

STRATEGIES FOR QUALITY ASSESSMENT: THE HUMAN FACTOR

por WILLEM VAN OS
Free University of Amsterdam

1. Introduction

I would like to start by showing you four statements, reflecting my opinions about the assessment of the quality of higher education.

1. *It is not easy to assess the quality of education;*
2. *There is no such thing as one ideal way to assess the quality of education;*
3. *Nevertheless, it is necessary to assess the quality of education;*
4. *The usefulness of the attempts depends heavily on the acceptance and cooperation of both teachers and students: the human factor.*

Underlying these statements there is one basic proposition:

The purpose of quality assessment should be the improvement of the quality of higher education.

The choice for one kind of quality control or another must therefore take into account the rate at which such a procedure contributes to that quality improvement; all other reflection about the correctness of a particular way of quality assessment are of minor importance.

In this hour I shall discuss these statements and the basic assumption in more detail, ending up with some conclusions about the practical consequences for the chosen strategy of quality assessment.

Before doing so, it is necessary to stress that this paper is directed only to the quality of *the educational processes*. By this I mean the way in which teachers give their lectures, their testing practices, the amount of feedback given to students, the way in which they stimulate learning, and so on. In other words: «the art of teaching».

This implies I shall not talk about the *content of courses and programmes*, that is to say whether or not that content meets accepted scientific standards. I take that for granted.

2. *No easy way*

Growing interest

The topic of quality assessment is receiving much attention these days. In the Netherlands, for instance, at least five conferences have been devoted to this theme in the last two or three years, and this summer course is another example of the growing interest in the possibilities and problems associated with the subject of educational quality. Moreover, this interest is not restricted to the European countries but extends to the American continent, including South America, and also to Australia.

The beginning of this development lies in the early seventies, but the longest tradition in the search for educational quality assessment can without doubt be found in the United States. Over there as early as the beginning of this century much attention has been given to manners to assess the quality of institutions. This fact has probably to do with the specific American way of life, in which the initiation of activities is always closely bound to the expected effectiveness or profit of these activities, but also with the structure of American higher education. I mean with that structure the way in which universities are founded, and the connected struggle for research grants and students.

Baby-boom

The overwhelming interest arising in the last 5 or 10 years is, after all, not very surprising. In all western countries the decade between let us say 1960-1970 was the period in which the great number of children born just after the second world war, arrived at higher education's doors. In a short time, universities were confronted with lots of students, much more than they were used to, at first without having sufficient means to deal with them. It may be suspected that in that time many teachers and researchers were appointed who did not reach the educational or academic standards needed in all respects. Anyhow, universities were growing very rapidly, became massive institutions costing a lot of money. This was an educational problem first, but very soon it became an economic and political problem too when in the seventies the world economics declined. The question arose how to

manage and pay higher education in a time in which the limits of economic growth were reached.

Retrenchments

In almost all western oriented countries the question was posed if all the financial means, needed for higher education, were rightly spent, soon followed by the question how to lower the expenditures. The second question was far more easier to answer than the first one. In the Netherlands, for instance, the costs per student in 1983 are only 70 % of that in 1975 (Ministry of Education, 1987), reached by a dramatic lowering of university budgets and an even more painful reduction in academic personnel and salaries. Besides, the freedom for students in spending their days on higher education was limited, and the price they have to pay it tenfolded.

Many of these measures were connected, that is to say in a political sense, with expectations that in this way the quality of higher education might raise too, but as you may guess most academics held the opposite view. Anyhow, from both sides—the universities and the governmental administration—the need for some form of quality assessment was felt. The universities in the conviction that in this way the injustice or even the stupidity of the political decisions could be proven, the administration on the other hand in the hope that more specific cutbacks could be possible, meaning for instance the reduction in the budgets of low-quality departments, or even the abolishment of them.

