Australian Council for Educational Research (ACER)

From the SelectedWorks of Dr Lawrence Ingvarson

1984

Consensus moderation observed

Lawrence Ingvarson

Available at: https://works.bepress.com/lawrence_ingvarson1/109/

CONSENSUS MODERATION OBSERVED

Lawrence Ingvarson
Monash University

Kevin Hall
Melbourne College of Advanced Education

Gary Pollard
Education Department of Victoria

ABSTRACT

This paper reports the findings of a three-year evaluation study of consensus moderation procedures as applied to the school-assessed component in H.S.C. Geography in Victoria.

Since 1981, Group 1 subjects in the assessment program of the Victorian Institute of Secondary Education (V.I.S.E.) at Year 12 level have included a school-assessed component in addition to an external examination. For most subjects, the method of moderating the school-assessed component is by statistical moderation. Moderation by visitation and consensus moderation are the two other methods currently used in Victoria.

As V.I.S.E. is proposing to extend the number of subjects for which consensus moderation procedures may be used, the implications of the experience are discussed in some detail. The findings of the evaluation will be of interest to public examining authorities in other Australian States.

In the context of public examinations, 'moderation' refers to methods used to bring about comparability between teachers in the assessments they make of their own students' work within a particular school subject. The need for moderation arises when, in addition to an assessment based upon the conventional external examination, some proportion of the final examination mark is based upon teachers' marks of work they have set for their own students in school during the year.

Consensus Moderation was introduced in 1981 for a small number of Group 1 Higher School Certificate (H.S.C.) subjects by the Victorian Institute of Secondary Education (VISE), the statutory authority responsible for the administration of public examinations in Victoria. An evaluation of consensus moderation procedures, as used in one of these H.S.C. subjects, Geography, was conducted from 1981 to 1984 the first three years of this complex innovation. It was commissioned by the VISE Geography Subject committee. This paper summarizes some of the main findings from the evaluation.

The evaluation aimed to provide the committee and the teachers involved in implementing the new procedures with an independent source of information and judgements relevant to their own evaluations and the decisions they had to make from time to time to improve the operation. In this sense the evaluation team adopted a 'responsive' approach, after Stake (1978)

Since its introduction, consensus moderation has been the subject of considerable controversy. Doubts concerning its fairness and reliability have been raised together with many other issues such as its effects on teacher-student relationships and the workload of both teachers and students. Practical problems also arose for those involved in the course of learning how to implement the new procedures as there were virtually no precedents to follow. It was a case of learning how the principle of consensus moderation might be made workable in practice, and a matter of refining the concept of consensus moderation based upon learnings from practice. The role of evaluation was seen to be one of facilitating this learning by the main actors involved. As issues arose the evaluation team gathered information relevant to that concern which was fed into the deliberation process when it was needed in the form of informal brief reports and tabulations.

One of the main findings of the evaluation was the importance of conducting such studies over an extended period of time. At the end of the first year of implementation the balance of evidence in favour of the new procedures was precarious. The difficulty of putting them into practice had been underestimated. But after three years the situation had changed significantly, to the extent that it was possible to say that

consensus moderation has been a successful innovation with considerable long-term side-benefits for the professional development of the teachers involved. Another significant feature of consensus moderation is that it provides a practical means of moving from a system based on external examinations to, perhaps eventually, a system of credentialling based entirely on school-based assessment procedure. It does this by providing widespread upgrading of the skills teachers possess in curriculum and assessment and, consequently, public confidence that school-based assessments can be conducted in a manner consistent with important criteria such as fairness, reliability and educational quality. Without such a means for enhancing skills and confidence it seems foolhardy to advocate, as some are doing, rapid transition to total school-based assessments for credentials such as a Certificate of General (or Secondary) Education, assuming that our society will continue to expect secondary education to provide some type of credential for most of its graduates.

Types of Moderation used in Victoria

Since 1981, all fifty-four VISE Higher School Certificate (i.e. Year 12) Group 1 subjects include at least one school-assessed component which is 'moderated' by one of three methods (to be described below). For most subjects, the component which is assessed by teachers is an optional unit worth 30 per cent of the final mark. In Geography, however, students are also required to undertake an 'Individual Research Project' component which is assessed by their theachers and worth an additional 20 per cent toward their final mark. This means that for Geography the external examination contributes only 50 per cent to the final assessment, compared with 70 per cent for most Group 1 HSC subjects. Moderation should be distinguished from standardization, which refers to statistical methods used by public examination authorities in an attempt to create some comparability between marks awarded to groups of students across different subjects. One of the reasons for standardization in Victoria is to enable a student's marks in a number of different subjects to be added together to form an aggregate score, such as the 'Anderson Score' in Victoria. The Anderson Score is a common criterion used for selection purposes by tertiary institutions.

