control
2. Internal vs. External	Internal or external motives
3. Qualitative vs. Quantitative	An operative connection between what the evaluation is to show and how
4. System level vs. Process level	Evaluation "of" or "in" a system

Re. 1

Is the evaluation intended as a tool for control for e.g. management or is the intention of the evaluation continuing development of individual teaching? If the intention of the evaluation is not clear it may be difficult to follow up on its results. In connection with the implementation of the evaluation the management plays an important part. The purpose of the evaluation must be clear to all actors – and it must be clarified whether the evaluation is controlling or developing in nature. If the intention of the evaluation remains unclear it is not realistic to expect improvements of quality. Thus, prior to the commencement of an evaluation it must be clear what consequences the evaluation will imply for as well the actors as for the institution and the management. Otherwise, the individual teacher may feel that his or her teaching is subject to control, which will lead to less trust between the teacher and the management.

Re. 2

Concerning this dimension it is important to be aware of whether an evaluation is initiated as a result of internal efforts or as a result of an external request. It is also important whether the evaluation is carried out by the institution's own employees or if it is carried out by external evaluators. In either case however, an evaluation may easily take on a strictly legitimising character.

Re. 3

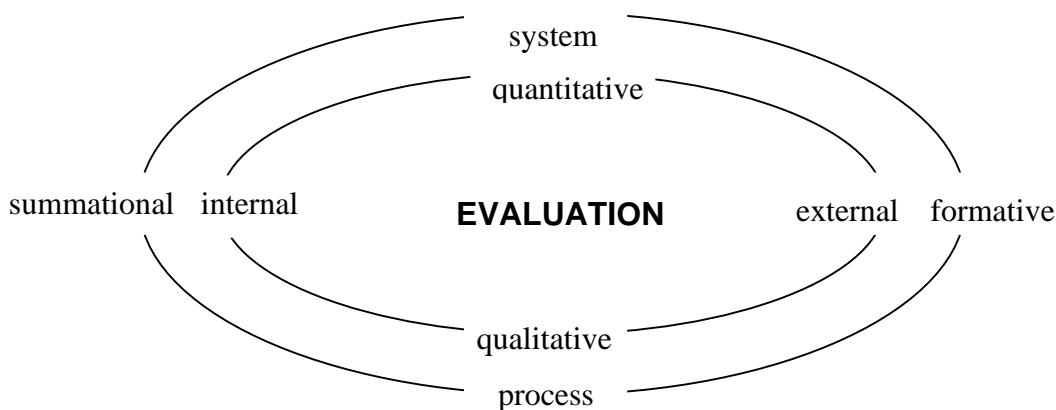
The choice of either qualitative or quantitative methods of evaluation (that is *how* one evaluates) must depend on *what* is to be evaluated and *why*. The choice of method should thus (in the

operationalisation) relate to reflections on which aspects of the change process one wishes to examine – and perhaps correlate and correct. The question of application of qualitative or quantitative methods is a question of whether one wishes a deeper or broader understanding (retrospectively or prospectively) of the viewpoints.

Re. 4

In most of the educational systems some kind of evaluation already exists, e.g. via different forms of tests and examinations. In this context evaluation can be seen as a quality ensuring initiative between the system level and the process level. The system level is the evaluation *of* a system and its ability to carry out qualified education and training, while examinations must be perceived as evaluation *within* a system – that is evaluation of processes of change, e.g. the students' profit. There is, however, no reason why process related material can be used in a system evaluation, e.g. the grades of the students can be correlated with other variables.

Figure 2: The dimensions of evaluation



Evaluation in moderation

As the heading of this paragraph could imply, it is my opinion that evaluations and their results should not automatically and as a matter of course be accepted at face value, since there are, as appears from the above, a series of factors, which supervene in relation to the credibility, validity and applicability of the evaluation. A recurring problem of many evaluations is related to the respondents' abilities to provide useful information. Justly, one can question the basis of experience that e.g. a first-year student at a university may have. When asking these students how the teaching *should* be carried out, only a few will have an adequate frame of reference on which to base their statements. Precisely in this case I believe that an evaluation in its starting point should examine and take due account of expectations and basis of experience.

I am, of course, no patron of the students' opinions being of no value. On the contrary, I believe that they constitute the most valuable source of information in an educational system. But student evaluations should not take the form of public opinion polls. Evaluations must be carried out carefully and in this connection it is important to adhere to the teacher's didactical autonomy in relation to subject matter, implementation and the continued development of teaching. An educational situation is not – and should not be conducted as – a commercial business, in which the

students to a certain extent are perceived as customers, whose immediate needs must be satisfied to the greatest possible extent.

However, this does not contrast with the fact that the students obviously have the right to demand quality education. As a makeweight to this right the students are obliged to contribute to producing this quality in a responsible manner. The ultimate responsibility for establishing conditions to enhance quality in the education rests solely with the educational institution and the collegiate community of teachers.

Another important question is how representative the population of the evaluation should be? Is it e.g. a “fair” evaluation if it only involves the opinions from those students, who live up to their end of the bargain by e.g. preparation? Rather than an aggregation of opinions on the quality of the education irrespective of the students’ effort, activity or preparation, an alternative form of response from the students could be that selected groups in some form or another provide the teacher with feedback. The advantage of this is a raised awareness – in the situation – of e.g. quality aspects of the education. Such an evaluation can hardly be characterised as being representative or objective, but it will without doubt uncover aspects of the education, which the “blind” and more representative evaluation has no possibility of uncovering.

Evaluation of practice – in practice

The discursively dominating belief in the distanced and disassociating production of knowledge as superior in comparison with the subjective and the practice anchored “everyday knowledge” can seem paradoxical. Especially considering how often we all in our everyday life make evaluative judgements of objects, situations and persons we meet. Everyday life is full of these kinds of judgements – a big part of them though, remain unreflected at a conscious immediate and explicate level: “Which apple tastes best?” or “How do I get home as fast as possible?” etc. To the individual teacher most of the activities of the everyday life are anchored in choices between different action alternatives rooted in more or less conscious didactical reflections. “How do I explain this or that in the best way?” “How do I keep the students interested?” etc. In the broadest sense the individual teacher’s or the teacher college’s practice must be viewed as an accumulation of these knowledge rooted experiences and judgements. (Schwandt, 1998, pp. 556-558).