Mutual suspicion versus mutual trust

The laborious discussion about quality assessment in higher education and the difficulty in reaching a consensus about the way to accomplish that, is to a great extent caused by the suspicion of the academic world that it is just another attempt to cut expenses, and the belief of the government or sometimes university boards that faculties and teachers just want to try to cover up their apparent weaknesses.

Any effort to initiate some form of quality assessment has to deal with this mutual suspicion, and any effort is doomed to failure if one does not succeed to replace it by mutual trust. This is one aspect of the human factor in quality assessment, and very difficult to overcome.

The mutual trust must be based on the conviction that quality assessment is and should be directed towards the maintenance or, even better, the improvement of that quality.

This point of view has some implications, especially with respect to the chosen parameters or so called performance indicators as I hope to

show. Before doing so, I would like to expand a little on the various possibilities in the process of quality assessment.

3. *Not one ideal way*

Definition

At first, I must provide you with a definition of the concept performance indicator. Following Cave et al. (1988) I define a performance indicator as «an authoritative measure —usually in quantitative form— of an attribute of the activity of a higher education institution» (p. 20). To this definition may be added, as Cuenin suggests, that performance indicators imply a point of reference and are therefore relative rather than absolute in character.

One may wonder why it should be necessary to rely on parameters, on indicators. Why not use the word quality itself? After all, anyone knows what it means. As a matter of fact, this exactly is the problem. In an earlier article (Drenth et al., 1987) we cited the Dutch scientist De Groot (1983), stating that «the term (...) can operate as a kind of refuge for those for whom open discussion about more concrete matters has become too difficult or unsatisfactory...».

Internal, external, operating indicators

The many approaches to quality assessment or, otherwise, the many different performance indicators suggest that every single, simple measure is insufficient, and most researchers believe that this is indeed the case. The concrete demonstration of quality as a multi-dimensional concept is possible in various ways.

For instance, a distinction is proposed between internal, external and operating performance indicators.

Internal indicators include variables reflecting either inputs into the institution (for instance, attractiveness of courses) or valuations internal to the institution (teaching quality).

External indicators concern the institution's market position: for instance the employment of graduates. External indicators are in other words directed at the output of the educational process.

Operating indicators include those variables, needed to accomplish the goals of the institution: libraries, computer facilities, etc. (Cave et al., p. 19). They are in a way conditional variables, necessary but not

sufficient, and in this paper I shall limit myself to the internal and external performance indicators.

I will not repeat here the numerous examples of indicators; those who are interested will find more details in Cave. Instead, I direct your attention to one indicator in particular, because in many ways it reflects the problems connected with our subject. I am aiming here at the so called value-added measure.

Value-added measure

This approach finds its origin in the idea that the quality of an instruction depends to a very high degree on the quality of students. By the quality of students I mean their aptitudes and attitudes. However, it can be argued that it is much easier to educate bright students than it is to coach less gifted pupils to meet high exam standards (I shall return to this subject in the third section). To use «graduate quality» as a measure for «institution quality» would be denying the effort the institution (and the students of course) puts in.

So, a measure that reflects the *difference in performance* of incoming and outgoing students seems a more fair one. (In this way the value-added approach can be viewed as a combination of an internal (entry level) and external (exit level) indicator.)

Cave et al. cite Astin, who writes that «true quality resides in the institution's ability to affect its students favourably, to make a positive difference in their intellectual and personal development» (p. 58).

The value-added approach seems attractive; it is logically sound and has considerable face-value. However, several weaknesses appear when we explore it further.

Some problems

1. In the first place, some evidence points to a weak or non-existent relationship between entry scores and final degree performance. This probably means that entry scores are a poor measure of student's actual level of attainment. If this is true, much work has to be done in finding reliable and valid measures of student's entry level, *work which is not directly aimed at the institution's true goals but is just in favour of the process of quality assessment.*

2. A similar problem arises when institutions would reorganise their programmes in order to score well in the value-added assessment. The disadvantage is that the procedure itself has *no direct positive benefit* for the students; instead, the results are used solely for the pur-

pose of evaluating institutions. This objection becomes more important when one realizes that considerable costs and efforts are connected to this kind of evaluation practices.

