

Effects of Quality Assessment in Dutch Higher Education

M. M. H. FREDERIKS, D. F. WESTERHEIJDEN & P. J. M. WEUSTHOF

Quality management [1] has been an issue in Dutch higher education since the 1980s. Quality assessment (QA) of education was introduced on the political agenda as part of the new policy of the government, with the policy paper Higher Education: Autonomy and Quality (1985). In exchange for a larger measure of administrative autonomy, the universities promised to retain and enhance their levels of quality in education. Quality assessment then appeared on a systematic and nationwide scale in 1988, when the Association of Universities in the Netherlands (VSNU) implemented its new responsibility. The VSNU QA system consists of an external Visiting Committee (VC) for each discipline or cluster of study programmes [2] operating nationwide.

To prepare for the VC, those responsible for each study programme are required to write a self-evaluation (we will take 'self-evaluation' and 'self-study' to be synonyms). When the VC has visited all study programmes in its 'area', it writes a report based on the self-evaluations that were its input and on the experience of the visits to the locations. These visits usually last for two intensive days, during which the VC talks to representatives of all actors involved in the study programme (including students) and at the end of which, based on the self-evaluation and the impressions of the visit, a preliminary comment and judgment about the study programme is given by the chair. This text, after being commented on by the study programme, is included in the report of the visiting committee. The comments, recommendations and judgments about the individual study programmes are preceded in the report by a chapter on the general 'state of the art', shared problems and so on in the discipline covered.

The self-evaluations and the report by the VC are the key documents in this quality spiral, of which the second cycle is starting in 1994. In 1990 a parallel system of QA was introduced by the VSNU counterpart, the HBO Council, for the non-university sector. The VC's judgments and self-evaluations are intended, in the first place, to enhance the quality of the study programme; accountability to the government and society in general is only the second goal (Vroeijsstijn & Acherman, 1990). This raises the question of to what extent the results of the evaluations are really utilised.

Commissioned by the Ministry of Education and Science, in parallel with a 'mid-term review' by the Inspectorate of Higher Education (1992), the authors started a research project, the central question of which can be formulated briefly as: *What are the effects of the QA system on the quality activities of higher education institutions?* This project was not intended as an evaluation of the evaluation system

in the Netherlands; we focused not on the QA system as such, but on the activities at the study programme and institutional levels.

The project consisted of two stages. First, all study programmes in higher education in the Netherlands were sent a questionnaire about the activities they perform regarding quality control and quality enhancement, focusing especially on follow-up activities after the self-evaluation and the external visitation. The second stage consisted of 12 case studies of study programmes that were selected on the basis of the questionnaire answers. The case studies comprise analyses of the self-study and VC reports, and a number of interviews (usually three per case) with QA experts on the study programme, faculty and institutional level.

The total response to the questionnaire [3] by university study programmes was 239 (60%), not quite evenly, but reasonably, distributed over the nine disciplinary sectors usually distinguished in Dutch higher education policy documents [4]. Also, all 13 universities in the Netherlands are represented. For the non-university sector the response was lower: 240 study programmes, which means a rate of 31%. All seven disciplinary sectors (biology and law cannot be identified as disciplinary sectors in the HBO institutions) are represented, although again not quite evenly. Of the total response, 64% of the university study programmes have been visited or were in the process of visitation (from the writing of the self-study to the actual visitation). For the HBOs the number of 'visited' study programmes amounts to 21%. Taking into consideration the fact that at the moment of our research about two-thirds of the universities and less than one-third of the HBO institutions were visited, the response rates can be seen as representative for the total population of visited study programmes.

We added case studies to obtain a more thorough understanding of the follow-up of QA. Therefore, eight universities and four HBO institutions were selected with the help of our survey data. Selection criteria were a reasonable distribution over the disciplinary sectors, the institutions, the extent of measures taken in response to the evaluations and, in the HBOs, the question of whether or not a visitation had taken place. As a result of this selection 35 QA experts on the study programme, faculty and institutional level were interviewed.

The present paper contains some of the final results of the questionnaire and case studies (see also Frederiks, Westerheijden & Weusthof, 1993). Before going into the results, we shall first give a brief account of the theoretical background of the project, indicating the most important relationships expected to explain the degree of effectiveness of the QA procedures. Following the confrontation of these hypothetical expectations with the facts, we shall finally discuss the meaning of the empirical data for the theory, and for the QA procedures in Dutch higher education. Apart from the testing of the hypotheses we will focus on the QA activities undertaken by universities and HBO institutions (also called 'polytechnics' or non-university sector), the use made of the self-evaluation and VC report and the level of satisfaction with the QA system.

Theoretical Framework and Hypotheses

A Conceptual Model of Quality Assessment

The Dutch quality assessment system may suggest that a linear relationship exists between being informed about evaluation results and the utilisation of these results

by individuals or groups belonging to higher education organisations. However, empirical research often shows that no use is made of evaluation results (e.g., Cooley & Bickel, 1986). In the United States some empirical studies were conducted on the nature of self-evaluation and on the factors which seem to be related to success in such efforts. A major empirical study of some 200 self-evaluation processes identified three factors as being related to perceived successful self-evaluations: the attitude of the chief executive or other institutional leaders, the motivation of institutional actors, and the organisational context (Kells & Kirkwood, 1979). Later a number of empirical studies pursued these matters. Harris (1984) found that effectiveness of self-evaluation seems to be related to the availability of data and the institutional research capacity. Cornett (1987) confirmed the findings of Kells & Kirkwood in the allied health fields. Although these studies give some insight into the practice of self-evaluation in higher education institutions in the United States, the theoretical knowledge about the characteristics of self-evaluation processes and the factors that influence the success of these processes is limited. In our study we have tried, on the basis of an organisational and political point of view, to develop a theoretical framework in which several factors are related to a successful evaluation.