According to the Victorian Institute of Secondary Education, moderation methods are

> designed to place different schools' assessments of student performance, in the same subject, on to a common scale so that the moderated school assessments may be validly combined with the external mark in that subject. (Cornish, 1983)

In Victoria three methods of moderation for the school-assessed component are used: statistical moderation, moderation by visitation, and consensus moderation. In 1983 consensus moderation was the preferred method in Ancient Greek, Art, Computer Science, Geography, Graphic Communication, Latin, Music A, Music B, and Spanish.

In Statistical Moderation the method is mathematical and involves the adjustment, on to a common scale of assessment, of scores awarded externally and internally, to each particular school group. This adjustment is statistically computed, using the external scores as the basis for adjustment. The assumption underlying this process of adjustment is that the work which teachers set for school-based assessment purposes tests the attainment of skills and understandings similar to those tested by the external examination.

In Moderation by visitation the process of ensuring comparability requires a person, or panel of persons, to visit schools. As well, teachers meet together early in the year to plan and discuss aspects of the course relevant to assessment procedures. At the end of the year, two moderators visit the schools and, in consultation with the teachers, review samples of student work and adjust assessments, if necessary. In 1983 the only subjects using this method were Physical Education and Physical Science.

In Consensus Moderation groups of teachers meet during the year to establish mutually accepted conditions and criteria to be applied to the moderation of student work within guidelines set by the relevant VISE Subject Committee and approved by the VISE Assessment Procedures Committee. At an end-of-year meeting these groups of teachers review ('moderate') samples of student work randomly selected by the convenor from each school in their group, adjusting marks where necessary in order to ensure that teacher assessments (or 'standards') are comparable from school to school. Each group is led by a convenor, usually a teacher, appointed by VISE, from outside its membership, who facilitates the process of consensus decision-making, ensures that the group works within guidelines established by the Geography Subject Committee, and provides a means of cross-group comparison as they usually convene two groups. Convenors alone meet as a group at the end of the year to make cross-group comparisons and adjustments, where necessary, based upon samples of student work from each consensus moderation group. After this final moderation meeting the moderated marks are added, without

further change, to the marks students have obtained in the external exam.

Consensus Moderation in Geography

In the case of Geography there are about thirty consensus moderation groups (C.M.Gs) based on regions across the State. Each group consists of teachers from eight to ten schools. All group members must attend a full day (in school time) meeting early in the year, an after-school meeting in the middle of the year and a final full day moderation meeting in November. If members cannot attend a meeting they must abide by decisions made in their absence. It is common for groups to arrange additional informal meetings to share ideas about assessment tasks such as field work and practical work and to assist each other with problems such as determining whether topics students have selected for their individual research projects (IRPs) meet the guidelines.

At the first meeting of the year each group must clarify the assessment criteria for the individual research project; each teacher must nominate the two-three assessment methods he/she will use for the Optional Unit (there are six methods to choose from including item bank, fieldwork report, practical work report, extended essay, individual research report and action research report); and each group must clarify and determine, within VISE quidelines, how they will implement the criteria for these assessment methods for the Optional Unit. The mid-year meeting is mainly to discuss whether any adjustments to the implementation of the assessment procedures are necessary and to provide teachers with advice and support.

Convenors meet regularly during the year with a VISE-appointed Chief Convenor nominated by the Geography Subject Committee. The Chief Convenor's role is to oversee the team of convenors, liaise with Chairman of the Panel of Examiners in the HSC Geography and review and report on the moderation procedures in each year to the VISE Assessment Procedures Committee. Both the Chief Convenor and the Convenors receive honoraria and allowances to cover administrative costs and travel expenses.