Goals of higher education

Now I come to a main issue of my views regarding quality assessment. This is that *there should be a proper balance between the primary goals of higher education —teaching and research— and the process of quality control*. Above all, neither teachers nor researchers should get the idea that quality assessment is considered (by the government or even the university board) more important than just doing their proper job: teaching and scientific research.

For this reason there should be an optimal relationship between the essentials of the educational process and the specific performance indicators. A correspondence which is clearly visible and easy to understand, the indicators having considerable face value and being fairly easy to measure. In my opinion the success rates of an institution of higher education meets these conditions in high degree. That is not to say that success rates are the only possible indicator, not even an ideal one; there are objections to be made, and they have been made as I will show you later on.

Besides, one has to remember that on performance indicator —including these success rates— can or should be a definite *judgment* about the real educational quality of an institution. At best, it is just what the word implies: an *indication* that something might be right or wrong, no more, no less. I shall continue on this topic in the next hour too.

Now, let me say something about the necessity of quality control and quality improvement.

4. *Necessary to do*

Accepting that the main purpose of quality control is the improvement of higher education, an objection can be made, stating that there is *no real need* for improvement. For instance, someone made a very sharp distinction between bright students and all others. The former group did not really need any education, he said; they succeed even without it. The latter ones, on the contrary, are just not good enough and will not be helped by education, anyhow not enough. Moreover, they *ought not be helped*; it is right if they fail. Their presence only even lowers the quality of the university or faculty.

This point of view is often heard, although mostly in a less extreme form. It is not quite untrue either and may lead to a critical and sometimes necessary and unevitable discussion about the very essentials of higher education. I return to this topic at the end of this section.

Lack of incentive, no training, low success rates

For the moment the question remains if the quality of higher education needs substantial improvement. I repeat that I am not speaking about the content of programmes or courses, but that I restrict myself to the educational process. Three remarks must be made.

1. In a recent review, Moses (1987) described the institutionalisation of educational development in Australia, Britain, the USA, West-Germany and Sweden. With some variation her opinion about the Australian situation may well apply to other countries mentioned, including the Netherlands and may be Spain: «For faculty selection, professional and/or research skills and performance are normally the main criteria, even though university faculty (...) are expected to do both research and teaching. Teaching experience may be an advantage, yet excellence in teaching is hardly ever assessed. For tenure and promotion often a minimum level of teaching competence is expected, while faculty perceive and experience that excellence in research earns promotion» (o.c., p. 450).

In her point of view there is *no incentive* for an outstanding educational performance; on the contrary, it may harm your career to spend much time on teaching instead of doing research.

2.. Secondly: in the Netherlands, and as far as I know in many other European countries, for teaching at the universities *no special training is required* apart from the competence in the subject to be taught. This differs very much from the requirements to become a teacher in for instance primary schools, for which students are trained for as much as three years.

It is also quite remarkable that neither the Ministerial bill of 1985 nor that of 1987, both of which are dedicated to the quality of higher education and the means to improve that quality, mention the lack of any formal training for university teachers.

The results are clearly visible. At my own university, we use standardized questionnaires in which students are requested to rate their teachers with respect to their didactic skills. The outcomes have led us to the conclusions that, firstly, every department or faculty has some bad and sometimes very bad teachers and, secondly, that it is possible to improve their performance substantially, for instance by giving them

a short training in a few basic didactic principles. Moreover: the mere fact that teachers are rated seems to be a motive to do better next time (Van Os, 1989)..

3.. In the third place: the need for improvement in higher education can be derived, at least in my own country, from the rather *disappointing number of students* who pass their examinations within the course time of 4 years or even the maximum allowed study time of 6 years. The next hour will be devoted to this topic.

My conclusion so far is that the quality of the educational processes can be raised. But would it make a difference, for instance with respect to success rates? Some people doubt that.