Figure 1 contains a conceptual model that is based on a systems theoretical approach, visualising the relationships among evaluations of the quality of teaching and the effects of these evaluations, as well as a cluster of explanatory variables. In this figure we see that information about educational characteristics is the input for evaluation and that the output consists of evaluation results (analyses and judgments). The extent to which results are used to improve the quality of education is, for our purpose, the effect of the quality management system. The characteristics of actors, together with characteristics of the organisation, are 'contextual' factors that influence the degree and form of utilisation.

We define the key dependent variable 'utilisation' as the extent to which the results of an internal or external evaluation are handled by a higher education organisation. We distinguish between *no utilisation*, *passive utilisation* and *active utilisation*. 'No utilisation' means that evaluation results are neglected by a higher education organisation. 'Passive utilisation' refers to use of results without immediately taking decisions to change (parts of) the curriculum or the organisation, but, for instance, to disseminate evaluation reports within the organisation, to discuss the results of the process in formal settings, or to make recommendations for future changes. 'Active utilisation' is defined as taking measures on the basis of available evaluation results. One can distinguish a ranking of utilisation from no utilisation to passive and active utilisation, but that does not imply a normative ranking: active utilisation is not necessarily 'better' than passive utilisation. It may well be that no active utilisation is called for, e.g., when the evaluation results are completely positive. Also, passive utilisation may lead to future changes that are not directly linked (in the consciousness of the organisation members) to the evaluation, but that would not have occurred without it. Another point is that opportunities for change may differ. Not only is active utilisation not expected when all evaluation results are positive, but also some VCs succeed better in formulating recommendations that can be used for decision making than other VCs.

In connection with the conceptual framework, a number of hypotheses are formulated. The argumentation for these hypotheses is based on the contingency

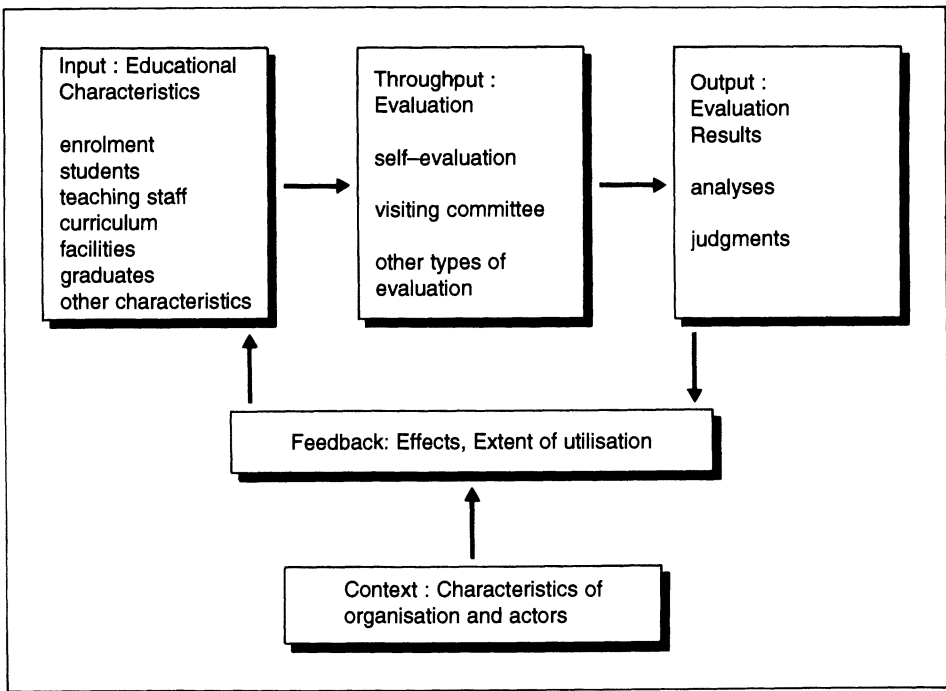


FIG. 1: Conceptual model of the quality management process

approach (Child, 1984) and the political economics approach (Lieshout, 1989). These two approaches can provide us with explanations for the relationship between, on the one hand, characteristics of higher education organisations and the actors who operate inside these organisations and, on the other hand, the extent of utilisation of the external (VC) and internal evaluation (self-study) results.

The Contingency Approach

The basic assumption of the contingency approach is that an effective organisation is optimally adjusted to (contingent with) specific environmental circumstances. The central question of this approach is: *Under what conditions can an organisation function most effectively?* Child (1984) distinguishes a number of factors which can be conceived as explanations for the sometimes large differences between organisations. The most important contingency factors in the context of our research are: the environment, the technology, the size and the centralisation of an organisation. We assume that environment, technology, size and centralisation are related to the extent of active utilisation of the internal and external evaluation results.

The environment of higher education organisations consists of, among others, potential students, employers and the financing bodies of higher education institutions, of which the national government is the most important. We see the visiting committee as another primary environmental factor for a higher education organisation. Because of the serious consequences that a negative judgment by a visiting

committee can have for a study programme (student enrolments may decrease or finance may be at risk), we expect more active utilisation of evaluation results on the basis of the VC report than on the basis of the self-evaluation report.

Hypothesis 1: External evaluations by visiting committees are more likely than internal self-evaluations to result in active utilisation of the evaluation report.

We distinguish two characteristics that refer to the structure of a higher education organisation, namely the size of the study programme and the extent of centralisation. The larger an organisation becomes, the more the need evolves to write down all kinds of previously unwritten rules and informal procedures: empirical research has confirmed that a positive relationship exists between the size of an organisation and the extent of formalisation (Mintzberg, 1979). One consequence of larger size may be that decision making on the basis of evaluation results (active utilisation) takes longer or even never takes place at all. It can be assumed that larger study programmes will mean more formal decision-making processes, with long advice and discussion tracks and more possibilities for delaying decisions.

