

## ASSESSMENT PRACTICES IN SECONDARY COLLEGES OF THE AUSTRALIAN CAPITAL TERRITORY

BRUCE McBRYDE

Paper read at the annual conference of the  
Australian Association for Research in Education  
Gold Coast, Queensland  
November 1991

### INTRODUCTION

When Canberra was established as the capital of Australia, an area of land was excised from the state of New South Wales (NSW) to form the Australian Capital Territory (ACT). Because of the geographical location of the ACT, a decision was taken to use the NSW education system in the ACT. In particular, students in Years 11 and 12 attended high schools which used the syllabuses developed for NSW students and sat for the NSW Higher School Certificate (HSC). As the size of the capital grew, the level of dissatisfaction with the NSW system increased, leading to increased pressures for the development and operation of an independent system.

Following recommendations from the Campbell Report, a new system of secondary education was implemented in 1976. Secondary colleges were built for students in Years 11-12, leaving high schools with the responsibility for students still in compulsory schooling, i.e. from Year 7 up to the completion of Year 10.

The step was a major break from the past. The new secondary colleges had new grounds and buildings. Staff were selected on the basis of criteria which reflected the philosophies of colleges. The colleges had responsibility for syllabus development and assessment within the frameworks established by the ACT Schools Accreditation Agency, the body established to monitor syllabus development, assessment practices, standards and related matters. Although the NSW syllabuses were used as the starting point for the new

curriculum development, the offerings of the secondary colleges diverged rapidly from these patterns and from each other. Although teachers in the new colleges had no previous experience in college-based curriculum development and assessment, they soon accepted the challenge producing a wide range of new and modified courses.

The procedures described above were developed for students attending schools provided by the government. However, non-government schools were invited into the system which involved accreditation of their procedures by the Accreditation Agency. All but one of the non-government secondary schools joined the new system. Canberra Grammar School (boys) declined the invitation to be included in the new system and has continued its participation with the NSW syllabuses and HSC examination.

The model of syllabus development, assessment and accreditation procedures was broadly based on the model implemented in Queensland a few years earlier (Department of Education, 1970). However the procedures implemented in the ACT represented a larger step away from central control than the changes in Queensland. In the ACT there were no systemwide syllabuses. Also, the moderation procedures implemented to achieve comparability of assessment across schools was much less demanding for the ACT system of consensus moderation involving the review of student work as there had been in Queensland. Both systems did use the Australian Scholastic Aptitude Test (ASAT) as a scaling device to provide a statistical form of moderation. Student achievement in appropriate courses (T-courses) scaled to ASAT provide the raw data for the calculation of the Tertiary Entrance Scores (TEs) which are the major measures used to determine entry into universities in both Queensland and the ACT. These procedures were developed in Queensland to supplement the data provided by internal assessment. The rationale for the procedures and the data used to validate their use are reported by McGaw, Warry and McBryde (1975). The procedures were validated using data for the 1976 cohort of

students in the ACT by Keeves, McBryde and Bennett (1977).

Over the years, the procedures for selecting students into universities have come under pressure. The pressure has resulted from a mismatch between the supply of and demand for university places. The magnitude of the mismatch has increased dramatically over recent years. This has resulted in a number of research studies and enquiries into the calculation and use of TE scores. However, little systematic information had been collected on assessment practices in colleges or the procedures used to calculate course scores, which along with the ASAT results are the essential ingredients of the TE score calculation.

#### BACKGROUND TO THE STUDY

The major purpose of the study was to provide systematic research information on assessment practices in specific subjects in individual colleges. The study sought to document and describe how colleges derive and award unit grades and course scores.

The major aims for the study were to:

- . map the assessment procedures used at the college and faculty level to award unit grade and course scores for English, Mathematics and Art in secondary colleges in the ACT;
- . comment on the appropriateness and fairness of outcomes at critical stages;
- . evaluate the assessment practices; and
- . comment on important observations made by teachers and students about assessment practices.