A most important point about consensus moderation, as it is practised in HSC Geography in Victoria, for any group interested in adopting it to appreciate is that it grew out of a broad-based reform of the secondary geography curriculum involving many teachers over an extended period of time. It was not an imposed change. Interlocking networks of teachers from the Geography Teachers Association of Victoria, the Curriculum Branch of the Education Department and the VISE Geography Subject Committee pressed for changes in the mode of HSC assessment which would better reflect the inquiry-based concept of Geography which they had progressively developed and implemented over the previous ten years with the assistance of sequential in-service activities involving many teachers. The rationale and the practical plans for consensus moderation were developed by a core group of geography teachers, with the assistance of VISE officers, because they felt that a sit-down external examination could not adequately assess attainment of many of the skills and abilities which the new course was trying to promote. As a consequence, when it came to implementing consensus moderation, there was a considerable number of teachers with the background understanding required to act as convenors and a majority of HSC teachers prepared to take a step into the unknown. It would be foolhardy for any public examination authority to adopt consensus moderation unless these pre-conditions were satisfied. Implementation of this complex innovation requires both skill and commitment. And it may be that for some subjects, such as mathematics, moderation by consensus is inappropriate or unnecessary.

PRINCIPAL TASKS UNDERTAKEN BY THE EVALUATION TEAM

The evaluation team placed its main emphasis on the following tasks.

(i) Identifying and clarifying the key questions and issues which the main groups involved felt the evaluation should address if it was to be useful for them. The groups included were: HSC students, teachers, convenors, members of the VISE Geography Subject Committee, VISE officers, and principals.

(ii) Constant probing of these groups for criteria by means of which the success or failure of consensus moderation might be judged.

(iii) Gathering data by means of: observation of the work of consensus moderation groups and convenors' meetings; interviews with HSC classes, teachers and convenors; surveys of students, teachers, and principals; analysis of statistical data provided by VISE arising from HSC assessments; documentation such as convenors' reports, minutes of convenors' meetings and VISE internal reports.

(iv) Regular feeding back to convenors' meetings of this data immediately it became available, using it to draw the attention of convenors to emerging concerns, such as, in the early years, the workload

pressure on students, and the varying interpretations of the guidelines being implemented across C.M.G.s.

(v) Provoking discussion at convenors' meetings to clarify the 'taken-for-granteds'. For example, it became clear at one point in the evaluation that it was assumed by convenors that they held a common understanding about what 'moderation' meant, when in fact we had seen different interpretations of moderation being practised across C.M.G.s.

MAIN FINDINGS OF THE EVALUATION

There is only space in this paper to report a selection of the findings. The length of this paper precludes any adequate treatment of the rich data from observations of consensus moderation groups or interviews.

Teachers' views on consensus moderation

Representative samples of HSC Geography teachers were surveyed at the end of each of the first three years of implementation.

The proportion of teachers who had confidence and trust in the ability of their C.M.G. to assess student work fairly and reliably had increased steadily over the first three years of consensus moderation. Assessment of work for the Option had not had the same level of confidence and trust as for the IRP, but even so the level had increased from 43 per cent in 1981, and 61 per cent in 1982, to 76 per cent in 1983. Only 6 per cent of teachers lacked confidence in their groups' ability to assess IRPs fairly and reliably in 1983.

There was a marked change in the level and nature of teacher concern over time which indicated the importance of regarding consensus moderation as a multi-faceted innovation requiring a considerable amount of un-learning and re-learning of skills and attitudes to do with the assessment of learning. It was appropriate to use an evaluation with a long time span focused on facilitating this kind of learning. The complexity of the change involved should not be underestimated and it represents a professional development exercise of considerable proportions. There is no such 'thing' as consensus moderation to be taken from a shelf, as it were, to be used. In Geography it required a long-term process of learning how assessment methods can more appropriately match the kind of learning which teachers aspire to promote when they regard their subject more as a field of enquiry than a body of facts.

Most teachers thought that involvement in consensus moderation had contributed significantly to the skills they possessed for assessing student learning, their ability to evaluate and improve their teaching and the quality of learning and independence in learning of their students. Most also felt that their access to useful ideas for teaching had markedly increased through membership of a moderation group.

Students' views on consensus moderation and school-based assessment

The initial high level of concern about workload amongst students in the first year of consensus moderation fell away significantly in the second year. In 1982, a substantial proportion (70%) of students thought that the workload in Geography was fair. The workload in Geography was perceived by students to be much the same as it was for other HSC subjects that they were taking.

