Elements of Effective Curriculum Evaluation: 
An analysis of a Queensland model of curriculum evaluation

Paper presented at ‘Framing the Future’, the September 1999 Biennial Conference of the Australian Curriculum Studies Association

E. J. Bell and R. J. Dudley

Abstract

This paper aims to explore the model of curriculum evaluation in use in the Queensland senior secondary school system in ways that lead to identification of some elements of effective curriculum evaluation. It does this by exploring the broad orientation of the Queensland senior system in the context of some key questions about the nature, development and conceptualisation of curriculum evaluation, and then by examining in some detail the model for curriculum evaluation developed by the Queensland Board of Senior Secondary School Studies. This operational detail emphasises some key implications for curriculum evaluation. The paper concludes with a discussion of some possible elements of effective curriculum evaluation.

Introduction

This paper is structured by the following key questions:

• What is curriculum evaluation?
• What is the history of curriculum evaluation?
• How can we conceptualise different models of curriculum evaluation?
• What model of curriculum evaluation has been used by the Queensland Board of Senior Secondary School Studies?
• What are some elements of effective curriculum evaluation?

The broad conclusion of this paper is that effective curriculum evaluation is research that can tell decision makers what is working, what is not working and how to make non-working things work. The discussion in this paper aims to explore what this means in practice in one system.

What is curriculum evaluation?

The term curriculum evaluation is hard to define, not least because the word curriculum is an elusive one in the research literature. As McCormick and James (1990) suggest, it can include the things that ought to be taught as well as what actually is taught, and all the factors that affect these.

In the Queensland senior curriculum there are Board syllabuses for senior Board subjects, and broader curriculum framework documents for other kinds of senior subjects called ‘Board-registered’ subjects (not included in the calculation of statewide rankings for tertiary entrance) that state their intentions explicitly. Substantial numbers of senior students in Queensland and other states also take
individualised learning programs with, for example, a remedial literacy and numeracy focus.

Evaluation of the senior curriculum by the Board is less concerned with the difference between what is and what ought to be, than the extent to which what ought to be taught is communicated effectively and works well for teachers and their students.

Evaluation is equally elusive in the research literature. It is often linked in writings on curriculum evaluation with notions of accountability. Kemmis (1986, pp. 117–118) and other commentators refer to a range of definitions, including definitions that suggest ‘the role of evaluation in informing action at discrete decision points … it thus has a major role to play in articulating justifications of action’.

This definition seems to resonate with the Queensland context discussed in this paper. Curriculum evaluation is a term that refers not to evaluation of teachers but to evaluation of a key tool for teaching—the syllabus documents. Curriculum evaluation is the process of providing evidence of relevant community experience and views, and recommendations based on that evidence that may be used in the curriculum decision-making process. Curriculum evaluation in this context is not decision making, but rather a specialised kind of research with its own particular protocols and procedures.

The research literature on curriculum evaluation often refers to Tyler’s 1949 definition of evaluation as the extent to which ‘educational objectives’ are realised. Tyler is sometimes represented as the ‘founding father’ of the educational objectives/outcomes/benchmarking movement. Syllabuses for Board subjects in the Queensland context do in fact provide a detailed account of what students ought to know and be able to do: this is the motivating logic of the unique system of education based on criteria and standards that is in place in senior schooling in Queensland. In this sense it is true that Queensland has had ‘outcomes’-based education since the 1970s, and that this outcomes-based education has undergone considerable refinement and development over this period. Curriculum evaluators are also concerned with objectives in the sense that they evaluate how effectively these objectives are communicated by the syllabus and translated into practice, where such translation offers information about the soundness of the syllabus document.

However, as will be argued, the Queensland model of curriculum evaluation cannot be simply or easily described in terms of the polarised arguments about the different models of curriculum evaluation that can be found in the research literature.

What is the history of curriculum evaluation?

The 1960s saw an increase in investment in education and a corresponding development of the rhetoric of accountability so that, by the early 1980s, as McCormick and James (1990, p 7) have observed, ‘the concept of educational evaluation was often so interwoven with the concept of accountability that the two were difficult to distinguish’. Yet, as they add, the development of education evaluation marched in time with other broader social changes in the West: a broader application of notions of accountability to public institutions that emerged somewhat earlier in the USA than in the UK. These notions of accountability were part of the response of Western societies to pressures for increased economic performance, and the accusation that education did not sufficiently meet the needs of industrialised societies. When we consider what Habermas (1974) has described as the ‘crisis of late capitalism’—the loss of faith in specialists—and what many commentators such as Atkin (1979) have described as the fragmentation of our society, a moving away from any notion of a common good, we can see the wider context in which these notions of accountability in education developed.