#### METHODOLOGY

The use of qualitative evaluation methods has been a feature of advances in the methodology of educational evaluation in the past twenty years. While quantitative evaluators tend to turn their observations into numbers, qualitative

evaluators record their data in the language of their subjects in order to convey the meaning systems of the participants which eventually become the results of findings of the research. The primary concern is with description and interpretation rather than measurements and prediction.

It is the belief of the author that there is no one best solution. There is no single factor or set of factors that can solve any complex social situation; no one answer to the fundamental philosophical question: why do people do what they do? In the same way as there is no single right answer to complex questions, there is no way to get people to achieve what is wanted in the same way. If an accrediting body sets guidelines with broad objectives, then people will attempt to achieve the objectives in a variety of ways. This diversity does not imply that one solution will be right and others wrong. Some solutions may be more than adequate in achieving the relevant criteria.

Thus, research must continue, not to search for the best program, but for situational solutions. What is the best program at this time in these circumstances? How can we accomplish specific goals? How can we identify, implement and fine-tune alternatives rather than prescribing universal prescriptions which treat people and situations as if they are the same? The search for solutions to the problem of enhancing program effectiveness does not rest on any single methodological approach, because no single uniform approach will do the job in all situations. This paradigm is based on diversity and adaptability, not uniformity and rigidity.

#### Data Gathering Procedures

Information on the procedures, policies and practices of each college was collected and analysed prior to visits to individual colleges. This information helped in identifying the specific issues which needed further elaboration. The data gathering strategies used sought to explore the diversity in approaches used to achieve the overall objectives for the system. This required that the data be collected and interpreted in the

context of specific colleges. Hence, the study did not require the formulation or testing of any hypotheses. Rather, it required a flexible but in-depth investigation into the practices, rationale and context in each college. These were interpreted through the degree of fit with the principles underlying the ACT system, and a number of general education issues.

In order to obtain systematic information about the assessment procedures, a series of interview schedules were developed. These interview schedules covered a range of issues, including:

- . the nature and range of courses offered;
- . methods used to guide students and parents on subject choice;
- . college policy on assessment;
- . the structure of assessment programs in English, Mathematics and Art;
- . preparation and administration of assessment instruments;
- . steps taken to achieve comparability across classes;
- . the calculation of unit and course scores;
- . the formation of moderation groups; and
- . general concerns and suggestions for change.

Information was collected from each college through interviews with staff and students.

All levels of staff, principals and other administrators, subject coordinators and teachers of English, Mathematics and Art were interviewed. A sample of students completed a short student questionnaire and were invited to express their views on a number of issues.

## MAJOR FINDINGS

### Assessment Procedures Used in English

English teachers in colleges reported using a diverse range of modes of assessment, ranging from formal tests under examination conditions through to informal

observations  
of contributions made by students during class time. The most frequent and heavily weighted modes of assessments were essays and formal examinations. However, four colleges did not use a formal end-of-semester examination.

Only two colleges used common assessment across classes for all assessments. A further four colleges had some common assessments, usually end-of-semester tests, while the remainder of the colleges had few or no common assessments.

In all but one college, the teachers of each unit played a major role in the setting and marking of assessments. This was generally complimented with discussion among other teachers of the unit both during test development and also at the time of marking. In arriving at course scores, this process of consultation was seen as an important part of the moderation procedures across units within colleges. Half the colleges used the results from this process of calculating unit scores directly without further statistical adjustment in the calculation of course scores. The other colleges either standardised unit scores or scaled them to historical ASAT parameters prior to aggregation.

When students studied more than the minimum number of units to qualify for a major, most colleges allowed students to delete the lowest scoring units when calculating course scores. Three colleges allowed students to delete a maximum of one unit while a further two colleges averaged scores in all units in the calculation of course scores.

