The constraints on delivering public goods – a response to Randy Bennett's 'What does it mean to be a nonprofit educational measurement organization in the 21st Century?'
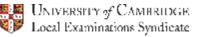
Tim Oates
Cambridge Assessment

A paper presented at the International Association for Educational Assessment
Annual Conference, Baku, Azerbaijan, September 2007.

Tim Oates,
Group Director
Assessment Research & Development
Cambridge Assessment
1 Hills Road
Cambridge
CB1 2EU
UK
Direct dial. +44 (0)1223 552750
Fax. +44 (0)1223 552700
Email: oates.t@cambridgeassessment.org.uk
www.cambridgeassessment.org.uk

UNIVERSITY of CAMBRIDGE
Local Examinations Syndicate

The constraints on delivering public goods – a response to Randy Bennett's 'What does it mean to be a nonprofit educational measurement organization in the 21st Century?'

Abstract

Very different structural arrangements exist, in different national settings, for the provision of assessments and qualifications. Whilst international convergence is occurring in the wake of global economic developments and the emergence of new public policy arrangements such as transnational qualifications frameworks, the unique composition of national systems remains worthy of analysis. Some national settings such as England are characterised by highly complex arrangements. In such settings, subtle relationships – heavily determined by their historical background - carry the responsibility for delivering public goods in respect of assessment and qualifications. This presentation will argue that issues of control, accountability, and sources of innovation are not always transparent, even to the key agencies within the system. Within this, the issue of the profit/not-for-profit status of organisations is less significant than their explicit and implicit structural positioning.

The paper argues that it is important, but insufficient, to analyse the mission and intentions of an assessment agency in respect of legislation which bears directly on it. It is essential to look beyond this, to understand the complexity and volatility of key factors in national structures and settings which constrain and empower agencies. The paper argues that extrapolation from one national setting to another only becomes possible when analysis is based on identification of 'control factors' and how they combine in specific national settings. Without this approach – drawn from mature transnational analysis methodology – there is an acute risk of reaching inappropriate conclusions about the 'space' for the operation of independent assessment agencies and viable modes of operation. Finally, the paper identifies subtle shifts in vital, informal accountability mechanisms.

Acknowledgements

Background Note

Awarding Bodies – assessment agencies
The three main awarding bodies offering the bulk of qualifications used in compulsory and 16-19 education in England are bodies independent of Government, but operate under the regulatory mechanisms operated by the Qualifications and Curriculum Authority – a body operating by the authority of national legislation. It is known as a 'Non-Departmental Public Body' (NDPB). Its board is appointed by the Secretary of State for Education.

The three main awarding bodies AQA, Edexcel and OCR differ in the nature of their constitution and governance.

> AQA is an educational charity and states its purpose as: *to advance education for the public benefit in ways which include operating GCSE, A Level and other qualifications and providing support for teachers and learners*. AQA was formed on 1 April 2000 following a merger between the Associated Examining Board (AEB) and the Northern Examinations and Assessment Board (NEAB).

> Edexcel was formed in 1996 by the merger of the Business & Technology Education Council (BTEC), the country's leading provider of vocational qualifications, and the University of London Examinations & Assessment Council (ULEAC). In June 2003 the Edexcel Foundation entered into a partnership arrangement with Pearson PLC, the biggest educational services company in the world, to set up a new company called London Qualifications Ltd, which trades as Edexcel.

> OCR is an educational charity formed in 1998, through the merger of Oxford & Cambridge Schools and the Royal Society of Arts. OCR is part of Cambridge Assessment, a department of the University of Cambridge. Until 2005, Cambridge Assessment was known as UCLES (founded in 1857). Its stated objective is *'to ensure all learners across the world are able to access the benefits of their education through receiving fair and accurate assessment'*.

For the sake of brevity, I refer to all three bodies in this paper as 'assessment agencies'.

Public examinations – qualifications

Developed and administered by AQA, Edexcel and OCR, GCSEs and GCEs are the main general academic qualifications in compulsory education in England. These are essentially subject-based examinations.