Good teaching pays

Hofstee for instance, a Dutch psychologist, states that the quality of education depends only to a minor degree on the quality of *teaching* and is much more dependent on the quality of *students*. The latter may be responsible for as much as 90 % of all differences between students, suggesting that the quality of teaching does not really count (Hofstee, 1985).

Even if Hofstee is right, even if the quality of teaching explains only 10 % of the variance between students, good teaching can make a lot of a difference as I shall show you with a hypothetical example.

TABLE 1
BENEFITS OF GOOD TEACHING

	STUDENTS										MEAN	PASS
	A	B	C	D	E	F	G	H	I	J		
GOOD TEACHING	3	3	3	4	4	5	5	6	7	7	4.7	30 %
BAD TEACHING	3	3	4	6	6	7	7	7	7	7	5.7	70 %
AMOUNT OF VARIANCE EXPLAINED: 9.4 %												

In this example both conditions (bad versus good teaching) are responsible for even less than 10 % of all differences between students, but because the performance of students in a very crucial part of the distribution has improved (in the Netherlands the borderline between pass and failure lies at mark 6) good teaching makes a great difference here.

Of course, it is a composed example lacking any empirical base, but the message is clear, I hope: the goal of our educational task should not only be excellence for a happy few, but also the success of those who would fail without help (Van Os, 1988).

Therefore, it is *necessary and useful* to improve the quality of higher education. Now then the fourth section: how to accomplish that.

5. *The human factor*

Free to evaluate

I began this hour with the observation that quality control is receiving much attention the last couple of years (see also Cook, 1988; Lynton, 1988).

However, the implicit suggestion that in former decades evaluation or quality assessment in general have never been topics of importance must be refuted, at least within the Dutch universities: many examples of evaluation studies prove the opposite.

On the other hand, it cannot be denied that the critical investigation of (educational) performance has been restricted to faculties themselves (and within faculties to particular teachers): neither the institution nor the government, played a substantial role in the stimulation or organisation of such activities. In fact, this is a crucial problem of institutional quality control: why do some faculties within some institutions have a long tradition of educational evaluation while others do not, and, to continue along this line: if some faculties within some universities have always been able to (and wanted to) bar the doors to a critical reflection on their performance, under what conditions will they change their position?

Outside pressure

I believe that in many cases some outside pressure (e.g. a governmental policy) is needed to stimulate institutional self-evaluation; that in many cases the same pressure (e.g. an instructional policy) is needed to stimulate faculty self-evaluation; and last but not least, that in many cases the same pressure (e.g. a faculty policy) is needed to stimulate teachers' self-evaluation.

In other words: *the announcement or start of some form of external evaluation, at whatever level, acts as a change agent to accomplish the internal self-evaluation at the following level.*

In the previous section I referred to Moses, stating that something is lacking to achieve an outstanding educational performance: an incentive.

One may wonder why research is, and teaching is not connected with a (be it fallible) reward system. A possible answer may be that the visi-

bility of the outcomes plays a role: scientific research has to result in a paper, an article, a book, something that is documented and publicly available. Traditionally, teaching lacked an output which could be discussed, questioned, approved or disapproved (apart from the students of course; one might even say that if there was a to be discussed output, in most instances it appeared to be a negative one). Therefore, the quality or even amount of teaching did not discriminate between those who teach. A main problem then is to reach consensus about an educational output, and a reliable, valid, objective and, most of all, acceptable performance indicator may also result in the introduction of an incentive for a better performance. I remind you to our observation that the use of standardized questionnaires probably acts as motive to improve, that is to obtain higher scores next time.

However, when there is *no external need* for a better teaching performance everything depends on the *personal willingness and capacities of individual teachers*, or on the opinion of people who are important in a particular faculty (and sometimes on the troubles caused by complaints of students!).

Many university teachers are indeed willing to improve their teaching, but also many of them are not, «hindered as one is by ignorance, self-overestimation, stubbornness, self-defence, arrogance or fear, or by an underestimation of the importance of good teaching» (Drenth et al., 1986, p. 276). It may even be suspected —although direct evidence is lacking— that a linear relationship exists between the quality of teaching and the wish to improve that teaching: the rich get richer, the poor remain poor.