In terms of workload, the school-based assessment and consensus moderation requirements have had effects which are more dependent on the ways in which particular teachers interpreted and implemented them than the procedures per se. There were many schools where all students felt that the workload was fair. On the other hand, there were some schools where every student thought that the workload was unreasonably high.

Students' written responses indicated that there was a need for teachers to share ideas with each other (and this applied particularly to those taking HSC Geography for the first time) for distributing the workload more evenly over the year. Their responses also indicated that attention should be given to ensuring that the amount of work demanded from students was more consistent from school to school. This applied within and across consensus moderation grups. (This data was fed back to convenors to provide a basis for discussion within consensus moderation grups about the ways in which the amount of work could be reduced as well as made more equal from school to school. Students' responses made it obvious that there was a need for teachers to share ideas about ways of setting tasks for practical work, field work, and research projects that were more economical in terms of time and effort, whilst still meeting the assessment requirements of the course.)

Only ten per cent of students would have preferred a system where their mark for Geography was totally based on an end-of-year examination. Eighty per cent felt that the way in which they had been assessed was fairer than it would have been with an end-of-year exam only.

Contrary to some predictions, consensus moderation had not had detrimental effects on teacher-student relationships. Only seven per cent of students felt that the fact that their teacher had assessed their Option and IRP work had produced any problems for them. The vivid comments of those students who felt that it had produced problems made it clear that it was more to do with the competence of their teacher than school-based assessment per se.

Principals' Views on Consensus Moderation

Most principals surveyed, after three years of consensus moderation, thought that the involvement of teachers from their school in the new system had made a positive contribution to their professional development and enhanced the quality of their teaching methods. It was a source of concern, however, that a third of principals perceived their staff to be overworked as a result of their involvement in consensus moderation. (Workload does appear to be dependent, in part, on the particular CMG that a teacher is in.) Seventy per cent of principals thought that consensus moderation was fairer to students than the previous system with an end-of-year examination only. However, there was a significant difference between principals from different systems on this point. Only 55 per cent of principals from Independent schools thought that it was fairer compared with 82 per cent and 71 per cent of principals from Catholic and State schools respectively. Also, principals from Catholic and State schools were more likely to want the present system retained or extended than those from Independent schools. Only nine per cent of principals overall wanted a return to an end-of-year examination without school-based assessment and consensus moderation.

Other Findings

When the distribution of marks for the external examination, the Option and the Individual Research Project were compared it became apparent that the school-based assessment components had produced a wider 'spread' of scores. That is, the standard deviation of the distribution of IRP and Option marks was greater than that for the external examination. In the traditional measurement sense, this meant that the school-based assessment component acted as a better 'discriminator' between students than the external component. On a scale from 0-200 the mean mark for the external component was 124.1 with a standard deviation of 22.9. The respective means and standard deviations for the IRP were 125.2 and 36.6 and for the option 130.8 and 26.3

The relative rankings of C.M.G.s in terms of external assessments and marks awarded for the school-based components revealed no patterns which might have suggested a tendency for standards to vary consistently between country and metropolitan groups, or between convenors. There was a tendency for high school students to do better on the school-based assessment, relative to the external examination, than students from Independent schools. There was a very high correlation between the rankings of C.M.G.s for each assessment component from one year to the next.

An analysis of the costs of consensus moderation conducted by VISE has shown that the average cost to VISE itself (this excludes any replacement costs which the employer may have had to meet) was about $7.00 per candidate per subject. The corresponding cost for subjects not using consensus moderation was about $16.00 per candidate per subject. Thus consensus moderation is about half as expensive which is not surprising when it is remembered that teachers in C.M.G.s are doing moderation for no payment and the whole paraphernalia of invigilators, hiring halls and paying markers has been removed. (One interpretation of these respective costs is that VISE should increase its payments to convenors and the Chief Convenors to match more realistically the time and effort they must put into their responsibilities.) In the opinion of the Chief Examiner for HSC Geography, the quality of students' geographical understanding as reflected in their answers in the external examination, had improved noticeably since the introduction of school-based assessment components such as fieldwork and individual research projects. One university volunteered the view that first-year Geography students who had come through the new system were markedly more competent in handling field work and writing research reports than their predecessors.

