
James Pounder

Abstract

Organisational performance assessment in higher education tends to employ criteria which reflect a view of effectiveness which is implicit and therefore untested (Pounder, 1997). However, the Competing Values Model or Framework (Quinn and Rohrbaugh, 1981, 1983) has been proposed as “a general paradigm of organisational effectiveness” This study examined the applicability of the Model to higher educational organisations. A set of scales suitable for the valid and reliable self assessment of the effectiveness of higher educational organisations in Hong Kong was developed. The scale development method was designed to maximise validity and reliability and to test the relevance of the Competing Values Model to Hong Kong higher education. The scales resulting from the study contained four of the Model’s nine original effectiveness dimensions which indicated possible core dimensions of an effectiveness model for higher educational organisations in Hong Kong.

Introduction

Despite the worldwide pressure on higher educational organisations to provide evidence of effective performance, organisational performance assessment in higher education continues to employ criteria which reflect a view of effectiveness which is implicit and therefore untested (Pounder, 1997). Higher education needs the solid theoretical base for organisational performance assessment that a generally accepted model of effectiveness could provide. The Competing Values Model of Organisational Effectiveness (Quinn and Rohrbaugh, 1981, 1983) may meet that need.

The Competing Values Model has been viewed as a “general paradigm” of organisational effectiveness (Quinn and Rohrbaugh, 1981). Its general paradigmatic status is based on the fact that the Model makes explicit major perspectives on organisational effectiveness taken by acknowledged experts in the field. Furthermore, the extent to which the Model has been employed in organisational and management studies provides support for that status. The Competing Values Model has provided an analytical framework for over 40 studies (Pounder, 1997).

This paper describes a study which sought to test the applicability of the Competing Values Model (Quinn and Rohrbaugh, 1981, 1983) to higher educational organisations. The study was conducted in the Hong Kong higher educational system and was completed in 1997. At the time of the field work, the Hong Kong higher education system comprised only nine accredited educational organisations which facilitated a high degree of participation in the study. Thus, Hong Kong provided a suitable base for an initial examination of the applicability of the Competing Values Model of Organisational Effectiveness to higher education. The relevance of the Model to higher educational organisations was tested through a method designed to produce scales capable of the valid and reliable self rating of organisational effectiveness. The scales were based on the effectiveness dimensions contained in the Competing Values Model.

Method

What follows is a summary of the method employed in the study to develop the rating scales. The method of scale development was a version of the procedure for constructing Behaviourally Anchored Rating Scales (BARS) (Smith and Kendall, 1963), modified to maximise validity and reliability.

1 Full details of the method can be obtained from the author.
Generation of Examples

A pilot study, involving a small number of senior academics and administrators in Hong Kong higher education, gauged the feasibility of generating behavioural examples for BARS in sufficient numbers for the main study to be feasible. Participants were requested to provide examples of good, average and poor organisational behaviour, for each of the nine effectiveness dimensions (i.e., Productivity-Efficiency, Quality, Cohesion, Adaptability-Readiness, Information Management-Communication, Growth, Planning-Goal Setting, Human Resources Development, Stability-Control) contained in the Competing Values Framework (Quinn and Rohrbaugh, 1981, 1983). Definitions of these effectiveness dimensions are contained in Appendix A. These definitions took the descriptions employed in an earlier experiment by Edwards (1986) conducted in a commercial context, and slightly modified them where necessary for a higher educational setting.

The results of the pilot study indicated the feasibility of the main study and all the nine accredited higher educational institutions in Hong Kong were then invited to participate in developing the instrument. Seven institutions accepted the invitation. Thirty two senior academic and administrative staff in Hong Kong higher education agreed to provide behavioural examples as in the pilot study. The rationale for inviting only senior staff was the need for participants to occupy positions affording them an overview of organisational performance in order for them to provide examples of behaviour at the organisational level. A total of 592 examples were generated.