Some teachers then will only evaluate (and improve) their performance if they are in some way forced to do so: in other words, when an outside need to do better is created. (It is not appropriate to explain here in detail the very complicated structure (by which I mean the distribution of power) within departments, faculties and the university as a whole in the Netherlands. It suffices to state that practically speaking only the Faculty Board can, and sometimes will, force teachers to participate in an evaluation procedure.)

In this way, self-evaluation of (unwilling) *teachers* will take place only if the Faculty Board decides to start a faculty-wide evaluation procedure. One should keep in mind that such an evaluation procedure can be called an internal self-evaluation from the viewpoint of the faculty, but surely will be perceived as an (imposed) external evaluation from the viewpoint of the teachers.

What has been said about particular teachers is also relevant with respect to *faculties* and Faculty Boards (partly consisting in fact, of the

same persons): some will start such an overall evaluation procedure at their own will, because they are convinced of the importance of continuous evaluation of the teaching performance within that faculty, others will not. Practically speaking only the university's Executive Board can, and sometimes will, force faculties to start an evaluation procedure.

Self-evaluation of (unwilling) faculties will take place only if the Executive Board decides to start an institution-wide evaluation procedure. Again: such a procedure can be called an internal self-evaluation from the viewpoint of the university, but will surely be perceived as an «imposed» external evaluation from the viewpoint of several faculties.

It is possible to continue up to the governmental level as an outside need to promote *institutional* self-evaluation, and I will do so in the next hour. Now it is important to realize that neither institutional self-evaluation, nor faculty self-evaluation, should be considered as a goal by itself. What matters is (should be) the self-evaluation of teachers, and of course, the improvement of teaching behaviour. In this respect institutional quality control can be seen as a necessary but not sufficient step, meant to achieve the short-term purpose of faculty self-evaluation and the long-term purpose of teacher self-evaluation.

6. Conclusion

I am arguing here for a top-down approach: from the government to the universities, from the university to the faculties, from the faculty to the teachers.

Two more things must be said.

1. Firstly, it must be kept in mind at each level —and I can't stress this often enough— that *the real quality of education lies in the interaction between teacher and student*. What matters are these two actors (of course, within the context of the educational programme), with the emphasis on the teachers. After all, if anything has to change, it is he or she who must do the job. For instance, improve his or her testing practices, review the written materials, train the didactic skills, and so on.

Therefore, any chosen performance indicator must be avoided which is not clearly relevant from the point of view of teachers or students. The reason for this is clear: why would any particular teacher put any effort in the improvement of his or her educational performance if the indicator used is not accepted as reliable and valid? The teacher must

be convinced that the indicator really indicates a possible lack in his or her performance.

2. Apart from this more internal motive, I remind you of the absence of any external motive for the improvement of teaching. To overcome this problem, one can imagine a change in the procedures in use for the promotion of university teachers, or a finer differentiation in salary scales, or the qualifications required for an appointment, and so on. As a matter of fact, in my own university changes like these are in discussion and some of them are even carried through already. But maybe it is not always necessary to go that far.

Several times I mentioned the use of standardized questionnaires. In one faculty the teacher who scores highest is declared «teacher of the year», and receives a challenge-cup. It is a joke in a way, but with a serious undertone: *teachers should be flattered by the public recognition that they are performing well, and they should be challenged to do even better*. Possibly this is the most important human factor after all, and the very essence of the problem of quality assessment in higher education.

Dirección del autor: Willem van Os, Educational Service Office, Free University of Amsterdam, Postbus 7161, 1007 MC Amsterdam, The Netherlands.

Fecha de recepción de la versión definitiva de este artículo: 10.VII.1990.