The essential question is then, how to make this knowledge and practice subject to conscious and evaluative reflections?

I believe that the answer to this question is to be found in “theory”. Not theory in the sense that a teacher must evaluate his performance according to certain theoretical models, but theory in the sense that the introduction of normative objectives will turn the evaluating self-reflection into an implicit part of practice. *Evaluation is thus, when practice is reflected in relation to theory.*

The point of departure is the practice knowledge of the individual teacher/teacher college – and not an external, objective and “superior” assessment of practice. (Schwandt, 1998, p. 556-558). The questions will thus remain at the conscious level: “How do I explain this or that in the best way, when I am aware that learning is rooted in the prior experiences of the individual learning actor?” or: “How do I keep the students interested, when I am conscious of their prerequisites of being part of learning and change processes?” etc.

When we evaluate the purpose is to create *meaning*. The underlying motive of any observation and evaluation is a motive of establishing some degree of meaningfulness in relation to the object of the evaluation and its nature. Meaning and meaningfulness are, however, not things or objects, which possess meaningful qualities within themselves. Meaning and meaningfulness are rather derivatives, which receive or obtain meaningfulness by way of a *connection* to other things, persons, events, theories, or the like (Carini, 1979). An observation is therefore only meaningful within the framework in which it is to be understood. The evaluator through his interpretation ascribes meaning to the observation – it is not contained in the isolated observation itself.

”For meaning to arise, there must be recognition. Hence, meaning addresses an underlying unity among persons, things, and the world, which the act of observation can make visible. Observing in the full power of vision is to discover what you recognize in the world, and, in discovering it, to find a part of yourself and your thought mirrored back through the world.” (Carini, 1979, p. 15).

Carini emphasises *recognition* as the basis of the creation of meaning. A recognition, which within the objective scientific discourse, is impeded by the observer’s passive and impartial evaluation. In the ideal objective scientific role the observer is cut off from the unique understanding and meaningfulness, which arises from connection, engagement and participation. (Carini, 1979).

An evaluator is an observer, who stands in the world among others, viewing them, and who maintains the ability to view himself viewing others. This is by no means easy, since such an observer must at the same time be receptive and evocative; embrace things and use them in an image and meaning creating process.

The discussion on evaluation and the role of the observer is ultimately a discussion on the epistemological norms of knowledge and understanding. The question is very simple: How can we know? Descartes, Rorty, Kuhn and Wittgenstein tell us that the world cannot be understood objectively from one privileged point and that understanding is determined by language, the social contexts and the discursively dominating scientific paradigms. Thus, there exists an almost infinite number of possible interpretations of any observed event: All things, objects and events have an infinite number of characteristics but what is then the veracity of the thing, the object or event? Often such understandings may form the basis of a path in the direction of different kinds of value nihilism. One understanding may be as valid and true as the next. However, I do not believe that all understandings are equally valid – even less equally true. The point is that the discussion on subjective versus objective knowledge is not a question of true or false. Objectivity is rather a question of which *methods* one uses to obtain knowledge. Objectivity is to say something about how and why this or that is true and *not that* this or that *is* true. One particular “pair of spectacles” does not necessarily improve the understanding, but the (self)critical consciousness of which pair of spectacles one uses certainly elevates one to a higher level of understanding⁹.

Concerning evaluative knowledge it is my opinion that focus must be on *understanding* in a broader sense – and not on *knowledge* in a more narrow and objectivistic sense of the word. This position is not an expression of a devaluation of the epistemological standards of knowledge but rather a manifestation of the recognition that human knowledge in its nature is subjective and a result of

⁹ Within law, arts, etc. uncertainty and ambiguity often bring about new understandings. Despite this we often criticise the ambiguity as a lack of precision – and as a manifestation of inferior knowledge.

construction and constructive processes. We understand the world by dividing it into different categories but these often cross one another, and it is by virtue of such cross-references that we obtain a more *detailed* understanding of the world surrounding us. The lack of objectivity appears most clearly where the subjectivity of the individual is repudiated.

Evaluation has always (in some form or another) been a part of education and learning and it is probably also the most effective tool for both teachers and administrators in the evaluation of learning – in answering the questions: Is this all right? Do we learn what we are supposed to? Do we get our money's worth? Within the educational systems there are at present, however, a growing recognition of the fact that learning and education cannot be understood out of context. This growing recognition represents a challenge to the many educational systems, as it means that the criteria on which evaluative knowledge is judged, instead of traditional concepts such as validity, reliability, generalisation, etc., must be questions relating to subjective understandings. In the modern educational institution, evaluative knowledge must therefore concern concepts such as progression in understanding, credibility and the connection between different understandings.

The area of evaluation is in relation to theory, method and practices an area that is particularly well described. This area has not only within educational research been the object of quite many reflections. Also within philosophy and psychology such connections, models and theories have been thematized and similar understandings of these areas have been elaborated. It is, however, important to emphasise that the intention of this paper has not been to formulate the final clarification of all problems concerning the validity, reliability and generalisation of evaluation, no more than the paper expresses any final clarification of my own understanding hereof. The development of this paper expresses rather a need for a personal clarification of selected aspects of the nature of evaluation in relation to applicability, organisation and logic. Thus, with these intentions in mind I hope to have contributed to a better understanding of the basic coherence of different aspects.

References

Carini, Patricia F. (1979). *The art of seeing and the visibility of the person*. Grand Forks, N.D: University of North Dakota.

Dansk Pædagogisk Tidsskrift nr. 4 (1998). Kbh.: Det pædagogiske Selskab og Pædagogisk Forum.

Jacobsen, Helge & Wied, Sidsel Rasborg (1996). *Kursusevaluering : Konferencerapport. Kvalitetsudvikling og evaluering af de videregående uddannelser*. Kbh: Evalueringscenteret.

Lundgren, Ulf P. (1978). *Educational Evaluation - A base line for a legitimation of Educational Policy*. - Stockholm: School of education, Dept. of educational research. (Research report / Department of educational research, School of education, Stockholm ; 1978 : 2)

Mannheim, Karl & Steward, W.A.C. (1962). *An Introduction to the Sociology of Education*. London : Routledge & Kegan Paul.