#### Assessment Procedures Used in Mathematics

The nature and range of assessment practices used by Mathematics teachers contrasted with those used in English. English teachers had concentrated on using a variety of modes of assessments, including tests, theory and practical assignments, class participation, oral performance, with much of the responsibility for setting and marking delegated to class teachers. Mathematics teachers used procedures which focused on

formal tests under examination conditions operating with centralised control within colleges.

In all colleges, written tests accounted for at least 80 per cent of the weighting for assessment in Tertiary Mathematics courses. Teachers generally wished to keep the weighting given to other forms of assessment as small as possible because of the difficulties encountered in achieving comparability across classes and because of the ease with which students could copy from each other. Teachers were more willing to use alternative forms of assessment in Accredited courses in Mathematics.

The belief by some Mathematics teachers that testing needed to be common to allow valid comparisons, led some teachers to suggest that tests should be common across colleges. Some also suggested that courses should also be common across colleges, while others, on further probing, gave inconsistent information. On one hand, they proposed common testing across colleges but, on the other hand, supported the flexibility which was available to colleges in course development and organisation.

Procedures used for marking assessments and for the calculation of unit grades and course scores also reflected the belief held by most Mathematics teachers that statistical procedures were required to achieve equivalence across groups. There appeared to be little support among Mathematics teachers for procedures involving collaboration and discussion to achieve equivalence across classes or colleges.

Mathematics is one subject area for which, at the time of the study, colleges could exercise the option of combining their courses into either one or two moderation groups. Many colleges have experimented with both options. Since this study was undertaken, a decision has been taken to standardise the procedures between colleges by eliminating the option of using two moderation groups. The advantage of using two groups was that the college did not have to make the judgements about the relativities between students studying the different levels. These relativities were determined by the

### ASAT scaling

procedures. This avoided conflicts among teachers and between teachers and students.

On the other hand, teachers who opted for one moderation group maintained that, although the task was difficult, it was a college responsibility which should be tackled by those who best understand the achievements of the students and the contexts within the college.

### Assessment Procedures Used in Art

In contrast to Mathematics, teachers of Art relied almost entirely on collaborative consultative procedures to achieve comparability among students. These procedures were easier to implement in Art because of the smaller number of students studying the subject.

It also reflected a fundamental difference between teachers of Mathematics and Art in their perception of teaching and assessment. Teachers of Art saw assessment as an integral component of the teacher/learning environment. The final products for students in Art, which for many colleges was the sole basis of the assessment, were generally developed through the semester with teachers providing advice and support along the way.

The notion of common tests, or even formal tests, was inconsistent with this pedagogy. In fact, the difficulty faced by many teachers was to separate the ideas and effort which students contributed to their work from the support and ideas provided by teachers and other students.

In most colleges, all aspects of the assessment process were collaborative. This included the development of the assessment program, setting of assessment tasks and marking. In some colleges teachers involved students in both peer and self-assessment. Six colleges reported using input from students extensively throughout the assessment process, while a further two colleges sought evaluative comments from students concerning the quality of their paintings. One college provided students with a self-evaluation

sheet. Students at several colleges had input into decisions concerning topics, weighting and timing of elements of the assessment program.

#### Views of Principals on Assessment Policies and Issues

Principals generally used assessment committees within their colleges to exert their influence on practices within colleges. The major purpose of these committees appeared to be to ensure that there was reasonable consistency across courses within the college. However, the committees provided flexibility for teachers within specific subject areas to implement alternative strategies to meet specific needs of their subject if they could justify the need for such alternatives.

Most of the decisions made by assessment committees were directed towards creating an equitable balance across subject areas, e.g. the number of assessment elements, length of tests, separation of assignment deadlines from test periods, the weighting of assessment elements in forming unit grades and the weighting of units in the calculation of course scores. Other decisions to which principals made substantial contributions were mechanisms for appeals, dealing with inappropriate student behaviour such as plagiarism and procedures to be used to compile course scores. Most principals also expressed a keen interest in examining the shape of distributions of course scores across subject areas before they were submitted to the ACT Accrediting Agency for further processing, leading to the calculation of Tertiary Entrance Scores.