Hypothesis 2: The larger a study programme, the less likely it is that active utilisation of evaluation results will occur.

A negative judgment by the VC might have severe consequences for the study programme. In anticipation of this potential threat, the management may seek to broaden its control over the study programme. In this way optimal conditions can be created for the external peer review, which enhances the chances of a positive judgment. Empirical research (Baldrige *et al.*, 1977) shows that a positive relationship exists between the presence of an external threat (in this case the VC) and centralisation. In case of a threat the formal decision makers will try to get more of a grip on the decision-making process (centralisation) and use their increased power to take measures (active utilisation of the evaluation results) that they conceive as necessary to cope with the external threat.

Hypothesis 3: The greater the centralisation of a study programme, the more likely it is that active utilisation of evaluation results will occur.

We can define technology informally as knowledge about ways of producing (Groot, 1988). Scheerens (1987) describes the technology of higher education as 'unclear': it is uncertain how the input of the educational process (teachers, students, teaching materials) is exactly transferred into output (graduates). This uncertainty will have a negative impact on the extent of active utilisation, since it is not known what difference it makes to change any single aspect. We have taken the availability of data required by the VSNU and HBO Council for the self-evaluations as indicators of the uncertainty about technology.

Hypothesis 4: The greater the availability of data required by VSNU and the HBO Council for the self-evaluations, the more likely it is that active utilisation of self-evaluation results will occur.

The Political Economics Approach

Apart from the characteristics of organisations, we can also identify characteristics of actors within the organisation. Especially important in this context are the reputations for power of these actors. The political economics approach is based

on the principle of maximisation of utility. Lieshout (1989) equates utility with the concept of power. This results in the proposition of an actor maximising power: the actor's perceptions of the power (hence *reputations* for power) of other relevant actors, combined with their striving to maximise their own power, forms the input for the actor's weighing of alternative courses of action. For our purpose, two types of actor can be distinguished in higher education institutions, namely, evaluators (those who carry out the internal evaluations, such as the self-study) and the decision makers, who have to decide on how to handle the evaluation results.

The reputation for power of the evaluators is decisive for the question of whether decision makers take the self-evaluation seriously. If the recommendations are threatening to the status quo and the reputation for power of the evaluators is relatively low, the decision makers have not much to fear from not following the advice.

Hypothesis 5: The lower the reputation for power of evaluators, the less likely it is that the results of the self-study will be actively used.

Not only do the reputations for power of the evaluators play an important role in determining the reaction to evaluations, the reputations for power of the decision makers count as well. Powerful actors are more likely not to use recommendations that might harm the status quo and therefore their interests.

Hypothesis 6: The higher the reputation for power of decision makers, the less likely it is that active utilisation of evaluation results will occur.

A final important point is related to the overlap between evaluators and decision makers. If the actors who perform the evaluation and the ones who have to decide on the utilisation are (partly) the same, then the consequence might be a more active utilisation, because this is in the interests of these evaluators-and-decision-makers.

Hypothesis 7: The more decision makers are involved in performing the self-evaluation, the more likely it is that results of the self-evaluation will be actively used.

Test of the Hypotheses

In this section we present the findings of the empirical tests of the seven hypotheses. To test these hypotheses we have calculated correlations between on the one hand the independent variables concerning contingency and power factors, and on the other the dependent variables 'active utilisation of the self-evaluation (SE) report' and 'active utilisation of the VC report'. This has been done for both the university and the non-university sector in Dutch higher education. We have used the number of measures taken in response to evaluations as an operationalisation of the extent of active utilisation.

As we can see from Table I, hypothesis 1, namely the proposition about the positive influence of environmental pressure (VC) on active utilisation, is partly corroborated. This does indeed appear to be the case for the universities. To make a comparison between the university and the non-university sector possible, active utilisation is calculated as the average number of measures taken following the report divided by the average number of recommendations. The contingency hypotheses 2 and 3, that relate active utilisation to size of study programme and centralisation respectively, are falsified. Interesting in this respect is the significant

TABLE I: Correlations between active utilisation and characteristics of the organisation and actors

Hypothesis	Active utilisation SE report universities	Active utilisation VC report universities	Active utilisation SE report HBO	Active utilisation VC report HBO
1 Active utilisation VC report > active utilisation SE report	6.3/9.6 = .66 N = 154	6.7/7.8 = .86 N = 106	7.8/13.5 = .56 N = 51	5.2/8.6 = .60 N = 31
2 Size of study programme → active utilisation (-)	r = .0930 N = 82 p < .406	r = .2595 N = 82 p < .019	r = .0890 N = 40 p < .585	r = -.2698 N = 21 p < .237
3 Centralisation → active utilisation	r = .146 N = 153 p < .070	r = .0683 N = 106 p < .487	r = .2427 N = 51 p < .086	r = .1490 N = 30 p < .432
4 Availability of data for SE report → active utilisation SE report	r = .1509 N = .105 p < .124		r = .4063 N = 40 p < .009	
5 Reputation for power of evaluators → active utilisation SE report	r = .0524 N = 132 p < .551		r = -.2376 N = 44 p < .120	
6 Reputation for power of decision makers → active utilisation (-)	r = .0963 N = 145 p < .249	r = .0315 N = 100 p < .756	r = -.3250 N = 50 p < .021	r = -.1425 N = 29 p < .481
7 Involvement of decision makers in SE → active utilisation SE report	r = -.0732 N = 145 p < .382		r = .3375 N = 50 p < .017	

Shaded background indicates significant correlation (p < .05)

correlation we found concerning the relationship between size of study programme (indicated by the number of first-year students) and active utilisation of the VC report in universities. This correlation turned out to be positive, while a negative relationship was expected. The most plausible explanation seems to be that in the response to our questionnaire the languages and cultural studies sector was over-represented in the university sample. This sector contains many small study programmes and their active utilisation was shown to be, on average, relatively low. This may have been caused by the low level of satisfaction with the VCs for language programmes [5]. So for this sector there is a positive relationship between active utilisation and size, which because of the over-representation of this sector may result in a positive correlation for the entire university population. The remaining hypothesis based on the contingency approach (number 4) is partly corroborated. Only for the non-university sector could a positive and significant correlation between the data available for the SE report and the extent of active utilisation of the SE report be identified. Further analysis shows that in the university sector the variation in answers tends to be smaller, which may explain why no significant correlation was established.