John and Donaher (1996, p. 2) have argued that the traditional objectivist approach developed by Tyler, specifically his series of lectures entitled ‘Basic principles of curriculum instruction’ which
offered an essentially ‘scientific’ and ‘instrumentalist’ way of evaluating curriculum on the basis of explicit objectives, has found particular appeal in the contemporary world:

The global economic crisis of the late 1980s and the subsequent reconceptualisation of education as a key player in national economic sustainability, created the perfect cultural climate for an increased dependence on ‘objectives’ evaluation approaches.

It seems true to say that the development of notions of education evaluation in the USA and the UK seem linked to the language of business and industry (inputs, outputs, outcomes, benchmarks). The development of describable outcomes has focused practices of curriculum evaluation, for such outcomes supposedly make clearer what curriculum is designed to achieve. That is, the establishment of describable outcomes brings with it a commensurate need to monitor their achievement.

John and Donaher (1996, p. 2) argue that such approaches to curriculum evaluation have been associated with what they call ‘mainstream, formal educational evaluation’ sponsored by ‘government, institutional and faculty leaders and educational administrators and managers’ who have particular perceptions of ‘quality and quality measurement’. They argue that since the 1970s, specifically the formulation of ‘the democratic alternative’ in late 1972 by Bob Stake and others at Cambridge, more democratic and participative models of evaluation have developed, involving participation of all stakeholders. The literature suggests varied opinions on the impact and influence of these alternative models.

An appreciation of the difference between the models involves some understanding of deeper theoretical differences, tensions and developments.

Over the last few decades the research conducted in the social sciences and in education has been influenced by the development of postmodernist approaches that are often juxtaposed with ‘positivist’ or ‘traditional liberal humanist’ approaches.

The positivist approach is informed by Western philosophies that, to use the simplest possible characterisation, emphasise a reality that exists outside the observer and can be objectively measured. Much scientific research has been informed by this world view.

The postmodernist approach emphasises the nature of reality as a social construct, and moreover as a reality that is actively constructed by both the writer and reader of a text, whether written or otherwise. It has been variously described as the critical–reflective approach and associated with philosophers such as Habermas, Derrida, and Foucault.

The distinction is a crude one, and has sometimes been associated with the distinction between quantitative and qualitative approaches, and the distinction between traditionalist and democratic approaches to evaluation, as well as the distinction between policymakers or deciders and those wanting to subjugate ‘dominant power structures’. In the evaluation literature (and in other literatures and disciplines), the tension between research that is what House (1986) describes as ‘intimately tied to the initiatives of governments’ and other kinds of intellectual endeavours and research is apparent.

What we can say with some certainty is that the evaluation literature seems to be defined by a tension, and a lack of dialogue between those who, like House (1986, pp. 6–7), state in various ways ‘that the

---

1 Why the distinction between quantitative and qualitative research is tangled up with this debate is suggested by House (1986, p. 8):

One of the hottest controversies has been the use of qualitative as opposed to quantitative evaluation techniques. Quantitative techniques have long been considered the sine qua non of modern social science, largely because their advocates believe that quantitative techniques ensure a degree of objectivity that any science must possess. This position has been assailed by advocates of qualitative studies, who have been successful in establishing qualitative approaches as more or less legitimate ways of evaluating.
industrialisation of education is not a trend of which I approve nor a use of evaluation which I endorse’ and those who conduct evaluations that can be described variously as evaluations of government initiatives. At stake are some deep philosophical and political differences, rather than superficial arguments about research techniques.

The evaluation literature seems, then, to be defined by a theoretical–political tension between positivism and postmodernism. Some commentators have argued that the postmodernist approach to evaluation has failed to powerfully influence the direction of education policy decision making. ‘The intent and profile of mainstream curriculum evaluation is now focused on a new era of “quality” measurement that is firmly embedded within a scientific, economic and technological rationale.’ (McCormick and James, 1990, p. 6) In contrast, March and Willis (1999, p. 270) argue more recently that ‘numerous forms of qualitative research and evaluation have been developed and now exist side by side with traditional quantitative forms’. Others will argue that curriculum evaluation is now dead.

That this polarised literature fails to account for the nature and complexity of at least some evaluation practices can be demonstrated by exploring the detail of the model of curriculum evaluation that exists for the Queensland senior curriculum. Curriculum evaluation in this state is very much alive, operating in a system in which the purpose, structures, procedures and protocols are highly defined. However, within this system, evaluators operate somewhat eclectically, using quantitative and qualitative approaches to triangulate their data, evaluating syllabuses with stated objectives but also evaluating many other things, involving teachers and students and other stakeholders in highly participative research, and working closely with decision makers to help inform judgments about curriculum changes needed.

How can we conceptualise different models of curriculum evaluation?