Schwandt, Thomas A. (1999). How should we think about evaluation. In: Weber, Kirsten (Ed.): *Lifelong Learning and Experience, Vol. 2. Papers and Perspectives of the 1st International Summer School of the Graduate School in Lifelong Learning. 2 bd*. Roskilde: Roskilde University and The Danish Research Academy. (Skriftserie fra Erhverv- og voksenuddannelsesgruppen ; 109)

Undervisningsministeriet m.fl. (1998). *Kvalitet i uddannelsessystemet*. Kbh.: Undervisningsministeriet.

Key Words:

Evaluation; development; control; differentiation; actors; reflection; objectivity; understanding; observation; practice.

The Impact of Quality Assessment on Education

Palle Rasmussen, Director, professor
Centre for the Interdisciplinary Study of Learning
Aalborg University, Denmark
palleras@vcl.auc.dk

During the last 10-15 years quality and assessment have become keywords in Danish educational policy, frequently encountered in official documents and debates. It was the then Minister of Education, Mr. Bertel Haarder, who in the mid-eighties ventured to establish quality as the main theme in education through a wave of debates, commissions and development projects. But today, almost 10 years after the collapse of the conservative-liberal coalition government of which Haarder was a part it seems that the concept of quality continues to have a stable place in educational policy. Quality is still the focus of many important policy documents, such as the sections on education in the government's "Welfare commission" (Ministry of Education 1995) and the later report on the educational system done together by the Ministries of Finance and Education (Ministries of Education and Finance and Education 1998). The triumph of the concept of quality has been accompanied by the growth and spreading of quality assessment activities in all parts of the educational system. Special institutions and procedures have been created and been made responsible for assessment. This happened first and most visibly in higher education, where the "Evaluation Centre" was established in 1992 (Thune 1997, Hansen 1999). But in all sectors of education far more resources are now being used on assessing the quality of the work of programmes, institutions and staff.

What has all this activity led to, and what will it lead to? That is the question I will discuss in this paper.

Why assessment?

As a political goal the concept of quality has a paradoxical character because it does not lay out any definite course for policy. It does not indicate any overriding objective for the functioning of

educational institutions; they just have to function well. Nevertheless quality has been widely accepted as just such an overriding objective. I see at least two reasons for this. First quality is a powerful “agenda setter”. It is difficult to voice disagreement with statements that quality should be prioritised (like it was difficult to disagree with statements that the public sector should be modern when the Danish government launched this as a political programme in the early eighties). If you try to counter the quality agenda with competing objectives, suspicions immediately arise that you may not be in favour of quality. Secondly the quality objective has had the role of contributing to deflecting the educational debate from other objectives without openly rejecting these. Not least this applies to an objective which had a strong place in policies of the sixties and seventies, that of social equality in education. In countries like the UK and the USA it has been obvious that the theme of quality was launched by the right as an alternative to the theme of equality.

The quality objective thus had a clear political mission, and to some extent it still has. But the pronounced growth in quality projects, evaluation studies and assessments of outcomes in the area of education and learning have a more general background. It reflects a comprehensive transformation of the forms of political governance, with less emphasis on detailed guidance through rules and more emphasis on control through objectives and budgets. It also reflects the fact that governance comes to rely more and more on a “scientific” approach, which does not trust that programmes and rules are carried out just because the politicians decided they should. This approach regards the implementation of policies as a problem area in itself, which means that evaluation studies and assessment of outcomes are given an important role.

These developments have occurred in many countries. In Denmark they have impacted relatively late. It was only towards the end of the seventies that greater parts of the public sector began to use evaluation studies. This is somewhat surprising in view of the fact that the public sector and welfare policies have had a key role in Danish society during most of the post-war period. But the explanation may be that in the sixties and seventies there was more or less consensus around the welfare state, and that evaluation studies are mobilised from the moment when this consensus begins to break down (Albæk & Winther 1990). If there is any truth in this, it also tells us something about the relationship between evaluation and control. It indicates that evaluation is not primarily a means of assessing the effectiveness of a coordinated and rational public policy, but rather a medium in which different actors (like ministries, institutions, professional groups, users) try out their scope for action and make visible their interpretations of given policies.

Assessing education

Many different things - like car engines, programmes for protection of the environment or educational processes - can be made objects of evaluation. Evaluation is a general approach to assessing given products or processes on the basis of given standards and objectives. But there are differences in the nature of products and processes, and the character of assessment should reflect this. In the case of education and learning it is important to remember that the product takes the form of changed human beings. Education is a process through which it is attempted to improve persons' knowledge, understanding, thinking, skills etc. in relationship to certain goals. Even though education is often called on to realise policy objectives at other levels (like preventing racism or improving employment) the realisation of these objectives depends on changing human beings in appropriate ways.

Processes of development and learning in human beings are multiple and complex. This is a well-known fact, but still a fact to bear in mind in the work of assessing education. For it means that

caution should always be exercised in drawing conclusions from factors in the educational environment (like the length of a course, the number of participants or the pedagogical methods used) to elements in the competence achieved (Madsen & Weber 1996). And in assessing the consequences of achieved competence for persons' living conditions one should be even more cautious.

Of course some educational processes are simpler to assess than others. In the case of many "practical" skills it is fairly safe to say that they cannot be developed without a minimum of training activity under instruction. For many work functions in industry it is also the case that some kind of training certificate is a precondition for getting access to jobs (and thus to improving one's competence through experience). This kind of simple and delimited educational tasks, which the Danish labour market training system was created to provide, may be assessed with relatively simple tools and criteria. But as labour market training has grown over the years into a comprehensive system of education offering long-term courses for comprehensive skills, evaluation and quality assessment must take account of the complexity of learning.

Another characteristic of human development processes is that they imply some degree of consciousness. Human beings define goals for their learning, and even when they are part of a course where the objectives, timetables and methods are fixed in detail, human beings develop their own interpretations and their own strategies. In evaluating educational processes and their outcomes it is crucial to achieve a minimum of understanding of the experiences, interpretations and strategies of the participants (Andersen 1996).

Assessing educational outcomes

Seen from the perspective of politicians and central administration the effects of education in working life and the labour market must be a core question in assessment of the quality of education. This certainly applies to vocational education and labour market training, because their stated purpose is to deliver work-related skills in appropriate places at the right time. Studies indicating that certain types of education have a positive or negative effect on employment will generally attract political attention.