The same argument may also apply to the partial corroboration of hypotheses 6 and 7, based on the political economics approach. Empirical support for the

expected negative relationship between the reputation for power of decision makers and active utilisation (hypothesis 6) was only found in the HBOs and then solely for the SE report. Possibly the reputation for power of decision makers in the non-university sector is more important than in the university sector, where regulation of decision making about educational matters is more extensive and formalised and hence the variation in reputation for power tends to be less. Furthermore, it seems plausible that reputations for power of decision makers are more important in response to SE reports and less so in response to VC reports, due to the relatively higher reputation for power of the VC. Further analysis shows that decision makers on the central level of the study programme are (by law) more involved in the self-evaluation in universities than in the HBO institutions. We can argue that the variation between universities and HBOs in degree of formalisation of the decision-making processes can also constitute an explanation for the falsification of hypothesis 7 for the universities (concerning the expected positive relation between involvement of decision makers in the self-study and the active utilisation of the SE report). Also, we have found no support for the expectation that a higher reputation for power of evaluators has a positive impact on active utilisation of the SE report: hypothesis 5 has been falsified.

In sum, we found some support for both the contingency as well as the political economics hypotheses. However, this support consists only of four partially corroborated hypotheses, while the other three hypotheses were falsified.

In the next three sections, we shall take a broader look at the research data, which may lead to a better insight into the way QA is embedded and to suggestions for alternative explanations, a subject to which we shall return in the final section. But first we shall give an overview of the existing quality management activities, followed by a section about the utilisation of the SE and VC reports and a section in which we will concentrate on satisfaction with assessment procedures. The universities and HBO institutions will be presented separately, primarily because the QA system in the HBO sector started later than in the university sector, so that experience with it and the time to implement changes were significantly less than in the university sector.

Quality Management Activities

Quality Management in the Universities

Internal Evaluations and the Self-Evaluation: Out of the valid responses to the survey, 95% of the university study programmes indicate that some kind of explicit quality activity is conducted within the organisation. In most cases (64%), this includes the VSNU procedure (the self-evaluation and/or the visitation). However, 69% indicate that some form of evaluation already existed, usually because it was the local tradition; special occasions, like governmental budget cuts or externally visible failure (unemployment), were not often given as reasons. These earlier evaluations were mostly concerned with the study programme as a whole (67%) or large parts (several years) of the programme (19%), so in extent they can be compared to the self-evaluations. The information resulting from previous internal evaluations could be used (partially or in full) for the VSNU self-evaluation in more than 80% of the study programmes.

Institutional Quality Management of Teaching: Internal quality evaluations (see

TABLE II: Subjects of quality activities

	Universities %	HBO %
Curriculum	94	98
Efficiency of teaching	93	89
Retention and completion rates	87	87
Aims of study programme	74	81
Characteristics of entering students	43	69

Table II) mostly focus upon the curriculum, or on the topics fashionable at the time, either because of public discussions or because of the financial incentives built into the government's funding system, i.e. on efficiency of teaching and on retention rates, dropout rates and completion time.

Slightly less often the aims and goals of the study programme are part of the internal quality activities. But the characteristics of students entering the programme are only mentioned in 43% of cases. This probably reflects the freedom of choice, and the almost unlimited right to enrol, that students have in the Dutch university system: study programmes or institutions have hardly any instruments to influence enrolment.

Internal quality evaluations are indeed to a large extent internal to the university study programmes in the points of view taken into account (see Table III). In practically all cases, student opinions and staff opinions are taken into account.

The views of other stakeholders, like faculty managers, alumni or employers, all score less than 50%.

How was the self-evaluation conducted: was it a task delegated to a single staff member? An individual staff member was indeed strongly or very strongly involved in 83% of the cases. However, that does not imply that it was an individual's study: in practically every study programme, the study programme committee (*studierichtingscommissie*) [7], the dean responsible for education, an ad hoc committee or other actors were strongly or very strongly involved as well. In 80% of the valid responses an ad hoc committee was installed. Ad hoc committees usually consist of members of the study programme committee and of academic staff, either as representatives of the departments (*vakgroepen*) or as individuals.

Moreover, the self-evaluation is usually indeed done by 'self': only in approximately 17% of the cases was some actor from outside the study programme involved. And those outside persons or agencies are, almost without exception, part of the larger faculty or of the university—in other words, part of the 'inner circle' of the environment.

Quality management of teaching used to be a responsibility primarily of the

TABLE III: Points of view used in quality activities

Opinions of	Universities %	HBO %
Students	100	97
Academic/teaching staff	99	95
Study programme or faculty/sector managers	37	42
Alumni	48	75
Employers/professionals	25	73

faculty or study programme level within Dutch universities. In most of the institutions involved in the case studies of our project, central quality management is 'reactive' rather than 'active'. A 'reactive' policy means that initiatives for quality management and enhancement are taken by the faculty or study programme, while the central actors are mostly in a supportive or stimulating role. This may have to do with the Continental tradition (Clark, 1983) of Dutch higher education: institutional management is not strongly developed, most competencies regarding teaching reside in the faculty or study programme and in this respect certainly the university is a 'loosely coupled' system (Birnbaum, 1989) with at least as many (management) cultures as there are faculties (Becher, 1989).