Screening of Examples

The behavioural examples generated above were subjected to an initial screening by a panel of four judges at senior institutional management level (i.e., Vice-President, Dean of Business Faculty, Institutional Consultant/Former Associate Vice President (Academic) and Head of Management Department) in one of the accredited institutions of higher education in Hong Kong. Panel members were advised to screen out any examples which fell into one or more of the following categories:

(a) did not describe behaviour at the organisational level
(b) expressed sentiments similar to another example perceived to be better constructed
(c) was a description of an intangible trait rather than a tangible illustration of behaviour
(d) was worded ambiguously
(e) described behaviour associated with more than one dimension of effectiveness.

Retranslation

A randomised list of the examples remaining after screening and a list of the nine effectiveness dimensions with descriptions enabled 162 academic and administrative staff in the seven participating institutions to undertake retranslation. This involved the reallocation of examples to relevant dimensions. A 60% agreement criterion (i.e., for each example, 60% of participants in the retranslation exercise agreeing that the example related to a particular effectiveness dimension) was employed which permitted the retention of sufficient examples to form scales in all the nine effectiveness dimensions of the Competing Values Framework. Bernardin et al (1976) have recommended the 60% criterion based on a comprehensive examination of the effects of procedural variations on the psychometric properties of BARS scales.

BARS Scaling

One hundred and thirty academic and administrative staff in the seven participating institutions engaged in the BARS scaling. BARS scaling involved each behavioural example being given a numerical rating by each participating staff member. This rating reflected the participant’s perception of the level of performance in a particular effectiveness dimension conveyed by the example in question. Thus, for instance, an example judged by a staff member to represent good performance in a particular effectiveness dimension was given a higher numerical rating than an example considered by that same staff member to represent average or poor performance. The objective of the scaling was to form rating scales in each of the nine effectiveness dimensions from only those examples with the lowest standard deviations.
consistent with retaining enough examples to form the scales. The BARS procedure required that those examples then be located as scale anchors according to the mean of their numerical ratings.

Landy and Farr (1980) concluded that the decision rules for scaling are often arbitrary and a survey of literature for the Hong Kong study did reveal variations in the number of scale points used for the numerical rating of examples and the standard deviation criteria employed for scaling. It was felt that numerical scales exceeding five points would tend to make the rating of behavioural examples too complex by requiring participants to make excessively fine judgments on levels of performance represented by the examples. Additionally, BARS literature contains a study which developed scales capable of valid measurement based on a five point scaling of anchors and a standard deviation of 1.00 or less (Dickinson and Zellinger, 1980). In developing the BARS for the Hong Kong study, a five point scale for rating behavioural examples was employed together with a more stringent standard deviation criterion than in the Dickinson and Zellinger study, namely 0.9 or less. This standard deviation criterion was applied to ensure that the scales developed in the Hong Kong study were at least comparable, and potentially superior, to Dickinson and Zellinger’s scales.

**Multidimensional Scaling**

Valid rating requires that rating scales be logically ordered (scalable). The scalability notion, originating with Guttman (1944), requires that scales produce response patterns “in which endorsement of the item reflecting the extreme position results also in endorsing all items which are less extreme” (Emory and Cooper, 1995, p. 223). Hence, scales should be ordered in such a way that a rater will not be faced with the dilemma of finding two anchor descriptions on the same scale which seem equally representative of a subject’s performance where one description purports to depict good performance and the other, poor performance.

Multidimensional scaling was employed in the Hong Kong study as a means of validating the ordering of scale anchors. Dissimilarity matrices were prepared for each of the nine effectiveness dimensions to enable pairwise dissimilarity judgments to be made for all the behavioural examples retained as scale anchors after BARS scaling had been completed. Dissimilarity judgments were made using a five point numerical scale. The selection of a scale confined to five point was to avoid the need for participants to make excessively fine judgments on degrees of dissimilarity between pairs of behavioural examples. For example, a participant might record a score of 1 in the appropriate cell of the matrices when a pair of behavioural statements was judged to be very similar and in the case of a pair of examples judged to be very dissimilar, a score of 5 might be given. It was recognised that multidimensional scaling requires careful attention to the task on the part of participants. This led to the selection of a panel of twenty five judges to undertake the exercise on the basis that:

i) they occupied senior academic or administrative roles in Hong Kong higher education, or;

ii) they had an academic background in organisational analysis and were currently involved in research or teaching in an area related to the study.