Most 16 year olds in England take General Certificate of Secondary Education (GCSE) examination(s) and coursework unit(s) in several school subjects at the end of compulsory schooling. Typically, pupils take 10 or more of these qualifications at age 16.

General Certificate of Education (GCE) A-level assessments are normally taken by 18 year olds in England, and results in these assessments are used by universities in student recruitment. Students typically take 4 subjects in the first year of their post-16 studies and sit the examinations for 4 AS subjects. They then continue with 3 of

these and sit the examinations for 3 A level subjects. This gives them 1 AS award and 3 A level awards at the point of admission to Higher Education.

National Curriculum - national (curriculum) assessment
The National Curriculum for England is a framework used by all maintained schools to ensure that teaching and learning is balanced and consistent.
It sets out:

- the subjects taught
- the knowledge, skills and understanding required in each subject
- standards or attainment targets in each subject - teachers can use these to measure each pupil's progress and plan the next steps in their learning
- how pupils' progress is assessed and reported

Within the framework of the National Curriculum, schools are free to plan and organise teaching and learning in the way that best meets the needs of their pupils. Many schools use the Qualifications and Curriculum Authority (QCA) Schemes of Work to plan their curriculum. These help to translate the National Curriculum objectives into teaching and learning activities.

The National Curriculum in England is divided into 4 'key stages' (ages 5-7; 7-11, 11-14, and 14-16). The critical mode of assessment is by 'end of key stage test' in maths, science and English. Typically taken at ages 7, 11, and 14. Key stage 4 learning typically is assessed through GCSEs (see above).

The constraints on delivering public goods – a response to Randy Bennett's 'What does it mean to be a nonprofit educational measurement organization in the 21st Century?'

In writing a response to Randy Bennett's illuminating analysis of the operation of ETS, I find myself suddenly subject to the methodological criticisms which I have recently made of others. In commenting on the forthcoming book from QCA (the Qualifications and Curriculum Authority) on comparability of public examinations, I stated (Oates T, 2007a) that whilst measurement specialists have a role in analyzing the historical, political, social and economic location of issues such as exam standards and comparability, it would be very helpful if more political scientists, historians and economists could overcome their apparent fear of the technical nature of assessment in order to make assessment an object of their own enquiries. In commenting on Bennett's paper, I am again an educationalist commenting on an educationalist's analysis - but it remains the case that analysis of the structural position and functioning of assessment agencies would benefit greatly from critique by those immersed in theory on the content and operation of national legislative frameworks, on conjunctions of market influences, on accountability arrangements, and so on. However, in the absence of any supply of this form of critique from those sources, my own analysis will offer an initial, albeit superficial overview. I make clear here the dimensions of a thoroughgoing analysis of the vital issues which Bennett raises, but in a short conference paper cannot do justice to the detail which the subject deserves.

My approach is as follows: I use the national operation of public examinations in England as a case study for the application and development of a framework of factors which enable adequate explanation of the 'space' which is created for the operation of independent assessment agencies. My contention is that the use of an extended framework of 'control factors' and wider cultural, political and economic contextual factors is essential for understanding the location of an assessment agency in a specific national setting, but also for enabling transnational comparison of the ways in which the 'space' for the operation of assessment agencies differs in different national settings. I argue that it is not enough to look at the mission of an assessment agency and the legislative arrangements which regulate this mission. I argue that public policy can frequently be driven by covert concepts which have not necessarily been exposed to public scrutiny, and that accountability structures are a

vital element of arrangements. I suggest that particular attention needs to be paid to informal mechanisms of accountability associated with the operations of assessment agencies. In focusing on English arrangements, I do not deal with the transnational operation of agencies such as Cambridge International Examinations (CIE), although using the framework of control factors to analyse transnational working would be a natural extension of the analysis I present in this paper.