Still 'reactive', but more 'active' than most, is one university where a central committee critically reads all self-evaluations for the VSNU QA procedure and comments on their content before they are sent to the VSNU visiting committee. Two of our seven case studies (out of 13 universities in the Netherlands) show 'active' central quality management of teaching, in that institution-wide goals are set. In one of the two, the completion rate is the prime indicator (in line with its central role in the Dutch funding model (Jongbloed & Westerheijden, 1994)), supporting study programmes to take measures to raise these if necessary. In the other, a policy of raising completion rates has evolved into a more encompassing quality policy where 'studiability' [8] of the curricula has become the central issue.

Quality Management of Teaching in HBO Institutions

Internal Evaluations and the Self-Evaluation: Quality activities are as frequent in the HBO institutions as in the universities: 95% say they have some form of quality management activity, even though in only 21% of the cases does this include the self-evaluation for the national QA procedure coordinated by the HBO Council (the QA procedure started two years later than in the universities). This means that the 'empirical base' for statements related to the SE is rather small (N = 51).

Most of the HBO institutions have come into existence in their present form as a result of a large-scale merger operation in the 1980s (Goedegebuure, 1992). It is, therefore, not surprising to find that quality management activities originated in recent years too. However, from the four institutions included in the case studies in this project, we tentatively conclude that the HBO institutions are more 'actively' involved in quality management of teaching than the central-level actors of the universities. In one of the cases included in our project, this is done through management contracts. Another interesting initiative in several HBO institutions is the organisation of 'trial visits': the self-evaluation is read by a committee appointed by the institution itself (consisting of institutional staff members from other study programmes and central-level quality managers, and of employers or professional representatives) prior to the official HBO Council visiting committee.

Utilisation of Self-Evaluation and VC Report

The VC report, together with the self-evaluation with its analyses, self-judgments and recommendations, is the input for the next phase of the process at the study programme level: the utilisation of the evaluation results (the feedback part of the loop in Figure 1). In Table IV we present a first impression of the effects of the

TABLE IV: The extent of utilisation of internal and external evaluation results

	SE report Universities	VC report Universities	SE report HBOs	VC report HBOs
Average number of actors discussing reports	3.8	3.1	3.4	2.7
Average number of recommendations	9.6	7.8	13.5	8.8
Average number of measures following report	6.3	6.7	7.5	5.2
Percentage of drastic measures	27	19	26	18

internal and external evaluations at the universities and HBOs. In the next sections we go into active, passive and no utilisation of the reports.

No Utilisation

One way to measure no utilisation has been when neither the self-evaluation report nor the VC report were discussed officially in any kind of meeting (passive utilisation), and no measures were taken in connection with these reports (active utilisation). These conditions, in conjunction, apply to four cases (two in universities and two in HBOs). This means, as far as we can observe, that participation in the national QA system for no more than four study programmes (out of a sample of 137) is nothing but a 'ritual dance'. However, a different picture emerges if we look at no utilisation of the SE and the VC report separately. No utilisation of the SE report occurs in nine out of 154 cases in universities, and in six out of 51 HBO study programmes. Passive as well as active utilisation of the VC report have been absent in 14 out of 106 university study programmes, and in six out of 31 HBO cases. Although no utilisation seems to be more widely spread than our first measurement indicated, there is no cause for exaggeration. Still, the vast majority of study programmes in both sectors utilise the reports either actively or passively: 'ritual dancing' remains a rarity.

Passive Utilisation

As we can see in Table IV the passive utilisation of both reports (indicated by the average number of actors discussing the reports: the maximum number of actors is seven) is higher for the universities than for the HBO institutions. In both sectors the SE rapport gives rise to more formal discussion than the VC report. However, further analysis of our survey data shows that the VC report is more often discussed by the faculty board (and its equivalent in the non-university sector) than the SE report, which is consistent with formal decision-making rules. From the fact that SE reports tend to be disseminated rather widely throughout the faculty, we can deduce the existence of a collective interest in the self-evaluation. This does not imply that usually all academic staff members and students receive copies of the report, but they can all have access to the information if they wish.

In universities the faculty boards, faculty councils, study programme commit-

tees and departments discussed both reports in more than 50% of the cases. The reports were also discussed in 50% or more of the cases by student representatives, but in only 15% (for the VC report) and 36% (for the SE report) by individual staff members. In contrast with these findings, passive utilisation by the equivalent of the faculty council in the HBO sector tends to be lower, but discussion of the reports by teachers is higher (50% or more).

From our interviews we formed the impression that on the institutional level the evaluation results are little utilised. This is not surprising, since responsibility for the content of the SE report and the follow-up of both the internal and external evaluation rests exclusively with the faculties. However, in a few cases the faculties were assisted at the institutional level in preparing the self-evaluation, but only at their request. Furthermore, although passive utilisation (formal discussion) of the SE report on the institutional level does not in general occur, the VC reports were in most cases discussed on the central level (by the Committee of Deans). This passive utilisation may also include discussion with the faculty and/or the study programme about the follow-up, especially when the judgments of the VC were rather negative.

Active Utilisation

Ultimately, one of the aims of evaluation is improving the quality of the study programme. As a rule, the evaluation reports show that something could be improved: out of the 31 categories presented, the average university study programme that has written an SE report indicates that on 9.6 points (HBOs: 13.5) some recommendation(s) was (were) made in the self-evaluation report, and on 7.8 (HBOs: 8.6) in the VC report (see Table IV). In Table V the percentage of study programmes that have taken measures on a certain subject following the SE and VC recommendations are confronted with the percentage of study programmes that received recommendations on that subject. From Table V it can be noticed that in both universities and HBOs the most 'popular' subjects mentioned in recommendations relate to the content and aims of the curriculum (including study progress). In universities considerably less advice is given regarding the qualifications of teachers and graduates. In the HBO sector, on the contrary, qualifications of teachers and graduates receive much more attention, especially in the self-evaluation. Here facilities (computers, libraries, etc.) are the less 'popular' subject for recommendation.