The large body of academic and government-commissioned research and curriculum evaluation provides an array of models with different and conflicting claims about their utility and value. Yet as the previous discussion might suggest this does not mean that those who conduct public sector research reflect on theoretical debates about evaluation or are confident in their knowledge and use of curriculum evaluation techniques. As Withers and Batten (1988, p 13) have argued:

A major gap appears to exist in the range of procedures available to departments and other authorities in conducting curriculum projects and fostering sound developments. This is the lack of—or lack of knowledge of—a wide range of models for evaluation of curriculum principles and procedures, for use at central and school levels.

Clearly, reflection on these models involves some enabling constructs, a language in which to conceptualise evaluation. There are, of course, different ways of defining the different models, but perhaps the first question is when is a model a model? As Stake (1981) pointed out almost 20 years ago, there has been much written about the different evaluation ‘models’ but these models often lack the detail that is implied by the term model and might more correctly be described as ‘persuasions’.

The detail offered on the Queensland model does suggest that it is a fully developed model in this sense: there is considerable complexity to this model.

Nevo (1986) defines ten somewhat unwieldy ‘dimensions’ for his analysis of evaluation literature, but a simpler method for distinguishing the models might be obtained using the words value, valuation and valuer. The intention here is to suggest the utility of these three simple terms to conceptualise evaluation activities, rather than offer a summary of available evaluation models—although a scan of at least some ‘traditions’ of curriculum evaluation, such as Tyler (1949), Stake (1967), Parlett & Hamilton (1972) and Eisner (1977), suggests that all these different models can be distinguished in
these terms. The conceptualisation suggests that curriculum evaluation is permeated with value, that the consideration of value is not some separate act that lies outside other research activities, but rather that everything the evaluator does is in some sense about values.

**Value**

*Value* refers to the primary value of the evaluation, whether explicitly or implicitly stated, and the subsidiary or related set of values.

The previous discussion linked evaluation to accountability. In the Queensland senior curriculum, evaluation is based on a particular notion of accountability. The Queensland Board of Senior Secondary School Studies (the Board) reports to Queensland’s Minister for Education, and the Queensland parliament, yet it is also in some direct ways accountable to a community of teachers and, ultimately, students. The primary value for curriculum evaluation is that curriculum documents must work well for teachers, students, and the community, and represent the best possible synthesis of sometimes disparate views about a subject. Curriculum evaluation is evaluation of one of the Board’s products—the curriculum it designs for teachers to deliver to students. This primary value explains why a key subsidiary value used to define research questions for curriculum evaluation is the view that syllabus documents must work well for teachers designing work programs for their students. (Work programs are the school’s translation of the syllabus into the school’s context, incorporating details of the learning and assessment that their particular students will undertake. Work programs are approved by the Board on the advice of its district and state review panels (panels of experienced teachers)).

**Valuation**

*Valuation* refers to the totality of what is evaluated: the content of the evaluation, what the evaluation will include and what it will not include. The content is a statement of the value that has been placed on the total possible field of things that could have been researched. Discussion of each thing included in the study is an act of valuation, of weighing the importance and significance of the evidence.

There are some important differences in what is included in evaluations. For example, one model that has received little attention is one that operates at the school level and involves schools evaluating the particular combination of curriculum options taken by their students i.e. a whole curriculum evaluation model. McCormick and James (1990, pp. 46–47) argue that in the UK (at least) ‘many teachers, whether in primary or secondary schools, are, by tradition, unfamiliar with discussions that involve the whole curriculum of the school’. They argue that, in the UK, public examinations based on subject distinctions have represented a major constraint to teachers taking a whole-curriculum approach.

In Queensland, while curriculum evaluation is, at the simplest level, about the syllabus document and its implementation, it includes anything and everything that the evaluator judges relevant to the sound development of the subject. Evaluators have detailed guidelines about what must be included, but they also have considerable scope to include other things. The content of the evaluation, whether it exists at the level of the report or as a spoken text when the findings are explained to the Board’s subject advisory committees and its Curriculum Committee, is actively weighed in terms of its value and significance.

**Valuer(s)**

Models can also be defined in terms not simply of who conducts the evaluation, but of all those involved in considering and assessing the evidence of the evaluation. The evaluator may be only one of a number of people who act as valuers.
Discussions in the literature of the ‘who’ question in curriculum evaluation have sometimes focused upon defining the evaluator as an ‘insider’ or an ‘outsider’ or as a ‘professional’ evaluator or ‘amateur’ evaluator (Nevo, 1986).

There has been little discussion of the significance of the total set of people who act as valuers, although the existence of these groups in the Queensland system is a crucial feature of its participative practices of curriculum development. Valuers can be described not simply as decision-makers, but rather the group of people who are involved in making value judgments about the curriculum being evaluated.

In Queensland then, the evaluator is only one of a few individuals who act as valuers in this sense. There are, as we shall see, a number of interlocking committees comprised of community members (primarily teachers) with which the evaluator interacts, and a process of collective monitoring of, and making judgments about, the curriculum under consideration.