The foreword to Bennett's paper presented in the ETS website version, and the abstract for Bennett's paper are couched in general terms: the idea of ETS as a case study of general trends, the importance of remembering the past, the significance of ETS as 'the largest of the non-profit educational measurement organisations'. However, I wish to suggest that we can only fully understand the precise location of ETS in its own national setting, and extrapolate from this, by use of the methodological tenets of mature transnational comparative method (Green A, 1990; Green A, 1997; Raffe D, 2005; Oates T, 2005). Outlining history is one thing, but analysing the full set of factors at work in another. This is a vital issue. Any generalization of the insights from the paper can only be made through the apprehension of the full set of conditions within the specific national setting which provide the context, drivers, sanctions, incentives and limits on the operation of the agency, and the state of political, economic, social and technical development at play in the specific national setting. This then allows other nations to examine (i) whether such factors are in operation in their own settings and (ii) the way in which these factors combine, in their own setting, to create the 'space' for particular modes of operation – if any – of independent non-profit making assessment agencies. Sophisticated transnational analysis is more than 'understanding things in context', it is concerned with apprehending the factors at play in deep and surface structures and using these factors to re-analyse other national systems – sometimes expressed as 'undertaking study of other nations in order to 'hold up a mirror' to one's own system'. Without this approach, the analysis is rather like the American baseball World Series – despite its name, essentially national rather than international.

Is this an unfair critique of the paper – after all, Bennett does not claim to provide transnational comparative analysis, so how can it be legitimate to use this body of theory to outline limitations of his approach? There are three reasons why I believe that these criticisms are legitimate. Firstly; whilst the paper provides a valuable historical record, and makes clear the relationship between the organizational mission, legislation and operation, it does not delineate the full range of factors

which constrain and legitimate the specific operation of ETS in its own national context. In other words, it falls short of a full explanation of the role and positioning of the organization in the US setting. Secondly, ETS operates in a variety of countries; the complex interactions in other national settings are not capable of being understood by virtue of the analysis supplied in the paper. Moral purpose becomes particularly relevant in countries where the purchasing of educational services has significant domestic economic implications, and the meshing of service provision and national public policy aims is a sensitive issue. Thirdly, the paper *does* present itself as an analysis of general interest and merit, and invites people to generalize from its analysis.

This legitimates the use of more extended frameworks for the analysis of the operation of independent assessment agencies. By way of contrast to the examination of the history and modus operandi of ETS, I will examine the role of such agencies – the examination boards – in the English setting, but through a framework of analysis which draws from Green's methodological work (Green A op cit). I will focus principally on general academic qualifications and assessments for compulsory schooling rather than vocational and occupational qualifications. The development and operation of the latter require even more extended frameworks of analysis due the complex interactions deriving from the labour market and global markets (Hayward G & Sturdy, 2005; Keep E & Mayhew K, 2001; Oates T 2005).

Green's framework of categories includes over twenty factors (including patterns of economic development; impact of economic development on family structures; labour market structures; demographic pressures; public policy imperatives; influencing roles of lobby groups; changing patterns of state function and control) as a means of adequately explaining the forms of education which have emerged in the US, Germany, France and the UK. The factor-based analysis approach of Green provided the basis for my own framework for analysis of the public policy associated with general and vocational education in England (Oates T, 2005). I will focus here on key control factors associated with compulsory schooling. This framework was originally developed to analyse the dimensions of public policy associated with the implementation of the National Curriculum in England, showing the differentials in focus, effort, and funding across different elements of the system (Oates T, op cit).

1. (national) curriculum content
2. assessment and qualifications

3. national framework for qualifications

4. inspection

5. pedagogy

6. professional development

7. institutional development

8. institutional forms and structures (eg size of schools, education phases)

9. allied social measures

10. funding

11. governance (autonomy versus direct control)

12. accountability arrangements

13. labour market/professional licensing

14. allied market regulation (eg health and safety legislation; insurance regulation)

15. deliberate conditioning of cultural climate (eg the importance of 'raising standards' in education and training)

Factor 13 is interesting since this overtly and heavily determines the shape of compulsory education in some countries, such as Germany, Austria and Switzerland. There, the high-participation vocational routes are determined by the nature of labour market regulation. It can be argued that the retention of GCSE qualifications in England (typically 10+ subjects taken at age 16) is in part explained by the extent of school leaving at 16, which is in turn the result of a continued labour market for 16 year olds (Hayward G, 2005).