As mentioned above I think it is difficult to evaluate external outcomes of specific educational initiatives. In principle it is possible, but it is very demanding. A high level of methodological reflection is needed because it will often be necessary to combine different methods of investigation. Considerable resources for data collection and analysis are also needed, because existing data and models will most often not be sufficiently sensitive to the specific task. Not least time and patience is needed, because many outcomes will not be visible in the short run.

If these demands can be met effects of specific educational initiatives can in principle be traced. But there is no guarantee that results from such studies can provide a basis for educational or labour market policies. Precisely because the interaction between education, learning and action is so complex types of education, which have had certain documented effects in a certain context, will not necessarily have the same effects in other contexts. This is related to a more general feature of the production of scientific knowledge: In the study of man and society it is much more difficult to generate reliable predictions than in the study of nature and technology.

The social science discipline which has associated itself most closely with the ideal of predictive scientific theory is economics. It is no coincidence that attempts to estimate the effects of labour

market training on employment has most often been undertaken by economists (see the survey of the literature in Gregersen & Holek 1996). These studies are characterised by a high degree of formalisation; they include many persons, but few items of information about each person. The information typically include labour market position for a given period, participation in education during the same period and some standard information on the person like gender, nationality, age, marital status, number of children and income. The analyses then try to trace statistically the connections between labour market position and participation in education, controlling for possible effects of other factors.

I have little hope that this kind of analysis will produce important new knowledge about the relationship between education and society, but there is no doubt that they contribute to general knowledge about the structures and processes of the labour market. The problem is, however, that in public and political debate the results from such surveys are often interpreted in a much too clear-cut and concrete way. This may result in the paradoxical situation that model calculations done by economists may provoke questions and debates in parliament about the quality of teaching in labour market training centres, while one training centre manager after the other tell the papers that almost all trainees get jobs after leaving the training centres.

Legitimacy and knowledge/power

In a rationalist conception of decision-making and control evaluation is accorded a specific role. Evaluation is a systematic investigation of the consequences of previous decisions, both as regards their administrative implementation and as regards their broader impact. The basis for evaluation is most often specific issues and programmes, but evaluation work may also be seen as a general contribution to the knowledge base of policies.

However, the knowledge base created through evaluation studies in areas of policy or institutional practise is often of limited consequence for decision-making in these areas. Policy decisions are influenced by many different factors and are always results of power relationships between different actors and interests. Of course evaluation studies provide knowledge, but in many cases they also provide something else which may best be called legitimacy. This means that evaluation studies and their results are used to assure the stakeholders, politicians or the public that decisions are being made on the basis of factual and impartial knowledge, and that policies or initiatives produce results.

In modern societies the use of power requires legitimacy. The kind of power which circulates in the control of public institutions ideally gets its legitimacy from the system of political democracy; but as administrative systems grow in scope and complexity they find it harder to legitimise their day-to-day practises in this way. Also, political democracy itself is increasingly a quasi-market, where competing political positions try to legitimise themselves in the most appealing and credible way. And in modern societies scientific knowledge often has strong credibility and can help secure legitimacy, just as religious concepts helped secure legitimacy in earlier societies.

The very fact that educational programmes and institutions are made the object of quality assessment contributes to the legitimacy of educational policy. Practises are being evaluated, often by impartial experts, and this is documented in comprehensive reports. The public sector is at work. It is only in the few cases where evaluation studies reveal undeniable failures that this basic legitimising effect is lost.

In the same way evaluation studies can provide legitimacy for decision-making at higher administrative levels they can also legitimise interests at lower levels of the educational system. The more quality assessment become part of institutional management, the more it also becomes an arena where different institutions and units act strategically and compete for resources and status. Thus assessment contributes to the constitution of social reality in the institutional setting (cf. Dahler-Larsen & Krogstrup 2001). To give an example: One of the areas of higher education to be assessed by the Danish evaluation centre was the schools of social work. These schools and their study programme were in fact not included in the original plans for the activities of the evaluation centre, but explicitly asked to be assessed. I think an important motive for this was that the schools of social work - which are almost all independent institutions outside the universities - were afraid to be “forgotten” when the recurrent assessments began to focus attention on the university study programmes (Rasmussen 1997).

Quality assessment and evaluation also have another function in present-day society. They reinforce and disseminate predominant form of knowledge/power. The concept of “knowledge/power” was originally developed by Michel Foucault, who called attention to the fact that the development of knowledge in the human and social sciences has historically been linked with normalisation and disciplining of social practises (Foucault 1980). Theories and concepts have been instrumental in shaping and disseminating conceptions of acceptable and normal behaviour in specific social settings, and in developing techniques for disciplining individuals to this behaviour. Seen in this light the main role of the growing body of “evaluation knowledge” may not be to provide decision-makers with valid knowledge about the institutions and practises they are supposed to control, but rather to install and reinforce among practitioners the norms endorsed by the decision-makers.

I will mention briefly two examples of this. One is from the field of university research. A few years ago the quality Danish research in economics and political science was assessed in a study (Ministry of Research 1997) commissioned by the Ministry of Research. The analyses and the conclusions reflected a strongly hierarchical view of scientific publications in which articles published in refereed British or American journals are assumed to have the highest academic quality, while for instance books published in Danish are assumed to be of considerably lower quality. This view is accepted in some circles, but it does not reflect a general opinion in the social science community and it is generally not supported by empirical research on the process of science. My point is that when the evaluation study concludes on this basis, and when the Ministry of Research accepts these conclusions, the hierarchical view of publication quality is laid down as a norm for all university researchers in the field, a norm which they will try (grudgingly, perhaps) to live up to. The other example is from the field of labour market training. During the nineties a number of evaluation studies illustrated that the value of courses organised through the Danish system of labour market training depends very much on the way in which companies use these courses, and that company decisions in the area of training are often ill informed and haphazard. The studies generally recommended that the training centres should establish dialogue with local business on educational planning, and this was taken up in the policy documents of the Ministry of Labour. The idea is attractive and rooted in the basic concept of an active labour market policy; but there was in fact little evidence that training centres and companies were capable of entering into this kind of dialogue. An abstract norm for the behaviour of training institutions was laid down because it conformed with general policy ideas, because other strategies had failed, and because there was no evidence to discredit it. The training centres launched numerous initiatives on company-based educational planning, but the results were often modest. It is only recently that

more substantial and realistic knowledge about the condition for educational planning (cfr. Lassen & Plougmann 1999) has begun to inform the training policy.