Figure 1 represents this conceptual device for distinguishing the different models of evaluation—an

![Figure 1: The evaluation triptych](image)

What model of curriculum evaluation has been used by the Queensland Board?

This section gives an account of the Queensland model that emphasises the ‘nuts and bolts’ of the system. Why bother with the detail of structures and procedures? Because it seems that people tend to agree about things when they are talking at the most abstract level: it is when we examine the detail of a system that we can see how it is different, and how abstractions and ideals have been translated into operational realities.
The context

The Board operates in a context in which evaluation of the workability of curriculum documents has a particular importance, for these documents are one of the cornerstones of its system of school-based assessment. Curriculum evaluation occurs in the context of a unique system of moderated school-based assessment: there are no public examinations in Queensland. Accordingly, syllabuses must be workable, not least because they have to communicate standards for student achievement that will be comparable across the state.

Syllabuses

Syllabuses are documents that provide a curriculum framework so that schools can prepare work programs for senior Board subjects. Board subjects are subjects included in the calculations for tertiary entrance results (Overall Positions (OPs), and Field Positions (FPs)). A syllabus for a Board subject will typically include:

- a rationale or justification for including the subject in the senior school curriculum
- global aims or broad, long-term achievements, attitudes and values to be developed by students (these are not directly assessed)
- general objectives which the school pursues directly and which are assessed: these divide into ‘process’ or cognitive skills, ‘content’ objectives (relating to factual knowledge), ‘skill’ objectives (relating to practical skills), and ‘affective’ objectives which are not assessed, but relate to important values and attitudes to be developed by study of the subject
- subject matter
- core requirements
- specifications such as minimum number of timetabled hours the subject must entail
- advice about construction of work programs
- advice about development of learning experiences
- advice about assessment (appendix 1 gives the Board’s policy on assessment): assessment techniques, criteria and standards for the award of each of five levels of achievement and any advice about ways of deciding levels of achievement; requirements for folios provided to district review panels as part of the Board’s moderation procedures
- work program requirements
- Board policy on language education
- policy on quantitative concepts and skills
- statement on educational equity
- a list of resource material.
Appendix 2 gives the criteria for a Board subject.

**Structures and procedures for syllabus development**

Processes of syllabus development can relate to small or large changes to the curriculum. The following discussion emphasises what happens with larger curriculum developments. There are many players in this process, which is highly participative: the evaluator, the range of community members with which the evaluator works to gather evidence and conduct the evaluation, the Office of the Board, the relevant subject advisory committee (SAC), the Curriculum Committee which reviews the advice of the SAC in the context of the whole curriculum, and the Board itself, which is a committee legally constituted as representative of community views.

The discussion that follows cannot hope to offer the detail of what each of these Board structures contributes, but rather focuses on key procedures of curriculum development, the implications of these for curriculum evaluation, and the role of the evaluator.
New Board subjects

Syllabus development involves subject advisory committees (SACs): much of the work of the Board is undertaken by these committees, comprised largely of members of the teaching community, usually teachers with considerable experience.

Every six years Board syllabuses are reviewed by an SAC which recommends to the Curriculum Committee a range of options from ‘no change needed’, through minor or major changes, to development of an entirely new syllabus. All new syllabuses require trialling and piloting. However, anyone may propose a change to an existing or new syllabus and this will be considered by the SAC and evaluated by the Curriculum Committee. In this sense the process of curriculum development has an open rather than closed nature.

The production of a new syllabus can be undertaken by the originators of an approved proposal, working with the SAC, or the syllabus may be developed by the SAC. If the Board accepts the new syllabus as being potentially viable, and if at least six schools make contractual arrangements with the Board, the syllabus is trialled. This highlights a key feature of the Queensland senior system central to the nature of curriculum evaluation in this state—a syllabus does not ‘get up’ if there is no grass-roots support for it.

Evaluators are appointed for a new syllabus trial, which occurs for a nominal period of two years. The report of the evaluator is presented in the first instance to the SAC (the report is never a real surprise as the evaluator works with an evaluation committee that includes members of the SAC). From this point the syllabus may be retrialled if it is found to be substantially flawed, or it may become a pilot syllabus. For this stage of syllabus development, 15–25 schools are required for an evaluation period of two years. This stage continues the ‘development and dissemination of the standards of assessment with the state review panel and district review panels’ (Dudley & Abbey 1999, p. 3). The Board’s moderation of standards in Board subjects occurs through over 3000 district review panels which scrutinise complete samples of student work and provide advice to schools about how well they have applied standards.

If the indications from the pilot stage support a decision of ‘minor alterations only’ recommended by the pilot evaluation, the revised syllabus is made available to all schools.