Elsewhere, I have stated that '…it is the interaction of (these) factors which determines the precise character of regulation in each national setting and the relative weight of specific mechanisms of regulation…these might best be characterised as 'control factors' which public policy may or not choose to make use of…'. (Oates T, 2005). It is understanding the interaction of these factors which allows an understanding of how each national context (and legislative framework) allows a different 'space' and set of consents for the operation of independent assessment agencies. Again, using England as an example, in introducing the National Curriculum in the late 80's, the Government chose to emphasise the following:

factor 1:

curriculum content (in the form of attainment statements in each subject, at 10 levels of attainment)

factor 2:

national assessment (principally external tests in English, maths and science at ages 7, 11 and 14)

factor 12:

school performance tables using the data from the national tests

this was accompanied by changes in

factor 4:

geared up inspection arrangements with the introduction of the Office for Standards in Education

factor 11:

school governance, with a decrease in control of schools by Local Education Authorities

These were the key elements of the policy changes (Graham D & Tytler D, 1993). The National Curriculum was explicitly NOT directed at pedagogy – the organizations responsible were directed away from material which stated anything other than the outcomes of, and targets for, learning (Graham D & Tytler D op cit; Rawling E, 1999). Likewise, in the first years of the implementation of the National Curriculum, there were extended criticisms of the lack of financial resource which was directed at factors 5, 6 and 7 – pedagogy and professional and institutional development. However, it is the quality of the educational experience which is considered by many to be the crucial issue in raising standards (Osborn et al, 2000; Sylva K et al, 2004; Warrington M & Younger M, 2006) – an issue tacitly acknowledged in the 1998 and 1999 Literacy and Numeracy strategies. These were launched as a principal means of raising standards in maths and English, and marked a profound departure from the public policy position adopted for the National Curriculum – and thus represent a different policy use of the control factors listed above. The strategies engaged directly with the processes of learning as well as the aims and outcomes; they were welcomed by some as clear, helpful and based on well-grounded objective-based pedagogy, and vilified by others as an illegitimate extension of state control (ATL, 2006; Earl L et al, 2001; Select Committee, 2005).

This highlights the shifting emphasis in public policy across these key factors. It is relevant to the mission and structural positioning of independent assessment agencies operating within this national setting since it determines the 'space' for their

activities, the implied and actual role in respect of public policy, and the form of the assessment instruments and approaches which can be deployed. Crucially, the English setting has been characterized for the last three decades by 'assessment-led change'(Broadfoot P, 1979; Gipps C & Murphy P, 1994, McGaw B, 2006; Mansell W, 2007) where policy developers have explicitly exploited the 'washback' effect from assessment (Gipps C & Murphy P op cit; Hamp-Lyons L & Kroll B, 1997; Weiss C, 1998). This has been a pre-occupation of change strategy in England, with increasing direct control by Government regulators of the form and content of public examinations, and a cycle of constant change in the specific regulations and requirements on the form and content of public examinations. This has been felt acutely in areas such as vocational qualifications (Keep E, 2004; Oates T, 2007b; Oates T, 2007c), where policy reversals on the form and content of qualifications, continuous disruptive change, and superficial change without change in deeper structures has been endemic. It can be argued that successive policy-makers have changed qualifications since (i) they have a profound faith in the mechanistic nature of the washback from assessment, a faith which is not born out by the precise effects (Oates T, 2007b; Mansell W, 2007); and (ii) that it is easier to change qualifications than to undertake change in areas such as pedagogy, structural capacity (eg numbers of teachers in mathematics etc), etc. This has made public examinations not only the focus of Government policy, but also the focus of successive key structural enquiries into 14-19 education and training (Dearing R, 1996; Tomlinson M, 2004).

During this period, assessment agencies have been obliged to constantly follow pendulum swings in the design criteria for public qualifications, with all the attendant resource implications. A change in qualifications form and content not only affects the agencies supplying those qualifications, but also those organizations supplying text books, materials and curriculum and professional development relating to those qualifications. Whilst Government has become increasingly interested in the direct costs of the provision of qualifications by independent agencies (QCA 2007a) there is as yet no attempt to place a cost on the constant changes (in some cases, reversals) imposed on public qualifications.

