

ADDING VALUE TO SCHOOLING:

THE USE OF ASSESSMENT CRITERIA

Dr Clive B. Kings
Educational Development Consultant
Institutional Planning & Development Directorate
Ngee Ann Polytechnic

Dr Rosalind Y. Mau
Teaching Fellow
Division of Policy and Management
Studies
National Institute of Education
Nanyang Technological University

Paper presented at the ERA-AARE Joint Conference, November 25-29 1996,
Educational Research: Building New Partnerships, Singapore Polytechnic,
Singapore

INTRODUCTION

The nature of assessing student performance will to a large extent influence how students learn (Ramsden, 1992; Kings, 1993). How students perform on assessment tasks will depend on how well they understand what is required of them.

The authors assert that student self-motivation has an enormous impact on student learning irrespective of home and school backgrounds. The means of enhancing self-motivation is through students having clear benchmarks or criteria by which they can judge their own performance.

In this paper a case study illustrates the usefulness of assessment criteria on individual performance. The underlying assumption of this case study is that if assessment criteria are embodied in the planning of learning, they will ipso facto provide direction for improving performance. The following questions are addressed:

What are the differences between use and non-use of specific criteria on markers' assessments?

What are the differences between the use and non-use of specific assessment criteria on student performance ?

What are some benefits of and barriers to using assessment criteria?

BACKGROUND

In everyday life individuals naturally engage in assessment. They check whether they are correctly attired by assessing if their clothes are branded, are suitable for the climate, are appropriate to the task or are worn by peers. They judge the appropriateness of their attire by different criteria. It is apparent that individuals will give varying salience to each criterion depending on how much they are influenced by such things as advertisement, peer approval or comfort. In a similar way individuals set criteria for the way they assess their performance on cooking meals or doing their hobbies. The key point is that individuals choose and set criteria naturally. Individuals often set quite demanding criteria for themselves in their hobbies, and reach relatively high standards. In such cases the criteria are adopted and internalised by individuals, and are not prescribed from outside.

These examples of individuals choosing criteria for assessment are in sharp contrast to what often happens in schools. Some students do consistently well on assessments because they seem to know what is required. The criteria by which students can judge their performance are often not made explicit to them by their teachers. Other students guess at assessments and are unclear about how to achieve good grades.

These uncertainties often lead to insufficient learning and poor assessment results. Some students give up the guessing game and are doomed to failure.

Thus in the classroom assessment is not a natural process. The reasons for this dissonance include:

The criteria used in assessment are often not made clear to students.

The criteria underscoring the learning strategies of the classroom are at variance with those underscoring the assessments, e.g. when only single word or short answers are required.

The criteria of success in the home may differ from those of the school with respect to co-operation/competition and language.

The criteria for success may relate to only one type of assessment and not accommodate different learning modes.

Since students try to naturally make sense of what they are learning, teachers should be clear about their achievement criteria. In this way students can receive feedback when they engage in checking their learning by themselves and with their peers or their teachers. Teachers can empower students when they provide explicitly identifiable criteria.

THE IMPORTANCE OF ASSESSMENT CRITERIA: A CASE STUDY

The purpose of this case study was to investigate the impact of assessment criteria on student learning and performance.

Subjects

The sample in this case study consisted of 46 Heads of Departments (HoDs) from schools. In Singapore HoDs are the instructional leaders of schools who have a major role in the maintenance of quality assessment in their departments. Their major duties include being responsible for planning the curriculum, vetting examinations, monitoring teaching and evaluating programs. There were 40 HoDs (15 males and 25 females) from primary schools throughout Singapore and six (one male and five females) from the Maldives. The HoDs' mean length of teaching experience was 16.7 years (Standard deviation 8.4). They were in charge of the following departments: Science (8), Tamil (3), Mathematics (10), English Language (12) and Physical Education/Extracurricular Activities (13).