The above detail about the process of developing a new Board subject suggests some important implications for curriculum evaluation of new subjects:

- the evaluator is evaluating a new curriculum option for which there is already some grass-roots school support
- the thing being evaluated—the new curriculum option as detailed in the syllabus document—is based on a protracted process of consensus-making involving community members, and teachers on committees drawn from the teaching community that may deliver the syllabus
- the evaluator, therefore, does not so much drive the process of curriculum evaluation as act as an important player assisting a large and diverse group of decision makers at several levels.

Major revisions of Board subjects

Once the Board has accepted a proposed major revision of an existing syllabus as potentially viable, the syllabus undergoes a ‘combined trial-pilot’ which involves 15–20 schools and an evaluation period of two years. If the evaluation procedure discloses substantial faults with the syllabus, the rewritten syllabus undergoes additional trial-pilot stages.

2 For syllabuses with embedded vocational and training components, a review occurs every four years.
A syllabus undergoes another trial-pilot. If the evidence of the evaluation suggests that only minor revisions are needed, the revised syllabus is made available to all schools for general implementation. The 'old' syllabus will be kept in use for three years after the approval of a revised syllabus for general implementation.

For curriculum evaluation the process of revision of existing syllabuses has the implication that even for very small changes, the evaluation exercise is not so much an instrument for telling the community what changes ought to be made, but rather a conduit from the community of users of a syllabus to groups of decision makers drawn from that community of users. That all changes, even very small changes, must be made via certain participative structures and processes means that evaluation ideally becomes less an exercise in academic judgment and more an exercise in gathering evidence and drawing connections between that evidence and the recommendations for change.

**New and revised study area specifications**

The process described above is a little different for subjects that are not Board subjects. In Queensland, at the time of writing in 1999, 31 new senior subjects exist in the form of 10 new study area specifications (SASs) with a high accredited vocational education content. Even though six Board subjects have 'embedded' accredited vocational education, many more students participate in vocational education by enrolling in SASs (over 20,000 students took at least one SAS in 1998). In May 1999 the Board published (Bell, Williams & Paties 1999) what is effectively one of the largest single studies of the implementation of VET in senior schools in the form of an evaluation report of 28 of these subjects.

Assessments in study area specifications are not moderated: the quality assurance for assessment standards is that which exists under current arrangements for the delivery of accredited vocational education and training. Consequently the emphasis of the evaluation of these subjects is a little different, although the broad procedures and structures for developing a new SAS are similar.

Figures 3, 4 and 5 represent these processes of syllabus development. Figure 3 shows the sequence of development of Board syllabuses, figure 4 shows the relationships of the different functions performed by the office of the Board during trialling, and figure 5 shows these relationships during piloting of a syllabus.
**Figure 3: Sequence of the development of Board syllabuses**

<table>
<thead>
<tr>
<th>Syllabus for general implementation</th>
<th>Available to all schools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-pilot</strong></td>
<td>Implemented by 6–14 schools. Must proceed to pilot when n=15.</td>
</tr>
<tr>
<td><strong>Pilot</strong></td>
<td>Implemented by at least 15 schools. Evaluated.</td>
</tr>
<tr>
<td><strong>Trial</strong></td>
<td>Implemented by 6–15 schools. Evaluated.</td>
</tr>
<tr>
<td><strong>Pre-trial</strong></td>
<td>Implemented by fewer than six schools. Must proceed to trial when n=6.</td>
</tr>
<tr>
<td><strong>Syllabus Restricted Use</strong></td>
<td>Advertised</td>
</tr>
<tr>
<td><strong>Major revisions of existing syllabuses</strong></td>
<td>Advertised</td>
</tr>
<tr>
<td><strong>Trial/pilot</strong></td>
<td>Implemented by 15–25 schools. Evaluated.</td>
</tr>
<tr>
<td><strong>Syllabus Restricted Use</strong></td>
<td>Available to all schools</td>
</tr>
</tbody>
</table>

Figure 4: Interrelationships among the Board’s functions during syllabus trialling

Figure 5: Interrelationships among the Board’s functions during syllabus piloting

The evaluation process

The purpose of evaluation in this context is to provide information to an SAC that will help develop an effective syllabus. As noted previously, an effective syllabus is one that, among other things, communicates to teachers in ways that allow them to prepare effective work programs.

The importance of this criterion for evaluating an effective syllabus cannot be overstated in the Queensland context, for Board syllabuses set the parameters for the preparation of work programs. Syllabuses do not provide the fine detail of learning and assessment experiences. Teachers must be able to translate the syllabus into detailed work programs that, among other things, represent a sound interpretation of the exit criteria and standards for awarding levels of achievement that are comparable across the state. Student work is seen as comparable if student work in the same subject across the state awarded the same level of achievement is substantively the same in the sense that it meets the same criteria of achievement (see appendix 1 for the Board’s assessment policy).