Most principals were generally satisfied with the procedures operating in their colleges. One area of concern was the formation of moderation groups within colleges. It appears that most colleges have experimented with combining courses with small enrolments, together with courses with a large enrolment to ease the problems of moderating the smaller group. However, subsequent experience showed that this often created more problems than it solved, so principals have moved away from forming larger

moderation

groups, except under special circumstances. These problems were that the courses were conceptually similar, that there was an overlap of teachers of the courses, that some students studied both courses and that the courses were scaled to the same sub-scale of ASAT. If these conditions were not met, then principals were most likely to recommend that the courses be treated separately in their college, even if this entailed the use of small group moderation procedures for one or both of the courses.

The majority of principals expressed confidence that small group moderation procedures had worked reasonably well in recent years and that, in most cases, their use was preferable to forcing combinations of courses which were not really compatible. On the other hand, a minority of principals reported that some small group moderators had made judgements about work from their colleges which they thought were inappropriate and unfair. However, this flexibility in procedures available to college principals is seen as a positive aspect of the system, as it allows college principals to select the methods which are pertinent to their situation.

While principals were generally satisfied with college procedures over which they had control, there were mixed reactions to those procedures over which they had little or no control. The use of ASAT to scale course assessment received qualified support from a number of principals. Some expressed the view that it was better than anything else which was available, and certainly better than external examinations. However, a number of reservations were expressed, including bias in favour of boys, bias in favour of Mathematics and Science courses, the use of ASAT sub-scales and the fact that ASAT measured aptitude rather than achievement.

Principals also expressed concerns about the effectiveness of Moderation Day. Many believed that part of the day should be devoted to issues relating to the comparability of standards in courses across colleges. There was also strong support for

the traditional function of providing opportunities for the professional development of teachers. Pursuit of the priorities relating to comparability of standards would involve a greater emphasis on the comparison of student scripts and further development of teacher skills in the areas of assessment and moderation.

#### Views of Teachers on Assessment Policies and Issues

Although the major focus of the study was to provide a 'detailed map' of assessment practices in colleges, the collection of these data inevitably led to the respondents making evaluative comments on the topics which were discussed. These evaluative comments are summarised in this section.

In general, teachers were confident that the activities for which they had responsibility achieved the desired goals. This included procedures to achieve comparability of grades and scores across units and classes and comparability between students studying at different depths (i.e. between minors, majors, major-minors and double majors). They were less confident that comparability was achieved across colleges, an area over which they were able to exercise less control, although they appeared to have more confidence in the effectiveness of ASAT than in any activities undertaken during Moderation Day to achieve such comparability. In spite of these reservations about across-college moderation procedures, few teachers saw external examinations as a viable option to achieve this goal.

Most of the teachers who supported external examinations were Mathematics teachers. It would appear that of the three subjects analysed in this project, that Mathematics has characteristics which distinguish it from the other two and that these characteristics have implications for the types of assessment procedures which are consistent with the types of objectives and teacher practices associated with these subjects. Informal comments from teachers would suggest that the physical sciences may be similar to

Mathematics, in that the major objectives are able to be assessed using a narrow range of assessment techniques. For the other subjects in the curriculum, it would appear that teachers would argue for a wide range of assessment techniques.

Related to this variety of assessment procedures is teacher perception of teaching, learning and assessment and the range of course objectives achieved. Teachers of English and Art claimed to be placing a greater emphasis in these two areas than did Mathematics teachers.