Assessment as illumination of practise

I have argued above that the increasing use of quality assessment and evaluation studies in education takes questionable form and has questionable consequences. But quality assessment also opens new possibilities for institutions and professional groups in education. Evaluation studies provide new knowledge about educational practises, and quality assessment often challenges the different partners (like teachers, students, planners, companies) to define their educational tasks, to assess how they live up to these tasks, and to reflect whether they could do better. Quality assessment can help to illuminate existing practises in education, and thus to improve practises.

So the question is not whether quality assessment should be done but how. I offer three general guidelines:

Firstly the relationship between quality assessment and decision-making should be as clear and open as possible. In my opinion this means that assessment schemes should in principle be independent of decisions about resource allocation. If assessment and evaluation become arenas for continuous struggles about funding schools and different groups of staff will develop excessively strategic attitudes and behaviour, and will not exploit the positive possibilities for reflection and improvement of educational practise.

Secondly assessment should not focus too much on evaluating the external effects of education. As mentioned above this task is very demanding both on time and resources. Even in cases where significant results are produced it may not be easy to make use of these in the development of educational programmes. I think quality assessment should focus mainly on the day-to-day practise of educational institutions and programmes. This may be done by trying to answer in a systematic and well-documented way some elementary questions like: What were the aims or objectives of the educational programme? What actually happened, who did what? How did the different groups of participants experience the process and its results?

Thirdly much of the assessment activity should be organised as internal evaluation, i.e. responsibility should be placed in the educational institutions and the staff of these institutions should conduct much of the work involved. In this way the persons who are made objects of evaluation are themselves involved in discussions of assessment criteria and procedures. This no doubt increases the probability that the results of the assessment are in fact used to improve educational practises.

References

Albæk, Erik & Winter, Søren (1990). Evaluation in Denmark: the State of the Art. In: Ray C. Rist (ed.), *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations*. New Brunswick: Transaction Publishers.

Andersen, Anders Siig (1996). The Subjective Meaning of Teaching Content. In: Henning Salling Olesen & Palle Rasmussen (eds.), *Theoretical Issues in Adult Education: Danish research and experiences*. Frederiksberg C.: Roskilde University Press.

Dahler-Larsen, Peter & Krogstrup, Hanne Kathrine (2001). Evalueringers konstitutive virkninger. [The Constitutive Effects of Evaluation]. I: Peter Dahler-Larsen & Hanne Kathrine Krogstrup (red.), *Tendenser i evaluering* [Trends in evaluation]. Odense: Odense Universitets forlag.

Undervisningsministeriet m.fl. (1998). *Kvalitet i uddannelsessystemet* [Quality in the Educational System]. Kbh.: Undervisningsministeriet.

Foucault, Michel (1980). *Power/Knowledge: selected interviews and other writings, 1972-1977*. Edited by Colin Gordon. Brighton: Harvester Press.

Gregersen, Ole & Holek, Lothar (1996). *Arbejdsmarkedssuddannelserne: en vidensopsamling* [Labour Market Training - A Survey of Existing Knowledge]. Kbh.: Dansk Teknologisk Institut.

Foss Hansen, Hanne (1999). Construction of Educational Evaluation in Higher Education: The Case of Denmark. In: Ingemar Fägerlind, İçara Holmesland, Görel Strömqvist (eds.) *Higher Education at the Crossroads: tradition or transformation?* Stockholm: Institute of International Education.

Lassen, Morten & Peter Plougmann (1999). Kompetenceudvikling: et nyt rum for arbejdspladsens konflikter og samarbejde [Developing Competence – A New Setting for Cooperation and Conflict in the Workplace]. *Dansk Pædagogisk Tidsskrift no. 99:3*.

Madsen, Ulla & Kirsten Weber (1996). Evaluation Studies and Qualitative Designs. In: Henning Salling Olesen & Palle Rasmussen (eds.): *Theoretical Issues in Adult Education: Danish research and experiences*. Frederiksberg C.: Roskilde University Press.

Ministry of Research (1997). *International evaluation of social sciences research*. Copenhagen: Ministry of Research Publications.

Rasmussen, Palle (1997). A Danish Approach to Quality in Higher Education: The Case of Aalborg University. In: John Brennan; Peter de Vries and Ruth Williams (eds.) *Standards and Quality in Higher Education*. London: Jessica Kingsley Publishers.

Thune, Christian (1997). The Balance Between Accountability and Improvement: the Danish Experience. In: John Brennan; Peter de Vries and Ruth Williams (eds.) *Standards and Quality in Higher Education*. London: Jessica Kingsley Publishers.

Erhvervsministeriet (red.) & Undervisningsministeriet (1995). *Uddannelse og kvalitet* [Education and Quality]. Kbh.:Erhvervsministeriet. (Velstand og Velfærd, bilag 9).

Key Words:

Education; educational policy; quality; assessment.

Apprenticeship and assessment

Thomas Aastrup Rømer, Ph.D. Student
Centre for the Interdisciplinary Study of Learning
Aalborg University, Denmark
taar@vcl.auc.dk

Within the last decade we have seen an extensive interest in the concept of situated learning. This interest has mainly been accentuated by the works of Jean Lave and Etienne Wenger (Lave & Wenger 1991, Wenger 1998). The ideas of situated learning have been used to put forward rather stimulating practical and theoretical ideas in a plurality of empirical settings – from primary schools to the educations of smiths – of how to understand and design the learning proces.

However, there is one field of interest that has not yet been focus for a deeper discussion - namely the assesment of the learning which is and which is not taking place. This shortage seems important because assessment procedures are taking place constantly in a variety of situations; in formal settings like external and internal tests as well as in connection with general instruction, presentations, discussions and so forth. My point of interest is therefore to outline what seems to be Lave and Wengers explicit and implicit positions on this subject.

In order to discuss assessment it is necessary to mention an implicit distinction. On the one hand we will speak of learning or knowledge as the object of assessment. I will have to make some remarks on the ideas of the theory of situated learning on this matter. On the other hand we need to comment on the concept of assessment. In the good old days discussions of assessment were primarily a discussion of method. It was simple in the sense that the criteria of assessment were given a priority, and the assessment was about a more or less mechanical application of criteria on the object of assessment¹. This – rather safe and bureaucratic situation - has been altered due to matters like decentralization of the curriculum and a general epistemological relativism. We should therefore put into question, how different theories of learning, including the theories here mentioned, are dressed to face this theoretical and empirical complexity.