It is clear from Table IV that the average number of recommendations (both for the SE and the VC report) is lower for the universities than for the HBO sector. The figures concerning the average number of measures taken following the reports are more diversified. Following the SE report more measures are taken in HBO institutions (7.5 vs. 6.3), while in response to the VC report the average number of measures is higher in universities (6.7 vs. 5.2). However, this comparison between the university and non-university sector is not quite fair: to compare active utilisation within the two sectors it seems appropriate to take the average number of recommendations into account. The active utilisation, corrected for the average number of recommendations, can be seen in Table I (hypothesis 1). Then a different picture emerges: the active utilisation in both sectors is higher for the VC than for the SE report.

But the important question, determining the actual effects, must be: are

TABLE V: Active utilisation per subject in QA reports (%)

	Universities		HBO institutions	
	measures/ recommen- dations SE report	measures/ recommen- dations VC report	measures/ recommen- dations SE report	measures/ recommen- dations VC report
Student enrolment	20/46	11/29	20/57	10/32
Propaedeutic programme	10/60	19/74	6/71	23/52
Master's/bachelor's programme	13/67	23/81	24/78	26/74
Teachers	8/28	7/30	20/53	10/35
Facilities	14/53	12/43	14/45	3/29
Graduates	10/39	14/37	24/55	13/45

measures taken as a consequence of the evaluation process? There certainly is not a one-to-one ('linear') relationship between recommendations and measures. Indeed, the number of times that recommendations on a certain subject in the reports correspond with measures on the same subject are very low, as we can also see in Table V.

The measures taken are, as a rule, not very drastic (see Table IV). On the contrary, in both sectors and following both the self-evaluation and the VC report, the median class of answers is 'not far-reaching' (the lowest of the five possibilities for answering) for all subjects except the master's (HBO: bachelor's) and propaedeutic programme. The subjects 'student enrolment', 'graduates' and 'teachers' have the lowest averages for drastic measures. The percentages of drastic measures (average for all categories in a study programme: four or more out of five) are higher, both in universities and in HBO institutions, for measures taken following the SE report. This seems logical if we assume that as a rule drastic measures are already taken in anticipation of the visitation.

Satisfaction with the QA System

Overall Satisfaction

Included in our questionnaire was a question concerning the appreciation by the study programme about the different types of evaluation. On a scale of one to ten (the usual rating scale in Dutch education: one being the minimum and ten the maximum) the respondents could give their judgment about internal quality management, the visitation and the self-evaluation. Respondents were also asked to motivate their judgment. The marks that the respondents gave were taken as an indicator for the degree of satisfaction with the elements of the QA system. This can be considered both as a side effect of the present round of the VSNU and HBO Council systems and as a first indication for the willingness of the study programmes to cooperate in the second round of evaluations. Apparently, the median class of answers being seven, the study programmes are fairly satisfied with the existing QA system (see Table VI). At this point we wish to point out the

TABLE VI: Satisfaction of respondents with different types of evaluation

	Universities (total response)	Universities (‘visited’ response)	HBO (total response)	HBO (‘visited’ response)
Internal quality management	7.2 (N = 198)	7.2 (N = 136)	7.0 (N = 139)	7.4 (N = 49)
Self-evaluation	7.1 (N = 168)	7.2 (N = 139)	7.0 (N = 100)	7.3 (N = 49)
Visitation	6.4 (N = 156)	6.5 (N = 129)	6.9 (N = 85)	7.3 (N = 40)

Note: Under ‘visited’ we include the study programmes that were visited at the time of our survey or that were in the process of visitation, viz. that had finished their self-evaluation

differences between the total response and the so-called ‘visited’ response (the study programmes that were visited at the time of our survey or that were in the process of visitation, viz., that had finished their self-evaluation). It can be noted from Table VI that, while in universities the scores for the total and the ‘visited’ response hardly differ, in the HBO sector the ‘visited’ study programmes are more satisfied about QA than the total response. However, in comparing these averages we must be careful, because the non-university response is considerably lower than the response of universities.

Satisfaction with Internal Quality Management

There is little difference in satisfaction with internal quality management between the two sectors: HBOs are slightly more satisfied than universities (‘visited response’). Further analysis (correlations) shows that study programmes are more satisfied if they use judgments of the faculty board and of students in their internal evaluations (both sectors) and if they take characteristics of student enrolment into account (only universities). For the HBO sector we found a positive and significant ($p < .05$) correlation between satisfaction and the utilisation of judgments of alumni.

In their motivations respondents in both sectors indicated that regular evaluations and feedback to teachers, regular meetings between teachers and departments, participation of students, a small size of study programme, utilisation of information systems, modularisation of courses, a critical self-evaluation and the external pressure of the visitation all contributed to a positive attitude towards internal quality management. Obstacles to internal evaluations are also mentioned: financial problems, shortages of teachers and the fact that evaluations take a lot of time. A few respondents from HBO institutions point to the risk that evaluations may be perceived as a threat to the autonomy of teachers and express doubts about the appropriation of quantitative evaluation methods for small study programmes. From the comments of respondents we got the impression that many study programmes in the non-university sector are still in the starting phase of quality management.

Satisfaction with the Self-Evaluation

The degree of satisfaction with self-evaluation, as for internal quality management,

hardly differs between universities and HBO institutions. Study programmes tend to be more positive towards SE if they already have a tradition of evaluation and if the SE report has been formally discussed by the faculty board and faculty council (universities) and by the teachers and the students' committee (HBOs).