Procedures

Modellers

Two groups each of 12 HoDs were given 20 minutes to make a model of an animal using Playdoh, a malleable material which can be modelled into different shapes. These students were referred to as modellers. One group of modellers was given the assessment criteria specified below while the other group of modellers was not given any criteria. The criteria specified were:

modelled a quadruped
have used at least three colours
smoothness
clarity of the face
proportionality
Markers

A total of 22 HoDs were assigned to two groups of 11 a markers. The markers had 20 minutes to assess the modelled animals. One group of markers had to individually devise their assessment criteria and the other group of markers was provided with the following criteria:

modelled a quadruped
5has legs of equal length and width
3has 4 legs but not uniform legs
0does not have 4 legs
have used at least three colours
5has at least three colours
0does not have three colours
smoothness
5smooth all over
3relatively smooth
0rough
clarity of the face
5highly detailed
3medium detail
0indistinguishable detail
proportionality
5well proportioned
3reasonably proportioned
0inappropriate proportions

In the second part of the study, the 46 HoDs were asked to respond to

four open-ended questions after they had completed the prescribed tasks. They were allocated 15 minutes to respond to the value and usefulness of assessment criteria. The four questions were:
How did you feel about your role, either as a modeller or a marker?
After doing this exercise about using assessment criteria, do you think that teachers should clearly state criteria in their assessment?
How will you use what you have learned from this exercise?
Are there any barriers to using assessment criteria?

Their responses were coded, collated and analysed.

RESULTS OF THE CASE STUDY

The impact of using criteria on student learning and performance have been investigated for groups of modellers and markers who use criteria and groups who do not use criteria. The following table shows the mean scores and standard deviations of markers with and without criteria as well as modellers with and without criteria. For example, when modellers with criteria are assessed by markers with criteria, the mean score is 17.2. In contrast, modellers without criteria who are assessed by markers with criteria, the mean score is 10.5.

Mean Scores and Standard Deviations of Modellers With/Without Criteria by Markers With/Without Criteria

| | Markers with criteria | | Markers without criteria | |
|----------------------------|-----------------------|----------|--------------------------|----------|
| | Mean | St. Dev. | Mean | St. Dev. |
| Modellers with criteria | 17.2 | 3.63 | 17.4 | 3.73 |
| Modellers without criteria | 10.5 | 5.07 | 17.3 | 3.63 |

N=46

When the markers were given specific assessment criteria, modellers without criteria had a mean score of 10.5 and did not perform as well as modellers with criteria who had a mean score of 17.2.

Modellers who knew what was expected through the explicit criteria given achieved a higher mean score. The modellers without criteria had a lower mean score because they produced their models according to what they thought was required. For example, modellers were given only two colours of Playdoh. Those modellers who followed the criterion to use three colours achieved the third colour by mixing two colours. It would have been by chance that modellers without the criteria would have used three colours. So when the criteria were clear there was a higher level of achievement.

A majority of responses from an open-ended question about how modellers without criteria felt during the modelling exercise indicated that they

felt lost or confused about how to approach the task. A minority enjoyed the freedom of not being constrained by specific modelling requirements, but felt unfairly assessed by markers with criteria since their strong points were not valued. The modellers with criteria generally felt comfortable and confident.

2When the markers were given specific assessment criteria, modellers without criteria had a greater range of scores with a standard deviation of 5.1 compared to modellers with criteria with a standard

deviation of 4.6.

Modellers were more focused on completing the task based on the assessment criteria and therefore had a narrower range of variation in the levels of achievement.

After the modelling and marking exercise markers with criteria stated they felt more comfortable in the allocation of marks because they had guidelines with specific values to be given to certain qualities.

3When markers did not have criteria the mean scores for modellers were similar for modellers with and without criteria, i.e. having mean scores of 17.3 and 17.4.