Turning back to the theme of Bennett's paper – the importance of clarification of mission – my analysis suggests that it is important not only to define that mission with clarity, but to understand what factors in a national context *condition* an agency's ability to exercise that mission. National regulation can assume a form

which supports that mission (eg by sanctioning and funding specific qualifications) or can prevent its execution (eg by excluding certain forms of examination and tests from the national system). Bennett also focuses on charity law in respect of the functioning of ETS. Whilst this is clearly of importance, the legislative frameworks which need to be considered when examining the function and explaining the strategic decisions of independent assessment agencies are not only those which bear directly on their activities (charitable status; tax exemptions; etc) but also those which condition their products (qualifications criteria; national approval processes; etc) and those which condition the context in which they operate (funding of education and training; public policy aims regarding creation of a single European labour market; etc).

Using this general framework to analyse the English context suggests that it is vital to understand that:

1

uptake and regulation of qualifications occurs in very different ways in different country contexts

2

regulation operates through a complex mix of formal processes (law etc), non-formal processes (culture, expectations, legacy/traditions)

3

the peculiar country- specific mix of these formal and informal mechanisms is critical, and only by understanding the peculiar mix in each setting can the real operation of regulation be understood

4

education and training is regulated not only through legislation etc which is specific to the education and training arena. Incentives and drivers in the labour market (ranging from the state of the economy to labour market regulation) affect the operation and uptake of both general academic and vocational qualifications.

On 4; some countries link economic development strategy, labour market regulation and education and training policy (Singapore, Korea) through integrated policy

(Wong J 2006; Yong-Shik L 2006). England does not have a track-record of success in integrated policy of this type, despite the commitment to 'joined up Government' made in 1997 by the in-coming Labour administration in Westminster. The potential for management of incentives and drivers through public policy is thus not realised in the English setting, and policies frequently are in tension – eg tensions between Government policy on decreasing labour market regulation in order to decrease unemployment (by introducing more flexible labour rates) versus desires to increase the uptake of vocational qualifications (which would be delivered through greater regulation regarding licence to practice) (Keep E & Mayhew K 2001; Oates T, 2004; West J & Steedman H 2003).

Lodged firmly in Green's methodological paradigm, Raffe's 'home international' comparisons uses this kind of factor analysis to indicate the contrasts in system management:

England, Scotland and Wales have similar policy objectives for 14-19 education and training…(but) they are pursuing these objectives through contrasting strategies and measures. After a brief flirtation with a unified framework, England is pursuing a twin-track model, but with 'linkages' between the tracks. Wales is developing a unified framework of 16-19 pathways. Scotland introduced a unified system of academic and vocational post-16 learning in 1999. (Raffe D 2005 p1)

But this throws a further issue into sharp relief. The identification of control factors is insufficient to explain the reasons why particular blocks of public policy focus on specific control factors rather than others. In the wake of the first PISA results, Germany experienced a wave of public and political concerns regarding disparities in attainment which were linked to social class and, as a result, have undertaken development of federal tests in core areas of the school curriculum (Zadja J et al, 2005). Similarly, Australia has embarked on production of national tests which will be administered over and above the tests currently developed and operated by the individual states and territories (Tognolini J, 2006). This interplay of state-federal relations over testing in compulsory schooling has also been subject to interesting oscillation in the US context, with the operation of NAEP and long-term NAEP (IES 2007); the federal SCANS program (Dougherty J, 2001); and the continuing tensions over state versus federal responsibilities of 'No Child Left Behind' (US Dept of Education 2007).

However, this focuses on the 'what' of developments rather than the 'why'. I will move into this territory by examining 'key driving concepts'. Again I will exemplify this by looking at dominant constructs which are driving the particular constellation of control factors in the English setting. Considerable store has been placed, by the Qualifications and Curriculum Authority in England, in the reforming power of a national qualifications framework (Oates T, 1999; Oates T, 2000; Behringer F & Coles M, 2003). Behind the design and implementation of this framework resides a powerful commitment to a notion of 'coherence' – ie coherence in a national system of qualifications. This term repeatedly emerges in official presentations of the rationale for, and precise form of, the national qualifications framework in England (NQF), and the policy associated with ensuring that the framework is put in place (QCA 2007b).