The theory of Lave and Wenger is grounded on an interpretation of apprenticeship learning, but in actual fact they have created a general theory about learning and the conditions hereof. That is also – I guess – the reason why they have disregarded the vocabulary of "master" and "pupil" in favour of the concepts of "full" and "peripheral" participation. This means that the relation of practice is not between individuals, but between different moments of access to participation in a community of practice. "To learn" is therefore to participate more competently or more fully in the communities of practice offered by a particular profession. To understand learning this way Lave and Wenger must deny any concept of learning taking place as an internalization, whatever internalization is understood according to the principles of activity theory, whereupon the individual is – so to speak – taking over structures of communication in the social world, or according to principles of transmission, whereupon the individual is receiving and processing information coming from the classroom desk. The most basic point of Lave and Wenger is that learning is always taking place in a community of practice with transparency between full and peripheral participation and the possibility to move between different points in the community towards a more full identity of a particular profession.

School learning is a subject for a critical look from the perspective of situated learning. The main reason for this is that the pupils are participating purely as pupils, which means that all you learn in schools is to go to school. The moments of full participation in schools are occupied by teachers, and pupils are – by definition – unable to become teachers, which again means that it is impossible for pupils to become full participants. In this sense the school is an odd community of practice, breaking all principles of learning. The general solution to these kind of problems has been to understand the school as a community of inquiring persons (Bruner 1998, Barab & Duffy 2000) or to work with the idea of didactic design, where the school is understood as an organizer of movements within and between communities of practice (Wenger 1998).

The consequence of this is – to be strict – that we cannot speak about assessment of individual knowledge, because knowledge cannot be located inside the head of the individual, but is distributed across the whole community of practice. Traditional forms of evaluation, in which the assessment is taking place in laboratorylike situations disconnected from the community of practice, will never be able to demonstrate knowledge, according to how knowledge is understood by Lave and Wenger. This conclusion is somewhat depressing, partly because relations of apprenticeships are imbued by assessment and partly because this conclusion prohibits us from producing an interpretation of the multiple judgements that are actually taking place within the educational system. It does not seem surprising then to realize, that we will find an increasing amount of references to the individual (Dreier 1999) if we take a look at some of the more recent developments of the theories of situated learning. The same tendency can be witnessed in some of the empirical work by Lave and Wenger and others. An example of this can be seen in the following quotation from a smithapprentice commenting on the assessments of the smiths:

"Then they look at it and say: "Oh, it looks good". Well in that case I've done my job. That is the way we figure out, what is considered proper". (Elmholt & Winsløw 1999, p. 109).

Having included these reservations we might ask about the characteristics of the assessments taking place in communities of practice? Well, these assessments signify a relation between full and peripheral participation. I believe there are two interpretations to this answer - a traditional and a modern one. According to *the traditional interpretation of full participation* there is an idea of a center, i.e. is one particular idea of full participation, towards which other forms of participation is

directed. The traditional interpretation are based on quotations from for instance Jean Laves research among Yucatec midwives:

"Girls in such families, without being identified as apprentice midwives, absorb the essence of midwifery practice as well as specific knowledge about many procedures, simply in the process of growing up" (Lave & Wenger 1991, p. 68).

The traditional interpretation can also be based on Laves idea of the function of language, which is considered to be of only minor importance, for instance:

" .. learning to become a legitimate participant in a community involves learning how to talk (and be silent) in matter of full participants". (Lave & Wenger 1991, p. 105).

The common trait in these two quotations is the idea about the right or the proper way of speaking and acting. The pupil must learn to find his bearings within an essence. Looking at the more well defined tradition of assessment theory, we can also find formulations in accordance with the traditional interpretation just mentioned:

"The teacher must possess, first, a concept of quality appropriate to the task and the student group; second, an ability to judge the student work in relation to that concept and a ..." (Morgan & Wyatt-Smith 2000, pp. 132-133).

Finally, the criteria of the assessor are found in a community of knowers:

" ..guild knowledge refers to knowledge that teachers acquire as their history of qualitative judgements developes over time and they collaborate, as a community of knowers, in making such assessments of the students' products." (Ibid., p. 134).

and the criteria are tacit and intuitively grasped:

"..guild knowledge exists largely in unarticulated form as "lore" ". (Ibid., p. 134).

The traditional interpretation emphasizes the singular and unequivocal character of the concept of full participation and stresses the idea of a community of knowers as a condition for making judgments about student knowledge. I will allow myself to combine the different ideas to the conclusion, that full participation is an essence of participation and a particular way of using language. This essence forms the basis of the different kind of assessments, that the student by and large has to accomodate to.

There is another interpretation of full participation that in my judgment can be extracted from Lave and Wengers books. I will call it the *modern interpretation*. The basis of this can be derived from:

"There is no place in a community of practice designated "the periphery", and most emphatically it has no single core or center". (Lave & Wenger 1991, p. 110).

Here, apparently, we should understand the community of practice as consisting of different and multiple "full participations", while none of these should be regarded as more full than others. An

exemplification of what this might mean to the pupil can be seen in this quotation from the research on smith-apprentices already mentioned:

"Peter has three masters, and they evaluate differently. One of the masters stresses the importance of the density of welding while others emphasize the beauty of the work done." (Elmholt & Winsløw 1999, p. 110).

Of course the idea of assessment seems to be complicated by this, as a particular judgement cannot be legitimized purely by intuition anymore. The question is how we can describe the relations between the different criteria (or places for full participation), that are being used in this modern interpretation? Lave and Wenger are not particularly explicit on this matter, but I believe that they end up by focusing on a process, the result of which is being constituted by negotiations and – in the last resort – by struggles of power between different positions within and between communities of practice, a situation that might be described as an "articulation of hegemonies" in the language of discourse analysis (Laclau & Mouffe 1985). The concept of negotiation is emphasized in:

"Participation is always based on situated negotiation and renegotiation of meaning in the world". (Lave & Wenger 1991, p. 51).

and the concept of power is being activated in a statement like:

"Each threatens the fulfillment of the others destiny, just as it is essential to it". (ibid., p. 116).