Positive judgments are motivated by pointing out that the self-evaluation increases awareness of the importance of quality management, enables the study programme to get a thorough overview of the 'state of the art' concerning educational matters and enhances the chance of critical reflection on the educational process. Negative conclusions in the report seem to be accepted better if teachers are more involved and have more influence on the follow-up. Many respondents who are negative about the self-evaluation motivate this by pointing to the heavy burden that making the SE report imposes on the study programme. The fact that some reports are made by one or just a few people and consequently lack the necessary support of larger groups within the faculty is another reason for dissatisfaction. Some respondents admit that the external pressure of the VC may stimulate the completion of the self-evaluation, but they say this also entails the risk that the SE may relate too much to the framework and references of the VC, thereby diminishing the importance of the self-evaluation as an instrument for internal quality management out of sight. Furthermore, the utilisation of data from institutional information systems as well as support from the central level are sometimes problematic in constructing the self-evaluation.

Not surprisingly then, the QA experts questioned in the interviews propose changes for the second round that are aimed at broader involvement of teachers in writing the report, an explication of the preferences of the VC and more attention being paid to the internal function of the SE as an instrument for internal quality management.

Satisfaction with the Visitation

The most significant deviation in Table VI that draws our attention is the relatively low appreciation of the visitation in universities, both compared to the level of satisfaction with the other elements of QA and compared to the higher scores for the visitation in the non-university sector. An analysis of the scores per discipline shows that this is partially caused by the negative judgment (5.6) of the VC in the university discipline of languages and cultural studies. In this discipline many study programmes were dissatisfied about the broad range of study programmes clustered under a few visiting committees. Further analysis shows that study programmes in both sectors have a more positive attitude towards the VC if a tradition of evaluation exists.

As motivations for a positive attitude towards the visitation, the importance of assessments by external experts, the stimulation of a 'quality culture' and the fact that VC recommendations often support desirable changes are mentioned. Factors that contribute to a negative attitude are that recommendations are often not very precise, reports are sometimes inconsistent, comparisons of not comparable study programmes are made, visitations sometimes coincide with reorganisations or holidays and the political views of the VC can have a negative influence on an independent assessment. Another major factor is the belief that if the self-evaluation is carried out properly, the VC has little more to add.

All experts, with one exception, who were interviewed held the opinion that the

present QA system should be maintained. This does not mean that changes are not desirable: most pleas were made for a less intensive procedure because the procedures take too much time now. Furthermore, the experts felt that the preferences of the VC should be more explicated, recommendations should be formulated more precisely and, in the second round, more attention should be given to the follow-up. University QA experts would like to see a less prominent role for retired professors as members of the VC and a more important role given to foreign experts. Several respondents from the HBO sector point to the importance of having more educational scientists on the committee.

Conclusions

Our first, and main, conclusion is that quality management of teaching is an issue that has obtained much more attention than before in Dutch higher education. In a previous research project of the university sector of Dutch higher education, before the policy changes took effect, one of the present authors found that internal quality management was a rare thing (Weusthof, 1989). The outcomes of that project suggest that self-evaluations mostly took place at the lower levels of universities (teacher evaluation, course unit evaluation) and that the results of these evaluations were rarely used in a structured way in faculty decision-making and planning processes. Now, on the contrary, many university and non-university study programmes or faculties have special committees or specially appointed staff members for quality management of education, and many more have written a self-evaluation: quality of education has certainly gained an important place on the agenda of decision makers. Another remarkable difference between this study and the earlier research project concerns the content of the self-evaluation. Apparently, as a consequence of the introduction of the VSNU and HBO Council, it is not only characteristics of the educational process (e.g. didactic qualities of teachers) which have become an object of evaluation, as proved to be the case in the previous research project, but also input and output characteristics of education (giving information to potential students, acquiring data about labour market prospects for graduates). In general we can say that a 'quality culture' may be putting out roots in the universities and HBO institutions. Involvement of the central-level actors of higher education institutions is less marked, which is in accordance with the procedures set up by the VSNU and the HBO Council, for these are focused at the study programme level.

Second, not only is quality on the agenda, but something is being done about it. Measures are taken in connection with the self-evaluations and the VC reports. However, we cannot say that the large amount of resources invested leads immediately to an equally large improvement in the quality of education; measures are not taken in response to every recommendation, nor are the measures taken drastic measures. Quality of education, however important, is only one of a multitude of issues in the institutions for higher education. Less far-reaching measures may have greater effects in future years. Furthermore, taking measures does not *necessarily* lead to improvement of education. The relation between taking measures and observing improvement is obscure. This originates in the complexity of the education process. There are many factors in this process that are very difficult to control. More important still is the lack of knowledge about cause and effect relationships in education; taking measures seems to be a matter of trial and error and of imitating behaviour.

Third, the level of satisfaction with the implementation of the Dutch quality management system within the institutions is fairly high. This is quite remarkable, taking into account that higher education organisations are often described as relatively autonomous organisations with little inclination to appreciate outside scrutiny. From this point of view the implementation of the present quality management system can be seen as a break-in into the ivory tower. Considering the relatively high level of satisfaction, the higher education institutions certainly do not shut their doors in the face of this attempt to break in. The level of satisfaction is also remarkable in view of the frequent remarks about how heavy a burden it is to write a self-evaluation. From our interviews it seems doubtful whether the same enthusiasm can be mobilised again for the second round of visits.

Finally, we have to conclude that our effort to explain differences in utilisation from a contingency theory perspective and from a political economics perspective have been partially successful. Further research can put us on the path to more adequate explanations of utilisation of the results of self-evaluations and visitations.