Although some models of curriculum evaluation focus on evaluation of teachers, the thrust of curriculum evaluation here is on scrutinising how well a key teaching tool—the syllabus—works for teachers. Essentially the Board is evaluating how well one of its products is working for the community of teachers for which that product was designed.

Evaluation of trial Board syllabuses

There are some key aspects of the syllabus that evaluators are asked to evaluate in a trial evaluation:

- internal consistency
- communication of intentions
- breadth and depth
- sequencing of subject matter
- relevance of the subject to students
- appropriateness of assessment requirements
- resources schools find useful.

In August of each year of a trial, the Board makes a decision, on the basis of the evidence provided in the June interim report of the evaluation, to maintain, speed up, or slow down the progress of the trial. In November of the final year of the trial a progress report is provided by the evaluator to signal further recommendations and facilitate redrafting before the March meeting of the SAC the following year.

Evaluation of pilot syllabuses

When a pilot syllabus has been approved, applications for evaluation are called. The pilot evaluation focuses on:

- an examination of work programs: their nature and development, and implementation in different schools, including relevant resources
- clarifying the standards for student achievement.

The reports, which are produced under the same schedule used for trial syllabuses, will reflect this focus.
Evaluation of trial-pilot syllabuses

As noted previously, evaluation of trial-pilot syllabuses relates to major revision of an existing syllabus. The processes described above for both the trial and pilot syllabuses apply here, as the purposes are a combination of both.

Evaluation research protocols

Scrutiny of procedures manuals and evaluation reports, as well as experience over the last few years as an evaluator of new subjects developed under SASs (Bell, Williams and Paties, 1999), suggests a few key Board protocols of curriculum evaluation that can be used as a preface to discussion about some features of reports produced by evaluators:

• **Evaluators operate with clearly defined research questions.** The form and nature of the evaluation report is clearly defined. The nature of any information required (such as a list of schools involved in the trial) is specified under particular sections, as well as the form of the recommendations (evaluators are required to make specific suggestions for modifications).

• **Evaluators are expected to use a broad range of research techniques to find the answers to specific questions.** The form and nature of the report is defined, but the techniques of research are not, except in the sense that they must represent a valid basis for recommendations.

• **Evaluators must accurately convey the views and experiences of teachers in schools as well as other community members (such as students, parents, employers, guidance officers, school administrators) to the SAC and the Curriculum Committee.** In this sense what matters is not so much what curriculum development the evaluator thinks should take place, but rather what sound judgments about the evidence for curriculum development the evaluator can make, in a context in which that evidence is the views and experiences of stakeholders.

Some features of reports produced by evaluators

In the evaluation literature it seems there is an abundance of theoretical reflections but a paucity of actual reports of evaluations of curriculum, a paucity that has been noted by commentators for at least two decades.

This section offers a summary of the results of a scan of evaluation reports to determine some of the key features of research reports produced by evaluators in Queensland since 1995. Such a scan can be part of a consideration of the Queensland model of evaluation leading to identification of some elements of effective curriculum evaluation.

There are, first of all, clear differences in the extent to which evaluators ‘theorise’ the research they undertake, and the extent to which they reflect explicitly in their reports upon the model of curriculum evaluation they are using. Gifford’s 1995 evaluation of the trial-pilot senior syllabuses in music offers a brief discussion of different models for curriculum evaluation, making some mention of the weaknesses and strengths of particular models and finally settling on an ‘eclectic approach’ (pp. 6–7). Some evaluators do refer explicitly to particular models—two examples of action research are given below—but when the broad body of curriculum evaluations is considered, their theoretical orientations differ such that collectively they cannot easily be characterised in terms of a single approach. This is an arguable strength of the system: that within a highly structured system of curriculum development, there is a plurality of approaches, whether we are talking about implicit or explicit theoretical orientations.

Evaluation reports typically include a broad range of modes for collecting information: closed or open ended questionnaires (qualitative and quantitative), structured or unstructured interviews,
observational techniques, archival research, analysis of published or unpublished documents (such as work programs, student work, teaching materials, student records, and teacher notebooks). Again, this plurality of techniques seems to be a healthy element in the Queensland system of curriculum evaluation.

The emphasis in most evaluations is on triangulation of data (a key word in the language of evaluators), not only through the range of modes, but also use of quantitative and qualitative data (see Beasley and Jess’s 1998 evaluation of the trial-pilot Agricultural Science syllabus and also the 1998 evaluation of the pilot syllabus in Engineering Technology by Butler and Jess). Data are also typically collected through several stages that are carefully documented by the evaluator (see, for example, Norris’s 1996 evaluation of the trial-pilot senior syllabus in Film and Television, p. 4).