This notion of 'coherence' has never been articulated in such a way that it could be subject to public scrutiny. It has been defined operationally by virtue of the strategies and policy measures put in place in its name. At the heart of the concept is a simple notion that there are too many qualifications in the English system, and that locating qualifications into a national framework of 8 levels will aid transparency and progression (QCA 2007c). These assumptions require scrutiny. Firstly, the view of an excessive number of qualifications derives from a figure of 20,000, produced by the Further Education and Funding Council, in 1996. This was a defective count, based on the total number of qualifications included on the Council's website, which included redundant qualifications, duplicate records, identical qualifications offered by different bodies, and previous versions of revised qualifications. The baseline figure for accredited qualifications in the England, Northern Ireland and Wales is provided by the National Database of Accredited Qualifications (NDAQ) (QCA 2007d) and this yielded a figure of 5,850 at May 2007. Germany has a very different general education structure in comparison with England, but the degree of within-nation variation is comparable – it cannot be claimed that one is startlingly more coherent than the other. In respect of vocational qualifications, there exist approximately 4800 vocational and other qualifications in Germany: 150 state-specific qualifications, 420 in the Dual System of apprenticeship, 3,000 chamber of commerce qualifications and 500 continuing education qualifications (Source: BIBB). There is thus no basis for using the figure of 20,000 as the benchmark for England, nor for using this as a rationale for a claim that the system is incoherent.

A further articulation of 'coherence' which has been present in the policy imperatives of QCA relates to progression – that placing all qualifications in a common framework increases the coherence of progression routes. But there is little evidence that users of qualifications in different occupational areas of the labour market are interested in identical sets of levels etc. Employers are interested in relatively narrow pools of qualified labour within occupations, and when moving outside this in times of skills shortages, turn to related areas, not to broad recruitment from any sector. To suggest that system-wide coherence is a desirable structural characteristic runs counter to the sector-specific trends in economic development (Bjornavold J, 2004) and to the activity of skills ecosystems (Payne J, 2007). Such analyses suggest that system-wide imposed (restrictive) frameworks are more an expression of policy-makers' sense of 'system-tidiness' than a concern to genuinely enable better economic, personal and social utilization of skills, knowledge and understanding (Oates T & Coles M, 2004). Yet the concept of 'coherence' has been used to justify continued increase in the purchase and application of national regulatory criteria in England. It is intriguing to contrast the 'coherence' argument with the objectives of one of QCA's predecessor bodies, the Schools Council, which was principally interested in innovation and responsiveness. Even in the context of a greater number of assessment agencies/examining boards (then 8 GCE boards and 14 GCSE boards rather than the current 3 GCE-GCSE boards) the council stated its purpose as being '…to explore ways of improving the traditional Advanced level type of sixth form course and of adding to it new courses, both examined and non-examined, which will appreciably increase the ability of schools to offer curricula better matched to individual and general needs…' (Schools Council 1965 p9)

Notably, using the NQF, increasingly tight regulatory criteria and the underpinning notion of 'coherence' has not decreased the number of qualifications taken per pupil in statutory education, which has seen a steady increase since the 1980's (Mansell W, op cit) – including an increase in non-statutory diagnostic and benchmarking tests prior to the age of 16. The state's growing concern regarding *costs* of public examinations now accompanies the concern for 'coherence'.

It can be argued that – for the relatively stable market of public examinations in schools – the stability of arrangements in the 1970's and 1980's gave rise to a desirably balanced set of incentives and drivers:

None of the eight GCE examination boards…receives Government funds, although some do draw monies for non-examination purposes. All eight boards rely on the fees that they receive for the examinations that they conduct. Like any other large business they cannot afford to make a loss and competition between them renders it extremely difficult to make a satisfactory surplus. All eight boards are thus in a state of financial and academic competition. None can afford to set their fees too high. Similarly, none can afford to have their syllabuses viewed as out-of-date, uninteresting, or of excessively high standard; if they do then they will fail to attract candidates. If their syllabuses are viewed as too easy or superficial, then they risk not being recognised by some universities for entrance purposes and are thus devalued in a different way. It is this combination of the two forms of pressure upon the GCE boards that has led to the efficiency of the current system as a whole and the wide choice which it now offers to teachers.