NOTES

- [1] Quality *management* is defined as: 'that aspect of the overall management function that determines and implements the quality policy [intentions and direction of the organisation]'. Quality *assurance* is: 'all those planned and systematic actions necessary to provide *adequate confidence ...*' (our italics). This usage is in accordance with ISO 8402. Quality *assessment* is not defined there and will be taken to mean a systematic examination to determine whether quality activities comply with planned arrangements and whether the 'product' (the educational process) is implemented effectively and is suitable to achieve objectives.
- [2] 'Study programme' will be used in this paper to designate a course *and* its organisational setting (*studierichting*). The term 'discipline' will be used here in a broad sense, meaning each area or cluster distinguished in the QA procedures.
- [3] Much of the processing of the questionnaire has been done by Jan Donders, whom the authors wish to thank for his efforts.
- [4] The sectors and response rates are: agriculture (63%), technical/engineering (68%), natural sciences (66%), health (76%), economics (48%), behavioural and social (58%), languages and cultural (50%), education (75%), law (45%).
- [5] In the second round, this area of knowledge will be covered by many specialist VCs, instead of the two broad VCs that operated in the first round. A sense of dissatisfaction has apparently been noted in the VSNU as well.
- [6] Valid responses: all responses, minus 'not applicable' or left blank, but including 'don't know's'.
- [7] In the study programme committee, the academic staff and students are represented equally. It is the most important advisory body for the faculty council and board regarding educational matters.
- [8] 'Studiability' is a neologism introduced in a report on common problems and

solutions to lower the dropout rate in Dutch higher education (Wijnen *et al.*, 1992).

REFERENCES

- BALDRIDGE, J.V., CURTIS, D.V., EDKER, G. & RILEY, G.L. (1977) Diversity in Higher Education, *Journal of Higher Education*, 58, pp. 367–388.
- BECHER, T. (1989) *Academic tribes and territories. Intellectual enquiry and the cultures of disciplines* (Milton Keynes, SRHE/Open University Press).
- BIRNBAUM, R. (1988) *How colleges work: the cybernetics of academic organisation and leadership* (San Francisco, Jossey-Bass).
- CHILD, J. (1984) *Organisations* (London, Harper & Row).
- CLARK, B.R. (1983) *The Higher Education System* (California, University of California Press).
- COOLEY, W.W. & BICKEL, W.E. (1986) *Decision-Oriented Educational Research* (Boston, Kluwer-Nijhoff Publishing).
- CORNETT, P. (1987) *Accreditation-Related Self-Study in Allied Health and Nursing Education* (Rutgers University).
- FREDERIKS, M.M.H., WESTERHEIJDEN, D.F. & WEUSTHOF, P.J.M. (1993) *Interne zorg en externe prikkel. Onderwijskwaliteitszorg in Nederlandse universiteiten en hogescholen* (Zoetermeer, SDU).
- GOEDEGEBUURE, L.C.J. (1992) *Mergers in Higher Education: A Comparative Perspective* (Utrecht, Lemma).
- GROOT, T.C.L.M. (1988) *Management van universiteiten: Een onderzoek naar de mogelijkheden voor doelmatig en doeltreffend universitair bestuur* (Groningen, Rijksuniversiteit Groningen).
- HARRIS, E. (1984) *A multi-case study of the self-study component of the regional accreditation process: Identifying influential factors* (Syracuse University).
- INSPECTIE HOGER ONDERWIJS (1992) *De bestuurlijke hantering van de resultaten van de externe kwaliteitszorg 1989 in het wetenschappelijk onderwijs, 1992–98 Report* (Zoetermeer, Ministerie van Onderwijs en Wetenschappen).
- JONGBLOED, B.W.A. & WESTERHEIJDEN, D.F. (1994) Indicators of quality in Europe: performance indicators and quality assessment in European higher education, *New directions of institutional research* (forthcoming).
- KELLS, H.R. & KIRKWOOD, R. (1979) Institutional Self-Evaluation Processes, *Educational Record*, 60 (1).
- LIESHOUT, R.H. (1989) De harde kern van de zaak, *Acta Politica*, 4, pp. 461–482.
- LIESHOUT, W.C.M., VAN (1992) Visitaties: Een schakel in de kwaliteitszorg, in: W. JOCHEMS & A. PILOT (Eds) *Visitaties: Géén stilte na de storm* (Groningen, Wolters-Noordhoff).
- MINISTRY OF EDUCATION & SCIENCE (1985) *Hoger onderwijs: Autonomie en kwaliteit* ('s-Gravenhage, Staatsuitgeverij).
- MINTZBERG, H. (1979) *The structuring of organisations: A synthesis of the research* (Englewood Cliffs, Prentice-Hall).
- SCHEERENS, J. (1987) *Het evaluerend vermogen van onderwijsorganisaties* (Enschede, Universiteit Twente).
- SCRIVEN, M. (1973) The methodology of evaluation, in: B. R. WORTHEN & J. R. SANDERS (Eds) *Educational evaluation, theory and practice* (Wadsworth, Wadsworth Publishing Company).

- SCRIVEN, M. (1974) *Evaluation in education: current applications* (Berkeley, McCutchan).
- VROEIJENSTIJN, T. & ACHERMAN, H. (1990) Control-Oriented versus Improvement-Oriented Quality Assessment, in: L. C. J. GOEDEGEBUURE, P. A. M. MAASSEN & D. F. WESTERHEIJDEN (Eds): *Peer Review and Performance Indicators* (Culemborg, Lemma).
- WEUSTHOF, P.J.M. (1989) Zelfevaluatie: wat, hoe en waartoe?, *Tijdschrift voor Hoger Onderwijs*, 7 (4), pp. 148–159.
- WIJNEN, W.H.F.W., DE BIE, D., BROUWER, O.G., RUIJTER, C.T.A., VOS, P. & WOLFHAGEN, H.A.P. (1992) *'Te doen of niet te doen?'* *Advies over de studeerbaarheid van onderwijsprogramma's in het hoger onderwijs* (Zoetermeer, Ministerie van Onderwijs en Wetenschappen).