#### Student Views on Assessment Issues

The views of students were tapped from two sources, a student interview and a short questionnaire. The questionnaire sought their views on a wide range of issues concerning:

- . whether tests and assignments were returned after they were marked and, if so, how they were used to assist student learning;
- . student perception of the adequacy of the number of assessment elements used, including the use of tests which did not contribute to course scores;
- . whether assessments were marked by class teachers or other teachers of the subject;
- . student satisfaction with marking procedures; and
- . the comparability of scores between classes and colleges.

The data show that the majority of students were reasonably satisfied with the assessment practices in their colleges. Most students reported that their tests were marked and returned in a reasonable time, although the time period reported by students was somewhat longer than that reported by teachers. Most students were satisfied with the number of assessment elements. Some students expressed a desire for more assessment. This result, which was supported by the interview data, conflicts with the commonly held view that, in systems of internal assessment, students are

extensively over-assessed. Students were also reasonably satisfied with marking procedures and the achievement of comparability of scores across classes and even colleges.

In the student interview, four major topics were pursued:

- . the frequency of assessment;
- . student contribution toward their own assessment;
- . equity issues; and
- . attitude towards assessments which do not contribute towards course scores.

The issue of frequency of assessment has been dealt with above. The level of student contribution towards their own assessment was not great, although teachers of English and Art in some colleges were making a considerable effort to involve their students. It was interesting to note that students often had reservations about peer and self-assessment. Some felt that they were not qualified to make a contribution and, hence, their judgements would be invalid. Others expressed the view that teachers were opting out of their responsibility, rather than seeing it as an opportunity to make a contribution (no matter how small) or for their own personal growth. Equity issues were not major issues for students. A few students were concerned about differences in treatment between boys and girls, while others were concerned about the perceived differences in standards and procedures across colleges. The attitude towards assessment which did not contribute directly towards scores was mixed. One problem was to identify which, if any, assessment met this criterion. Most students felt that all assessment (including those traditionally seen as formative) contributed in some way to final scores. Thus, there was strong support for the notion that all assessment had to be taken seriously.

#### COMPARISON WITH STUDIES IN QUEENSLAND

It was not intended that the results of the present study be compared

directly with any previous study. However, the aims of the study were formulated from a study undertaken in Queensland in 1978 (McBryde and Lamont, 1980). The author was also involved in the 1974 Schools Under Radford Study (Fairbairn, McBryde and Rigby, 1976) which was undertaken in the early years of internal assessment in Queensland.

In general, the findings of the Canberra study were similar to those of the Queensland studies on those issues which overlapped across the studies. In fact, there was greater commonality across studies within particular subject areas in attitudes and assessment practices than between subjects. This is obviously partly due to the selection of subject areas to be investigated. The major criterion in the selection of subject areas in the Canberra study was that the study should include subjects which used different approaches towards assessment practices.

In all studies, mathematics teachers tended to use a narrow range of assessment instruments, focussed on end-of-semester written tests under exam conditions. While these types of tests were also used in English, and sometimes in Art, there was much less emphasis on this form of assessment. Teachers of Mathematics required rigid marking schemes and specific answers to questions in order to feel comfortable about comparability across classes. Teachers of English and Art were more likely to accept judgements made after reviewing samples of work from different classes and after discussions among the teachers concerned.

It was interesting to note that in all studies, respondents expressed concern about the level of comparability of assessment which was achieved across schools/colleges. The same conclusion was reached by Viviani (1990) in her study into Tertiary Entrance procedures. In her report she identified lack of confidence in the comparability between school assessments as a major problem, recommending two strategies for its resolution. She recommended more research to investigate the extent of comparability at

present and the establishment of a committee to review current practices and to recommend reforms. This committee (CRAMP) has been established and should table its findings later this year.

One of the major differences between the Queensland and Canberra data was probably due to the different levels of devolution of responsibility in course development. In Queensland, Board subjects (the equivalent of T-courses in Canberra) are developed and accredited centrally by the Board of Senior Secondary School Studies. Schools use the syllabuses for these subjects to develop work programs which are then accredited by the Board. In Canberra, there are no common syllabuses. Each college produces its own syllabus and work program for each course offered. This has led to a greater diversity of courses offered in Canberra. The data from the studies showed that there was a greater sense of ownership of syllabuses and work programs by teachers in Canberra than in Queensland.