Of course, a range of modes for collecting data and use of triangulation of quantitative and qualitative data are not all that is required for an effective evaluation: also important is the clarity with which the evaluation is conceptualised. In their 1995 evaluation of the pilot syllabus in Chinese, Peckman and Peckman describe a framework comprised of:

… the key elements of the curriculum development and implementation process which were seen to be instrumental in achieving effective outcomes; viz syllabus, work program, and the school program. It also includes four factors which influence the effectiveness of some of these elements; viz. the characteristics of the students, teachers and the school and the quality of resources available. (p. 3)

Figure 6 offers a visual representation of this framework which has been adopted by other Board evaluators (for example, Sadler in the 1998 evaluation of the trial-pilot syllabus in Business Communication and Technologies, and Underwood in the 1998 evaluation of the trial-pilot senior syllabus in Economics, and the 1998 evaluation of the trial senior syllabus in Futures).

![Figure 6: Conceptual model adopted for the evaluation](source: Peckman, R.S. & Peckman, G.I. 1995, Evaluation of the Pilot Senior)
The Peckmans used the framework shown in figure 6 to identify the variables (see appendix 3 for a list of these variables). Bushell uses the framework in her 1999 evaluation of another Board syllabus—the trial-pilot syllabus in English—as well as the trial extension Literature syllabus (also 1999). She implements this using an ‘action research model’ which she describes in the English evaluation in the following manner using the work of Kemmis as a point of departure:

Action research involves groups of teachers systematically analysing an educational problem of concern to them, planning action programs, executing them, evaluating their efforts, and then repeating the cycle if necessary … The evaluator is an external researcher who provides tools and opportunities for teachers to reflect upon their experiences in the implementation and who records, interprets and supplements this information with data obtained from further enquiry (p. 57).

Hockey (1998) also used a model of community-based action research, based on the work of Stringer (1996) for the evaluation of the trial senior syllabus in Aboriginal and Torres Strait Islander Studies. There is a sense in these evaluations that, where different theoretical models are being used, the evaluator is using the principle of ‘fitness for purpose’.

There are differences in the extent to which evaluators engage with the more theoretical pedagogical arguments in each subject teaching community, apparently related to the extent to which the subject is ‘theorised’ in the teaching community itself. The evaluator has a role in helping the Board synthesise a syllabus that will represent the best possible consensus. For example, Bushell gives a list of issues defining her ‘assignment’ in her evaluation of the trial senior syllabus in Literature, the first of which is “… the rhetorical and theoretical/pedagogical coherence and consistency of each syllabus document” as well as ‘how pedagogy is being shaped under the new syllabuses’ (Bushell 1999, p. 1).

The independence of evaluators

No discussion of curriculum evaluation would be complete without reference to the question of independence. How do evaluators of Board curriculum stay independent? What does it mean in this context to say research is independent?

The particular structures just described suggest some important ways in which the system aims to secure the independence of evaluators. For example, those who design or develop the curriculum document do not directly employ the evaluator, and this makes it difficult for those who have an investment in the curriculum document to inappropriately influence the evaluation. The evaluator is employed by the Office of the Board, and it is very much in the interests of that office to have workable curriculum documents. The Office of the Board has no interest in employing evaluators who provide a less than accurate assessment of the curriculum being developed—the Office expects evaluators to advise it when a curriculum document is not working. An evaluator who ‘gets it wrong’ is one who recommends that a syllabus document should proceed when that syllabus is not working in some important respects.

Further, the multilayered and diverse nature of the committee structures with which evaluators work makes it difficult for any one individual to inappropriately influence the reception and consideration of a report.

The work of Michel Foucault suggests that to understand power we must understand the particular details of a situation, not apply universalising abstractions about how power always works. In the case of Board evaluators, the checks and balances previously described suggest why we should question...
assumptions that public sector research is *always* severely compromised, and generalisations about public sector research that are mechanically applied across this whole sector. At the very least, dismissing the possibility of intellectual independence in public sector research makes anyone who is concerned about intellectual independence less receptive to exploring the particular nature of public sector situations where such independence is recognised and valued.

**Elements of effective curriculum evaluation**

What is effective curriculum evaluation in the Queensland context? The short answer is that effective curriculum evaluation is evaluation that tells us what’s working, what’s not working and how to make non-working things work.

The long answer is that the previous discussion offers a list of some possible elements of effective curriculum evaluation that help us understand what this means. The ‘triptych’ used earlier (*value, valuation* and *valuer*, see figure 1) could be used to organise each of these elements under some headings, although of course some elements seem to belong under more than one heading. A list of some of the elements of effective curriculum evaluation might include, but not be limited to, the list shown in the following section.