It was reasoned that a validly ordered scale would exhibit a pattern of anchor ordering such that anchors representing clear differences in level of performance according to BARS scaling would be judged to be clearly different according to multidimensional scaling. It followed also that anchors representing clearly similar levels of performance should be judged to be clearly similar. Correlation analysis was used to compare the ordering of anchors produced by the BARS and multidimensional scaling methods. Those scales which were judged to possess relatively weak correlations between the rank ordering of examples produced by both scaling methods, were eliminated from the study.

**Multitrait-multirater Analysis**

Valid measurement requires that measurement scales be unidimensional. This means that descriptive scale anchors should be statements capturing the content of the trait, or in the case of the Hong Kong study the effectiveness dimension, to which they relate and only that trait or effectiveness dimension. This aspect of
validity is frequently examined with reference to Campbell and Fiske’s (1959) multitrait-multimethod analysis. The multitrait-multiplicative variant of Campbell and Fiske’s analysis was employed in the Hong Kong study given that the analysis remains current as a tool for gauging the construct validity (unidimensionality) of human performance measurement scales (Henderson et al., 1995; Kinicki et al., 1985; Lascu et al., 1995; Schreisheim and Eisenbach, 1995; Spreitzer, 1995; Sullivan, 1995). Furthermore, psychometricians continue to confirm the efficacy of multitrait-multimethod analysis as a thorough approach to validating measurement employing rating scales (Cronbach, 1990; Murphy and Davidshofer, 1991; Gregory, 1992).

Multitrait-multiplicative analysis requires the production of a matrix which presents all intercorrelations when each of a number of traits (in the Hong Kong study, organisational effectiveness dimensions) is measured by more than one source of ratings. The pattern of intercorrelations indicates the presence, or absence, of construct validity. For the purpose of multitrait-multiplicative analysis, the scales remaining after the scaling stage of the procedure were formed into an organisational effectiveness self rating instrument and completed by 151 academic and administrative staff at various levels of responsibility within the seven participating institutions. Two categories of respondents were identified: Assistant/Associate Professor and Others. The first group was viewed as largely homogeneous in job nature, with a primary emphasis on teaching and research. The second group was considered less homogeneous than the former but commonly, those falling into this category, tended to perform organisational roles with a high administrative content. This categorisation and the spread of responses across institutions permitted the development of multitrait-multiplicative matrices for four out of the seven participating institutions (Pounder, 1997).

**Tests of Reliability**

The inter-rater reliability and test-retest reliability of rating employing each of the scales was gauged against the Nunnally and Bernstein (1994) recommended reliability coefficient of around 0.70. Scales which did not meet the recommended reliability coefficient were eliminated.

**Results and Discussion**

The method of scale development resulted in rating scales in four of the nine effectiveness dimensions contained in the Competing Values Model (Quinn and Rohrbaugh, 1981, 1983), namely, Information Management-Communications, Planning-Goal Setting, Productivity-Efficiency and Cohesion. Each of these scales were found to be scalable, and capable of valid rating measured in terms of convergent and discriminant validity. Given the nature of the BARS procedure which requires experts to provide the scale anchors, the scales were also content valid. Additionally, each of those scales met Nunnally and Bernstein’s (1994) threshold requirement for reliability. The four scales are reproduced in Appendix B. The method of scale development also served to establish the applicability of Information Management-Communications, Planning-Goal Setting, Productivity-Efficiency and Cohesion effectiveness dimensions to Hong Kong higher education. Given the rigour of the scale development method and the high degree of participation in the study, these effectiveness dimensions are possible core dimensions of a model of organisational effectiveness applicable to higher education in Hong Kong.