This solution is quite significant, because I suppose that most of the major professions are characterized by a plurality of full participations. Of course these multiple positions must produce a whole range of criteria of what we should consider as being full participation. The decision whether a particular product is good – this is now the position of Lave and Wenger - is considered through the concept of power, a category of sociology. This means that the criteria for good welding (or anything else) are established by the positions who are able to mobilize the greatest amount of strategic and political resources.

In more recent formulations of situated learning (Wenger 1998, Lave 1999) the pluralistic structure of the community and the placement of the individual is far more emphasized. Here we find that the concept of learning is being connected to the idea of "trajectories" according to which the identities are formed together with the articulation of individual projects, zigzagging through the possibilities and structures available in the community of practice. In the early 90'es the possibility of learning was conceptualized through transparency between peripheral and full participation in the community of practice. In the late 90'es we are supposed to understand learning as the possibility of free movements between the different subpractices and positions of the profession. Instead of the concept of community we seem to get the movements of individuals within partly overlapping fields of practice. This focus on plurality is indicated by the title of Wengers book, "communities of practice". We should note that "communities" are in the plural. Another shift is the concept of "reification", which is also introduced by Wenger. The concept refers to the fact that any practice is producing both forms of participation and conceptual "markers" that are meant to describe the world. These markers (for instance the idea of democracy) are being produced and negotiated in different and overlapping communities of practice, and I think we should look at these reifications as Wengers reintroduction of language as an important category.

Well, what does all this mean for the assessment of student work? It is certainly a consequence that we get into trouble if we continue to work with the idea of an intuition on behalf of the assessor. Let me give an example: A student has been interested in anthropology for a number of years. He then decides to become a nurse, and he ends up in a nursing school that is very tied up to the medical perspective on the nursing profession. His own criteria for good nursing have been constituted in the former community of anthropology, and now he clashes with a totally different set of assessment criteria. Suppose we accept the traditional interpretation of full participation; which means assessment understood as grounded in intuition. In this case the student will have to choose whether he will assimilate to the new community or simply leave the school. This seems to be an unacceptable result because of the fact that anthropology actually plays an important role in the reification practices of modern nursing. A strictly traditional practice produces tension if we accept the idea of a plural society. Turning to a modern interpretation of assessment the picture changes in the sense that the plurality of full participations is accepted, but there are no particularly interesting ideas about how to decide on the quality of student work except from the repeated standpoint of negotiations and power.

Accepting these conclusions let us look at some of the problems facing the theory of assessment, and let us look at some of the challenges facing practices influenced by the ideas of situated learning if we want to justify assessments, referring to something more than sociological concepts like power, tradition and strategy. Questions like these are of course too far-reaching for a full discussion here, and I should not pretend to be fully clarified myself, but some rough outline should be presented.

There are three types of implications that I would like to mention. The first implication is simple: If we accept the possibility of judging individual knowledge, we will have to reintroduce a distinction between object and subject. This distinction was eliminated by the early theory of Lave and Wenger, because both object and subject was constituted in a social practice. Here it is reintroduced, because "somebody" wants to say something about the quality of something else (student knowledge). Another way of putting this is by saying, that the causality between teaching and learning, which is being disputed by Lave and Wenger, is reborn in the context of evaluation. An implication of this is that we cannot be content with only studying transparency of the community of practices and by studying trajectories of the individual student, we also cannot be satisfied with only studying didactical designs. We also have to study the principles of assessment. Both the object and the subject of evaluation must be discussed not only regarding how they are related to practice but also as distinct entities.

The second and the third implication of the analysis are attached to the subject of evaluation; i.e. the location of the assessor. When first looking at the traditional interpretation of full participation, it seems as if we ought to put more interest to the conditions for making assessments. I am here thinking of the concept of "community of knowers", which was introduced in one of the earlier quotations. This means that the judgements made in the different contexts of assessments, cannot be made in any strict individuality, but must be made with a reference to a particular community. It is possible for the socially isolated assessor to say sentences like "this is good because it promotes my interest" but there seems to be a particular duty imposed on the assessor who is a part of a community of knowledge to express himself like for instance "this is good because the essay shows knowledge of the history of modern art". In the last example it is not indifferent what the student or what the assessor says, because it refers to a given tradition and its most important texts and the interplay of this with the ability of the student on "how to speak properly". This interpretation of the

concept of a community of knowledge as a knowledge and a handling of the most important texts ought to remind us of the ideas of hermeneutics; i.e. how texts are developed in an interpretive practice (Gadamer 1989, part 2, ch.II). "To speak properly" in a community of practice depends on an idea of "properly" which again is defined by the curriculum and the interpretive practice that surrounds it. We should note the radicality of this. What is being said is that without an idea of a curriculum it is impossible to make assessments of anything but simple and uninteresting statements. So we end up with an idea of "full participation" that is not sociologically founded. "Full participation" in this sense is a concept that covers a particular competent insight in the tradition or the curriculum of a profession or a society. By the words competence and insight I refer to the distinctions that are given credit by the community of knowledge. When I say "given credit" it should not be understood in a sociological sense. Rather we should understand it as derived from moral philosophy; particularly from the tradition of communitarianism (Macintyre 1981, Taylor 1994). An important message in this tradition is that the existence of morality (e.g. "a social worker ought to act like this" or "that is how we ought to work with knowledge") is conditioned by the existence of a community possessing a strong relation to tradition and to particular virtuesⁱⁱ. The conclusion is then: The condition for assessment and evaluation in the traditional interpretation of full participation is the assessors knowledge of and ability to relate to the canonical texts of a tradition and his general integration in a community of virtuesⁱⁱⁱ.