Since the markers did not have criteria on which to judge, they chose to mark what they thought was important. For example, some markers might have chosen to accentuate creativity while others might have chosen to accentuate proportionality. The effect is regression towards the mean as shown by similar mean scores for both sets of markers.

Open-ended responses showed that markers without criteria felt lost or confused. They developed the models according to the criteria they each considered to be important.

4When markers did not have criteria to judge a piece of work, a modeller received a score of 6.4 from one marker and a score of 18.1 from another marker.

All markers should have criteria for consistency of marks and to give modellers a comparable message about the value of their work.

Moreover, modellers were asked what their response would be if they had two different scores for the same piece of work. They indicated a feeling of confusion about how they had really performed and a lack of clarity about how to improve. Two other open-ended questions asked HoDs whether they felt that criteria were important in assessment and whether they would use assessment criteria.

The respondents unanimously affirmed the use of criteria when setting assessments, endorsing the fairness of the procedure of letting students know teachers' expectations about assessment tasks. The respondents unanimously indicated their intention to use assessment criteria in the future. Most respondents perceived no barriers to the implementation of assessment criteria in their departments, even though

there might be some resistance from a few members of staff. This minority staff resistance might be accounted for by the time required to develop and specify criteria; the feeling that teachers were giving students clues; or the feeling that criteria might constrain creativity.

DISCUSSION

The results of the case study endorse the key idea that the quality of student learning can be largely determined by the nature and quality of the assessment tasks students are set (Ramsden, 1992; Kings, 1993). When students have an understanding of the assessment criteria, they are clear about how to approach their task, are less anxious and feel more confident. This is understandable if we acknowledge that individuals in society are continuously assessing themselves - setting themselves benchmarks against which to judge how effectively a task is being done, and revising their approaches to improve performance.

Criteria to Improve Achievement

Responses from the open-ended questions in the case study showed that criteria clarify what is expected of learners so they can better complete assessment tasks. When criteria are not specified, students might think that criteria such as the following are the important ones: more is better, neatness and presentation are critical, regurgitation of given ideas, one-solution interpretation, or artistic flair. The students' inferred criteria may mislead them to meet the 'wrong' criteria and so perform badly.

If the objective is to measure how well students understand science processes then the criteria should not solely relate to the number of facts learned. Students who are clear about the assessment criteria can perform better. They have a clearer direction of what needs to be done and are less confused on how to achieve higher scores. The clarification of criteria ensures that assessment is developmental in the sense that it helps students learn rather than only providing summative judgment by the teacher.

Criteria for Consistency of Marking

When modellers had the criteria and their works were judged by markers without criteria the range of scores for one individual modeller ranged from 11 to 25. The wide range of scores for the same piece of work was because markers without criteria set their own values and criteria of what they thought were important. One modeller felt 'This is so confusing.' The consequence for learners in general is that if they are unclear about how the objectives of a lesson can be met, they will

have more difficulty improving their performance levels.

Results of the study show that markers require guidelines or criteria that will enable them to have some consistency in their marking. Teachers may be unclear about their own criteria and place different values on what is important in an assignment. If the criteria are not made explicit, teachers may mark the same work differently.

A way to avoid disparity of marks on the same assignment is through inter-rater reliability, that is, through collaboration. Teachers should identify and clarify the criteria by which a piece of work will be judged prior to giving the assignment. Before marking all the assignments they should compare their marks on agreed criteria, discuss discrepancies, review the interpretation of the criteria and how assignments should be judged. This is known as the process of moderation.

SOME ASSESSMENT GUIDELINES

The authors assert that since students naturally assess themselves, they should be empowered in their own learning by self- and peer-assessment (Kings, 1993; Mau, 1996). Self- and peer assessment involves the following processes:
identification of specific criteria for an assessment task,
clarification of assessment criteria, and
judgment of the extent to which criteria have been met, either by the individual student, self-assessment, or by the individual's peers,
peer-
assessment.