The opportunity now exists to build on that core. For example, the method of instrument development employed in the Hong Kong study could be replicated in higher educational systems in a number of countries with the possible result of moving research towards the development of a generally agreed model of organisational effectiveness in higher education. From an organisational performance assessment viewpoint, a generally agreed model of organisational effectiveness in higher education would anchor performance assessment to a sound theoretical base enabling, for example, comparative assessments of the effectiveness of higher educational organisations to be made. Equally, a generally agreed effectiveness model would mean that such comparisons need not be confined within national boundaries.
Appendix A

- **Productivity-Efficiency**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent to which it is concerned with the quantity or volume of what it produces and the cost of operation.

- **Quality**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent to which it is concerned with the quality of what it produces.

- **Cohesion**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent to which it is concerned with staff morale, interpersonal relationships, teamwork, and sense of belonging.

- **Adaptability-Readiness**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent of its ability to readily alter or adapt its structure, programmes, courses etc., in response to changing demands. In other words, the extent of the organisation’s readiness to adapt to change.

- **Information Management-Communication**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent of its ability to distribute timely and accurate information needed by its members to do their jobs.

- **Growth**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent of its ability to secure external support, acquire resources, and increase its capabilities.

- **Planning-Goal Setting**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent to which it is responsive to the individual needs of its staff. It also has to do with the extent to which the institution facilitates participation in decision making. Additionally, this aspect is concerned with behaviour relating to the hiring, training and development of staff.

- **Stability-Control**: This aspect of an organisation’s performance has to do with behaviour that reflects the extent of its ability to control the flow of work, to direct the behaviour of its members, and to maintain the organisation’s continuity particularly under periods of pressure or threat.
Performance Dimension: Information Management - Communication

Definition: This aspect of an institution’s performance has to do with behaviour that reflects the extent of its ability to distribute timely and accurate information needed by members to do their jobs.

Typically, in this institution, one would expect mechanisms to have been introduced for the express purpose of cascading information systematically from top to bottom of the organisational hierarchy.

Typically, in this institution, one would expect a management information provision unit to have been established which is constantly consulting information users on their present and future needs.

Typically, in this institution, one would expect minutes of governing body meetings (e.g., Board of Governors/Council/Academic Board) to be made available to all staff, and actively circulated to those who need them.

Typically, in this institution, one would expect information bulletins from management, and meetings, to focus primarily on developments that have already taken place, as opposed to developments in the pipeline.

Typically, in this institution, one would expect there to be both formal and informal channels for information but information provision to be not always timely.

Typically, in this institution, one would expect there to be widespread use of the ‘confidential’ and ‘restricted’ stamp on documents and reports.

Typically, in this institution, one would expect information produced centrally to be consigned to the waste bin frequently by recipients because it is thought to serve no useful purpose.

Typically, in this institution, one would expect information provision to be ‘ad hoc’ in the sense of being provided when requested if one happens to know that it is available and the relevant party to contact.

Typically, in this institution, one would expect little or no publicity to be given to major developments such as the setting up of a new unit or the introduction of a new facility.

This space below is provided for raters to write down their own example (optional — see rater’s instructions):

______________________________________________________________________________________________

______________________________________________________________________________________________

Numerical Rating
Performance Dimension: Planning - Goal Setting

Definition: This aspect of an institution's performance has to do with behaviour that reflects the extent of its ability to set goals and objectives and systematically plan for the future.

Typically, in this institution, one would expect the institutional plan to be coherent, in the sense of moving from mission statement to broad aims and then to specific objectives, and action oriented, in the sense of providing a framework which allows each lower level unit to define its own objectives and action plans.

Typically, in this institution, one would expect long term goals to have been established that are consistent with the institution's traditions and the likely future of society, and mechanisms for regular review to have been set up.

Typically, in this institution, one would expect an institutional plan, comprising mission, goals and strategies to be revised and updated each year and circulated widely amongst members of staff.

Typically, in this institution, one would expect academic and administrative unit heads to prepare their own budgets which reflect their plans for the coming year.

Typically, in this institution, one would expect a planning committee to have been established, comprising members of the senior management team, which regularly reviews the institution's mission, objectives and strategies.

Typically, in this institution, one would expect there to be an approximate plan or view on future direction but with implementation not fully worked out.