CONTINUING OR EMERGING ISSUES

1. Replacement and training of convenors

Moderation by consensus, compared with statistical moderation, is crucially dependent on the social skills and personal qualities of the people who act as convenors for each moderation group. It must be possible for any HSC Subject Committee planning to use consensus moderation procedures to be able to call upon a group of teachers willing to spend a considerable amount of their own time planning and orchestrating the work of (usually) two C.M.G.s. As HSC teachers, many convenors are also members of a third C.M.G. Such an involvement would require a convenor-teacher to be absent from their own schools for at least six days each year, quite apart from numerous late afternoon and evening meetings. They need to be 'experts' with respect to the course description, and the regulations which govern the consensus moderation process (which are extensive). In addition, they need to be trained in verbal and problem solving skills to facilitate decision-making based on consensus and to be able to mould a number of teachers into a congenial task-oriented group. Interviews with convenors revealed that they had gained satisfaction from undertaking the role and learning to handle the new responsibility, but it was demanding. In addition to planning and conducting the scheduled meetings of a CMG throughout the year, participating in informal out-of hours meetings and attending meetings of convenors, they often fulfil a 'resource person' role, much like a consultant, to teachers in the group. Because of the wide-ranging nature of the convenor's role some find it difficult to know where to draw the line with regard to their availability and form of assistance. Commitment and a sense of responsibility for the course has drawn most people into the role of a convenor, but many of the original convenors now feel that they need to withdraw from the role to avoid 'burn-out', and also to prevent consensus moderation becoming dependent upon particular convenors. Convenors have suggested a possible list of attributes which could form the basis of a set of selection criteria. Induction programs for new convenors were seen as very important as was the need for a professional development program for current convenors. People can only be expected to act in the capacity of convenors for a limited period of time. Therefore there needs to be forward planning for the identification and preparation of new convenors. So far, it has not proved difficult to find new teachers willing to take on the frugally rewarding pastime of being a convenor as the old-timers retire.

2. The need to regard consensus moderation as a year-long process.

For many participants in the early years of the process, moderation was what happened at the November meeting, and all else was preliminary. There is an emerging view that moderation also includes ensuring comparability in the interpretation of objectives, the tasks set, the conditions under which students do the tasks, the assessment techniques used, and the teachers' views of 'standards'. In this view, moderation includes the context as well as the outcomes. It appears that, if the contextual aspects are monitored during the year, C.M.G.s have fewer difficulties in moderating the assessments at the final meeting.

With the experience of the past three years, we have come to understand that consensus moderation may be regarded as a process by means of which trust is built up amongst a group of teachers: trust based on a mutual understanding of the activities each is using to assess the degree of achievement of the objectives of the course, and why. This includes confidence that conditions, such as those listed above, are comparable for students from different schools. When members of a C.M.G. circulate amongst themselves their proposals for field work, practical work, objective tests, essays, and student research projects for review and approval during the year, validity has been promoted in that a check has been carried out on the extent to which these tasks and assessment items are relevant to the objectives of the course. Fairness has been promoted to the degree that the group has ensured comparability in the conditions under which students from different schools demonstrate what they can do. Teacher accountability to one another is high through the sharing and review of each other's plans and assessment items regularly during the year.

3. The autonomy of Consensus Moderation Groups.

There has been a small, but persistant, tension between the twin values of group autonomy and the need for specific criteria and guidelines to be laid down centrally for each of the assessment items. Within the criteria and guidelines, provided by VISE, C.M.G.s have quite a degree of discretion. It is the role of the convenor to facilitate decision-making and consensus in the group, within the guidelines, and to pay respect to the integrity of the groups, not to be examiners.

Yet there is another responsibility to ensure comparability across C.M.G.s. The 1983 survey of teachers showed that teachers were almost equally divided about whether more specific criteria should be laid down for moderation of the optional unit (i.e., field work, practical work, etc.) such as those which applied to the I.R.P. Some of the convenors felt a pressure to play down this role of facilitating adaptations to local needs and conditions to act instead as local representatives ensuring only compliance with central examining authority criteria. As one convenor put it, "Some of the groups will become very resentful if we do move into an examining role, because they are increasingly competent to make judgements and, also, to help one another. They're seeking advice impersonally, within the peer groups, which I think is the real outcome of the whole consensus moderation process". This tension is inevitable and it is to be hoped that it will remain a productive one.

Reference

Stake, R.E. "Responsive Evaluation". In D. Hamilton and others (Eds.) Beyond the Numbers Game. Berkeley, Calif.: McCutchan, 1978.