**Effective curriculum evaluation**

Effective curriculum evaluation involves:

1) **making values explicit such as:**
   - agreed-on primary motivating values of the system of curriculum evaluation in ways that demonstrate links to research questions, research expectations and protocols (for example, the freedom to adopt different theoretical orientations)
   - values implicit in the curriculum

2) **valuation decisions about:**
   - what to include in a synthesis of different perspectives in a consensus-making process involving community members and teachers that, while not being free of values itself, helps to resolve value conflicts
   - which whole-school issues to include
   - which data to gather at what different stages in the use and development of a syllabus
   - what teaching methods are relevant to an understanding of the effectiveness of the curriculum being evaluated
   - what theoretical models of evaluation to include on the basis of a ‘fitness for purpose’ principle
   - what theoretical pedagogical debates to include on the basis of a ‘fitness for purpose’ principle

3) **valuers who:**
   - seek clear conceptualisation of the evaluation and have a clear sense of the research variables
   - use a broad range of modes for data collection
   - use explicitly documented research techniques to assess the validity and significance of evidence
   - triangulate the evidence using quantitative and qualitative data
   - draw sound connections between evidence and the recommendations for change for the
consideration of groups of decision makers at several levels

- act as a conduit from the community of those with a stake in curriculum change to groups of decision makers drawn from this larger community.

Conclusion

In conclusion, and recalling our opening discussion of the polarised nature of evaluation literature, it seems that there has been little discussion of the detail of particular models of curriculum evaluation—the specific elements of those models that make them work well in particular contexts.

When we examine models of curriculum evaluation such as the one offered by the Queensland senior curriculum, we can see a small universe at work, in this case one that is shaped by the needs and nature of its highly participative system of school-based assessment.
References


Stake, R.E. 1967, ‘The countenance of education evaluation’, *Teachers College Record*, vol. 68, no. 7.


**Further reading**


APPENDIX 1: Assessment policy

Below are the underlying principles of the Board’s policy on assessment promulgated to schools:

- Exit assessment be devised to provide the fullest and latest information on a student’s achievement in the course of study.
- Assessment of a student’s achievement be in the significant aspects of the course of study identified in the syllabus and the school’s work program.
- Information be gathered through a process of continuous assessment.
- Selective updating of a student’s profile of achievement be undertaken over the course of study.
- Exit achievement levels be devised from student achievement in all areas identified in the syllabus as being mandatory.

**Balance of assessments is a balance over the course of study and not necessarily a balance within a semester or between semesters**

APPENDIX 2: Criteria for Board subjects

Educational criteria

The following criteria are to be addressed during the stages of syllabus development. All committees use this checklist in their recommendation of syllabuses for general or trial-pilot implementation.

A Board subject should contribute to the education of students by providing:

- opportunities to develop a range of intellectual, technological and operational skills, including the key competencies\(^3\)
- the best possible balance of practical experience, abstraction and reflection appropriate to the subject
- opportunities to acquire the specific knowledge and skills in the subject
- a sound basis for developing values and attitudes appropriate to students’ future participation in a democratic and pluralistic society, of which the work environment is a significant component
- challenges appropriate to the developmental level of the students for whom the subject is designed
- opportunities for students to develop the highest level of literacy and numeracy skills in the context of the subject\(^4\)
- opportunities which encourage the development of, and understanding and respect for, our heritage and cultural diversity, including Aboriginal and Torres Strait Islander and ethnic communities
- scope for critical thinking and the generation of questions, ideas, goals and visions of the future
- access for a range of students, including those with disabilities, while maintaining the integrity of the subject and the Senior Certificate\(^5\)
- appropriate articulation with P–10 syllabuses and post-school environments (higher education, further education, work and leisure).


---

\(^1\) The key competencies are: collecting, analysing and organising information; communicating ideas and information; planning and organising activities; working with others and in teams; using mathematical ideas and techniques; solving problems; and using technology.

\(^2\) See also the Board’s statements on Language Education and Quantitative Concepts and Skills

\(^3\) See also the Board’s statement on Educational Equity
APPENDIX 3: Evaluation Variables

Student characteristics
- Student gender
- Student background in Chinese language
- Students’ expectations of studying Chinese language
- Student interest in Chinese
- Students’ future occupational and study plans
- Students’ academic ability

Teacher characteristics
- Pre-service and in-service education
- Proficiency in teaching Chinese
- Proficiency in developing a work program
- Teaching experience

School characteristics
- Attitudes of school administration
- Expectations about Chinese students
- Facilities

Senior Chinese Syllabus
Adequacy of the syllabus:
- for developing a work program
- to guide in the selection/formulation of appropriate learning experiences
- to guide teachers in assessing students

Work program
- Quality of work programs
- Development process
- Adequacy of the learning experiences included
- Adequacy of assessment practices in the program
School programs
- Assessment procedures
- Adequacy of the learning programs
- Content selection

Teacher resources
- Availability of teaching resources
- Access to professional assistance
- Access to professional development

Outcomes
- Students’ and parents’ attitudes to Senior Chinese
- Students’ perceptions of their ability
- Students’ and teachers’ perceptions of the outcomes