#### CONCLUDING COMMENTS

It is my opinion that the system of curriculum development and assessment at the college level which was implemented in the ACT in 1977 is providing responsive, accountable, fair educational programs. Future strategies for improvement should be directed towards a fine-tuning of the present system rather than making a major change in policy or practice.

This view was supported to a large extent by those interviewed with recent experience in other systems, most of which had central curricula and a system of external examinations to achieve comparability across schools. These respondents reported significant educational benefits in the ACT secondary college system.

It is my view that teachers and senior staff in colleges operate in a professional way, taking responsibility for course development and the implementation of teaching strategies which cater for a diverse range of student needs.

The ACT system, which places responsibility for curriculum development and assessment at the college level, is based on the belief that decisions should be made by people who have the relevant information and understand the context in which the decision is made. This has led to decisions being made at a range of levels. Classroom teachers make decisions about the relevance of assessment programs and marking procedures in the context of their own teaching strategies and the characteristics of students in their classes.

Subject co-ordinators make decisions which promote consistency in assessment practices across classes. Principals, often with support from an assessment committee comprising senior staff, make decisions to promote consistency in assessment practices across courses within each college. The overall observation was that major decisions involved relevant staff working within guidelines set by those at a senior level, while being open to challenge from students, parents or other teachers who were affected by the decision. It is my view that these checks and balances make a significant contribution to the effectiveness of the system. These conclusions are based on the belief that there is no 'best solution' to the problems facing colleges in the development and refinement of their assessment procedures. Teachers and college administrators must continually review their procedures to ensure that their solutions are consistent with their objectives and priorities.

This is not to suggest that the Canberra system is superior to the Queensland system. The aim of this paper was to focus on the findings of the Canberra study. The comparisons with Queensland were based on studies undertaken over ten years ago. None of the studies were designed to provide comparative data, so the comparisons made must be interpreted with caution. In addition, a major reform resulting from the Review of School-based Assessment in Queensland (ROSBA) has been undertaken since those Queensland studies were undertaken.

It is anticipated that a future paper will deal with the events which led to the Viviani study in 1990 and the consequent reforms in Queensland.

This paper was based on the Research Report Student Assessment Project in Secondary Colleges of the Australian Capital Territory which was prepared for the Committee to Supervise Research into the calculation of ACT Tertiary Entrance Scores, chaired by Dr Don Anderson, Australian National University.

#### References:

Board of Senior Secondary School Studies, A Review of School-based Assessment in Queensland Secondary Schools (ROSBA), E. Scott, Chairman, Brisbane, 1978.

Department of Education, Public Examinations for Queensland Secondary Students, W.C. Radford, Chairman, Brisbane, 1970.

Fairbairn, K., McBryde, B. and Rigby, R., Schools Under Radford (SUR), Department of Education and Board of Senior Secondary School Studies, Brisbane, 1976.

Keeves, J., McBryde, B. and Bennett, L., The Validity of School Assessments at the HSC Level for Colleges and High Schools of the Australian Capital Territory, AARE Conference papers, Canberra, 1977.

McBryde, B. and Lamont, A., Mapping of Assessment Procedures in Schools, Department of Education, Queensland, 1980.

McBryde, B. and van der Hiede, G., Student Assessment Project in Secondary Colleges in the Australian Capital Territory, Department of Education, Canberra, 1989.

McGaw, B., Warry, R. and McBryde, B., Validation of Aptitude Measures for the Rescaling of School Assessments, Report No. 2, The Queensland Grade 12 Study, Research Branch, Department of Education, Brisbane, 1975.

Viviani, N., The Review of Tertiary Entrance in Queensland, Department of  
Education,  
Brisbane, 1990.