Identification of Specific Criteria

The criteria for any assessment task need to be spelled out so that students are clear about what they are expected to learn. Students should not have to guess at which criteria have to be met to achieve higher grades.

Ideally classroom assessment is fundamental to learning because it provides feedback to students on how much they have learned, and it guides teachers about how well they have taught. Feedback can be provided based on certain benchmarks or criteria. Judgments are made by teachers about how well students have performed on an oral assessment by articulation of assessment criteria such as: 'Use of appropriate language for professional scientists.'; 'Posing of an appropriate question for the audience to discuss.' ; or 'Use of appropriate visual aids.' Thus criteria provide directions for learning and reduce confusion.

Judicious Use of Criteria

Teachers should be judicious in choosing and developing criteria. They should be conscious of the students as learners in two ways. First, teachers should ensure that criteria are attainable, for example, the target time for a 100 metre dash. If students perceive they can never achieve the criteria they are more likely to give up learning, lose self-esteem and be more inclined to rebellious behaviour. Second, teachers should try to ensure that criteria are chosen that do not stifle creativity. For example, in creative dance the qualities of weight, time, space and flow provide an individual the opportunity to use creative flair.

Promoting Autonomous Learning

The development of student learning can be fostered by progressing from a stage where the teacher spells out the assessment criteria to students, to one in which the teacher negotiates the criteria with students, to one in which students can identify criteria either individually or in a group themselves. Thus students are encouraged to move from being dependent learners to being independent learners which replicates the situation in adult life and in the workplace.

Clarification of Criteria

Students need the opportunity of clarifying criteria so that they are clear about what should be achieved. This can be done through group discussion and through discussion with the teacher. This is an essential part of the learning process.

A caveat to the use of criteria is that teachers should avoid the overstatement of criteria which is susceptible to a dysfunctional mechanical approach to their use, since the use of criteria, a means of enhancing learning, can become an end in itself.

Judging Criteria Have Been Met

Once the criteria for assessment have been established teachers will need to provide a framework within which students will be able to assess for themselves how far they have met each criterion. Likewise, students will need a framework for the development of peer-assessment. In some cases, say in the development of a community project, the assessment may involve judgments being made by a number of participants against a number of agreed criteria. This situation is realistic preparation for students who in the future will have to address new situations and judge their own success in coping with change by establishing their own criteria.

In summary, as Kings (1993) and Mau (1995) have consistently stressed,

criteria should be clarified so that students are clear about what they are expected to learn. Lecturers in teacher education programs should engage in quality assessment planning, development of appropriate assessment techniques and procedures, use of assessment criteria and application of self- and peer-assessment processes. These will provide examples of good practice for their students. Furthermore, a periodic review of teacher education curricula with respect to these same assessment elements promotes value-added learning in the student community. As student teachers enter the workforce they can influence the lifelong learning of their students.

REFERENCES

- Kings, C. B. (1993). Effective assessment for improvement of student learning. *Research and Development in Higher Education*, 14, 144-150.
- Mau, R. (1995). Assessment in the classroom: The value of achievement targets. *Educational Theory and Practice*, 17(1), 59-64.
- Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.

Presenters

Dr Clive Kings is currently Educational Development Consultant at Ngee Ann Polytechnic, Singapore. He has worked in England, Australia and Singapore. His Ph.D. on higher education learning and assessment is from the Centre for the Study of Higher Education, The University of Melbourne. He has worked on examination boards and conducted programs on assessment in the UK, Australia, New Zealand, Singapore, Thailand and Malaysia. He has published extensively in the area of curriculum, assessment and evaluation.

Dr. Rosalind Y. Mau is currently working in the Division of Policy and Management Studies, National Institute of Education, Nanyang Technological University, Singapore. She has higher degrees in the sociology of education from the University of Hawaii, USA. She has taught assessment at the pre-service and in-service levels as well as over Hawaii Interactional Television, and has conducted courses in Micronesia and American Samoa.