Typically, in this institution, one would expect management to exhibit a reluctance to deal with open ended aspects of planning, preferring to plan within a framework which limits strategic options.

Typically, in this institution, one would expect management not to have given direction on priority areas for the present and future.

Typically, in this institution, one would expect powerful members of the senior management team to display little understanding of, or interest in, long term development. Consequently, one would expect such committees as are established to deal with change to have no obvious goals or functions.

Typically, in this institution, one would expect different committees to deal with different planning related activities in an uncoordinated way.

This space below is provided for raters to write down their own example (optional — see rater’s instructions):

______________________________________________________________________________________________

______________________________________________________________________________________________

Numerical Rating
Performance Dimension: Productivity-Efficiency

Definition: This aspect of an institution’s performance has to do with behaviour that reflects the extent to which it is concerned with the quantity or volume of what it produces and the cost of operation.

5

Typically, in this institution, one would expect there to be regular meetings at all levels in the organisation devoted to improving productivity and efficiency.

4.5

Typically, in this institution, one would expect that rationalisation of the organisation (e.g., grouping of departments into faculties or schools) and establishment of budget centres, have been carried out with a view to improving productivity and efficiency.

4

Typically, in this institution, one would expect departments and individuals to have been provided with incentives to use resources efficiently.

3.5

Typically, in this institution, one would expect periodic reviews of support units to be made with a view to establishing the extent to which the units give value for money. One would also expect appropriate action to be taken in cases where value for money is not established.

3

Typically, in this institution, one would expect that in new programme/course design, a great deal of emphasis would be placed on demand for the programme/course and economy in the use of resources.

2.5

Typically, in this institution, one would expect it to have been made clear to all unit heads and above that productivity - efficiency is a major criterion used to judge the worth of their decisions.

2

Typically, in this institution, one would expect evening and extension programmes/courses to be approved on the understanding that they cover costs.

1.5

Typically, in this institution, one would expect there to be a constant drive to achieve low unit costs.

1

Typically, in this institution, one would expect resources to be under-utilised in the evenings or weekends.

Typically, in this institution, one would expect that, so long as externally established norms are met, little attention is paid to efficiency.

Typically, in this institution, one would expect there to be a notable absence of mechanisms for assessing productivity and efficiency.

This space below is provided for raters to write down their own example (optional — see rater’s instructions):

_______________________________________________________________________________________________
_______________________________________________________________________________________________

Numerical Rating
Performance Dimension: Cohesion

Definition: This aspect of an institution’s performance has to do with behaviour that reflects the extent to which it is concerned with staff morale, interpersonal relationships, teamwork, and sense of belonging.

Typically, in this institution, one would expect senior management to have taken positive steps to create a climate in which employees at all levels are made to feel valuable members of the institution.

Typically, in this institution, one would expect staff to regularly refer to their sense of commitment to the institution.

Typically, in this institution, one would expect mechanisms to exist for staff to share problems and to work together.

Typically, in this institution, one would expect the senior management team to visit academic and administrative units regularly and talk freely and informally with members of staff.

Typically, in this institution, one would expect there to have been an effort either to foster allegiance to parts of the organisation in the face of growth or to limit expansion so as to maintain a sense of belonging.

Typically, in this institution, one would expect staff to demonstrate greater allegiance to the parent faculty or department than to the institution as a whole.

Typically, in this institution, one would expect senior management to express a commitment to the maintenance of staff morale but not to do sufficient to ensure that people at all levels have a sense of belonging.

Typically, in this institution, one would expect there to be frequent conflicts between the centre and sub-units, between academic and administrative units, between one academic or administrative unit and another, and between individuals.

Typically, in this institution, one would expect feelings of loyalty and sense of belonging to be undermined by an approach to human relations issues (e.g., to contract renewal) which generates insecurity.

Typically in this institution, one would expect there to be a general lack of informal contact amongst staff members demonstrated by inadequate staff common room life and too many closed office doors.

This space below is provided for raters to write down their own example (optional — see rater's instructions):

______________________________________________________________________________________________

______________________________________________________________________________________________ Numerical Rating
References