The third implication of the analysis of situated learning also relates to the later versions of the theory. It seems important to ask, what ought to characterize a community of knowledge in a postmodern educational culture, where the students have multiple, different and perhaps even odd trajectories. In situations like this "a curriculum" might even be problematic. The student of nursing colliding with the discourse of medicine, which was mentioned earlier, is a good example. Here the traditional interpretation becomes problematic because an assessment of this student due to the principles of medicine will be incapable of capturing the creativity on behalf of the student. Situations like the above mentioned, which I assume is becoming more and more normal in the educational culture, show that we are not only interested in talking properly, but in talking (and be silent) in new and interesting ways. As a consequence of this I see two lines of thought. First of all the assessor must get used to the idea that the product or the expressions of the student might not be what it appears to be. Another way of putting this is that the object of the assessment does not appear in its own right but through its representation in consciousness. We should conceive of the object as purely constituted by the re-presentation done by the assessor. The decisive criteria for the quality of a students expression of any kind should not be located in the expression itself, but in the assessor and his relation to the community. The second, but related, line of thought is that these distinctions seem to set free the imagination of the assessing community. It does not seem enough anymore to concentrate on the relation to the texts of the tradition. There must also be conversations with other traditions and with different possible re-presentations of the same object. The evaluator must – so to speak – be able to move between a wide range of language games and professional traditions. This should not take us by surprise. Within the last couple of decades we have been witnessing a heavy increase in epistemological interest in the relations between the particular object and the possibility of judging it^{iv}. An instance of such a non-essentialist and non-sociologist – but rather a philosophical – point of departure can be captured in this quotation from Richard Rorty:

"The attempt to edify (ourselves and others) may consist in the hermeneutic activity of making connections between (...) our own discipline and another discipline which seem to pursue incommensurable aims in an incommensurable vocabulary. But it may instead consist in the

"poetic" activity of thinking up such new aims, new words, new disciplines, followed by, so to speak, the inverse of hermeneutics: the attempt to reinterpret our familiar surroundings in the unfamiliar terms of our new inventions". (Rorty 1981, p. 360).

The idea is, that the evaluator should be dissatisfied with only his own vocabulary. He should also be discontent by only looking at other vocabularies through the spectacles of his own. In addition he also ought to look at his own vocabulary through the spectacles of the other. This is the activity which is named "poetic", and this is what opposes the tendency of the traditional community to closure. Metaphorically, we might say, that the condition for assessment in postmodern educational cultures is that the assessor is a "cognitive cosmopolitan".

The theories of situated learning suggest two different interpretations of what we ought to understand by the concept of full participation. A traditional idea, emphasizing the idea of a center or core which is termed full participation, and a modern idea, that stresses the multiple and the zigzagging relations between different kinds of full participation. Both ideas of participation tend to end up in a sociological reductionism, overemphasizing the concepts of power and negotiations. In this article, I have investigated the conditions for producing assessments. I have found, that it is possible to connect the traditional idea of full participation to an idea of a community of knowledge, which can be further described with the help of concepts like tradition, curriculum and virtue. I have also found that it is possible to connect recent theories of judgment and relativism to the modern version of full participation.

This article was originally published in danish in Dansk Pædagogisk Tidsskrift, no. 2, 2001, p. 82-90.

Notes:

ⁱ See e.g. Tyler (1970) and Bloom (1956).

ⁱⁱ The theories of communitarism are strongly influenced by aristotelian ethics. They are generated on the basis of a refusal of the non-naturalism of Stevenson (1940) and of modern liberalism (Nozick 1974, Rawls 1970).

ⁱⁱⁱ It is not at all usual within theories of learning and assessment to stress the connection between knowledge and the possibility of judgement in this way. Within the more general theory of education the perspective is presented in Eisner (1991). Eisner compares the evaluator with a wineconnoisseur, whose ability to relevant perception of the wines quality depends on his earlier experiences and his knowledge about the distinctions of the "wine-community" (e.g. "sweet", "color of grapes").

^{iv} See Bernstein (1983) for an overview of this discussion. See also Ahrendt (1992) for an interpretation of Kant's theory of judgement and Rorty (1981) for an interpretation of hermeneutics.

References:

Arendt, Hannah (1982). *Lectures on Kant's political philosophy*. Chicago: The University of Chicago Press,

Barab, Sasha & Duffy, Thomas (2000). From practice fields to communities of practice. I: David H. Jonassen & Susan Land (ed.), *Theoretical foundations of learning environments*. New Jersey: Lawrence Erlbaum Associates.

Bernstein, Richard J. (1983). *Beyond objectivism and relativism*. Oxford: Basil Blackwell.

Bloom Benjamin S. (1956). *Taxonomy of educational objectives: the classification of educational goals*. New York: David McKay Company.

Bruner, Jerome (1998). *Uddannelseskulturen*. København: Gyldendal.

Eisner, Elliot W. (1991). *The enlightened eye: qualitative inquiry and the enhancement of educational practice*. New Jersey: Prentice-Hall.

Elmholt, Claus & Winsløw, Jan Henrik (1999). Fra lærling til smed. I: Klaus Nielsen & Steinar Kvale (ed.). *Mesterlære: læring som social praksis*. Kbh.: Hans Reitzels Forlag.

Gadamer, Hans-Georg (1989). *Truth and method*. London: Sheed & Ward.

Laclau, Ernesto & Mouffe, Chanlan (1985). *Hegemony & socialist strategy: towards a radical democratic politics*. New York: Verso.

Macintyre, Alisdair (1981). *After virtue: a study in moral theory*. London: Duckworth.

Morgan, Wendy & Wyatt-Smith (2000). Im/proper Accountability: towards a theory of critical literacy and assessment. *Assessment in Education*, vol. 7, (no. 1) 2000, s.123-142.

Nozick, Robert (1974). *Anarchy, State and Utopia*. Oxford: Basil Blackwell.

Lave, Jean & Wenger Etienne (1991). *Situated learning: legitimate peripheral participation*. Cambridge: Cambridge University Press.

Rawls, John (1971). *A theory of justice*. Oxford: Oxford University Press.

Rorty, Richard (1981). *Philosophy and the mirror of nature*. Oxford: Basil Blackwell.

Stevenson, Charles Leslie (1994). *Ethics and language*. New Haven: Yale University Press.

Taylor, Charles (1994). Justice after virtue. In: John Horton & Susan Mendus (ed.), *After Macintyre: critical perspectives on the work of Alasdair MacIntyre*. Cambridge: Polity Press,

Tyler, Ralph W.(1977). *Undervisningsplanlægning*. Kbh.: Christian Ejlers Forlag.

Wenger, Etienne (1998). *Communities of practice: learning, meaning, and identity*. Cambridge: Cambridge University Press. (Learning in doing).

Key Words:

Apprenticeship; assessment; learning; participation; situated learning.