BIBLIOGRAFIA

- CAVE, M., HANNEY, S., KOGAN, M. y TREVETT, G. (1988) *The use of performance indicators in higher education* (London, Jessica Kingsley Publishers).
- COOK, C. M. (1988) *Internal and external evaluation: Reflections on the American experience*. Paper presented at the International Conference on Quality Assessment in Higher Education (Noordwijkerhout, The Netherlands).
- DRENTH, P. J. D., Van Os, W. y BERNAERT, G. F. (1986) Improvement of quality of education through internal evaluation, *International Journal of Institutional Management in Higher Education*, 10, 275-281.
- HOFSTEE, W. K. B. (1985) Notities over ondrewijsrendement en kwaliteit, *Universiteit en Hogeschool*, 32, 145-153.
- LYNTON, E. A. (1988) *The role of self-study in quality assessment of institutions of higher education*. Paper presented at the International Conference on Quality Assessment in Higher Education (Noordwijkerhout, The Netherlands).
- Ministry of Education (1985) *Higher Education: Autonomy and Quality* (The Hague: Staatsdrukkerij).
- Ministry of Education (1987) *Higher Education and Research Plan* (The Hague: Staatsdrukkerij).
- MOSES, I. (1987) Educational development units: A cross-cultural perspective, *Higher Education*, 16, 449-479.
- Os, W. van (1988) *Questions sur l'éducation: Nécessité, possibilité, utilité*. Paper

- presented at the Colloque International: Formation, Evaluation, Selection par Questionnaires Fermés (Paris, E.E.I.E.E.).
- (1989) *De evaluatie van cursussen binnen AMOS*. Paper presented at the National Congress Hoger Onderwijs (Groningen, The Netherlands).
- Os, W. van, DRENTH, P. J. D. y BERNAERT, G. F. (1987) AMOS: An evaluation model for institutions of higher education, *Higher Education*, 16, 243-256.

SUMMARY: STRATEGIES FOR QUALITY ASSESSMENT. THE HUMAN FACTOR.

The purpose of this paper is to provide some reflections about need to establish clear indicators that allow to evaluate the quality of Higher Education. The establishment of these indicators is a very complex task. The central theme running through the article is the importance of human factor —professors and pupils— which the author pay attention for to attain the necessary optimization of the process of teaching in the Universities.

KEY WORDS: Teacher evaluation, Quality of teaching, Human factor, indicators of quality.

SUMARIO: ESTRATEGIAS PARA EVALUAR LA CALIDAD: EL FACTOR HUMANO.

Me gustaría comenzar presentando cuatro afirmaciones que reflejan mi opinión acerca de la evaluación de la calidad de la enseñanza superior:

- no es fácil evaluar la calidad de la enseñanza superior
- no hay un método ideal para evaluar la calidad de la enseñanza superior
- a pesar de todo, es necesario evaluar la calidad de la enseñanza
- el éxito del proceso evaluador dependerá en gran medida de la disponibilidad para la cooperación entre profesores y estudiantes: el factor humano.

En estas afirmaciones subyace una proposición básica: el propósito de evaluar la enseñanza será la mejora de la calidad de dicha enseñanza. La preocupación por la calidad de la enseñanza es un tema que está actualmente recibiendo mucha atención por parte de las autoridades académicas y de los propios profesionales de la enseñanza. Aunque se trata de una vieja preocupación, la novedad actual es plantear su análisis desde una visión de conjunto, sobrepasando la dimensión estrictamente didáctica para incluir referencias económicas, políticas, sociales...

Sobre todo interesa llegar a establecer indicadores que nos permitan evaluar adecuadamente la institución, desde puntos de vista internos y externos a la propia institución. El establecimiento de estos indicadores es una tarea de difícil solución ya que no existe pleno acuerdo sobre los objetivos y finalidades de la enseñanza superior. Aun así, todo esfuerzo realizado en esta línea se juzga como positivo. Especialmente interesante me parece la consideración de la relación que se establezca entre los elementos humanos del sistema (profesores y alumnos) como requisito necesario para optimizar el proceso.