While the boards compete with each other for the available candidates, they also co-operate on academic matters. The GCE secretaries – the senior administrative officers of the boards – meet regularly to discuss matters of mutual interest and to formulate common policies. Meetings of research, computing and administrative staff are also organised on a regular basis. These exchanges constitute a major strength of the GCE system.

(Kingdon M, 1991 p8ff)

(My note: in referring to 'loss', Kingdon is referring to the overall financial viability of an agency rather than the viability of individual examinations. Loss-making examinations (eg in minority subjects) continue to be offered by agencies on the basis that there is an educational need which transcends commercial interests)

Kingdon thus highlights the importance of looking at factors other than state regulation. Whilst my analysis presented thus far focuses of the management of key control factors and gives some insight into the key driving concepts, it perpetuates the notion that the mission and accountabilities of independent assessment agencies should be analysed principally in terms of their relationship with state apparatus. But Kingdon reminds us that very different relationships obtained in the past, and it is vital not to overlook this.

During the 1970's and 1980's, substantial curriculum innovation was undertaken by linked groups of schools and leading curriculum developers working with schools (Salter's Science Project; Ridgeway History; French graded criteria project; Suffolk Science; Wessex Science; Nuffield Science). These projects originated in school practice (Stenhouse L, 1975) and the developers sought development relations with independent assessment bodies in order to provide assessment and certification arrangements. This partnership was thus focused on an assessment agency-school linkage rather than an assessment agency-state relationship. It is vital to note two things. Firstly, from these projects emerged many of the developments in

qualifications to which the state now places at the heart of qualifications criteria (modular qualifications; transparent assessment criteria; objective-based learning). Secondly, the accountabilities felt by the assessment agencies were felt in relation to schools rather than state agencies.

This introduces the final theme of this paper: accountability. Whilst it is vital to ensure that the explicit mission of an assessment agency reflects clear public goods, it is equally important to examine the way in which the organisation is held to account for executing this mission. I argue here that independent assessment agencies in England have a history of highly integrated accountability arrangements, and that this is a considerable advantage provided by their mode of operation. In addition, this highly integrated accountability tends to be ignored by those outside the agencies; it is pushed into the background by the highly prominent exchange between the state and the agencies over the formal regulation processes. These processes imply that the principal accountability exchange is between the state and the agencies, whereas the aims of the agencies place greater emphasis on the exchange between the agencies and the communities which they serve – principally education providers and the learners within them.

'Accountability' should not be exclusively associated with formal mechanisms - such as consultative forums (committees, groups etc) which are underpinned by statute or constitution. Nor should it been seen exclusively as formal 'accountability systems' such as the provision of school performance tables (see Mansell W opcit p200ff). More contemporary theorists define accountability as 'a social relation in which an actor feels an obligation to explain or justify his or her conduct to some significant other' (Lerner J & Tetlock P, 1999; Strom K, 2000; McCandless H, 2001; Pollit C, 2003; Bovens et al, 2006). This literature emphasizes the distinction between 'accountability', 'responsiveness' and 'transparency' (Mulgan R 2003; Bovens M 2005); with a crucial distinction being that accountability requires '…a justification of conduct…' (Bovens M et al 2006 p5).

There are strongly competing models of public accountability in democratic settings. Some models emphasise the importance of 'top down' specification of what assessment and qualifications should look like. That is, education forms part of the manifesto statements of political parties and Governments are elected on the basis of their manifesto statements. Assessment and qualifications are a key part of education and training provision, and it is therefore only right that any assessment

agency should accord with the policies of democratically-elected national government. But this simple picture of the structure of accountabilities is rendered far more complex by the following:

1

In countries with a state-federal system, there are persistent tensions between state and federal governments in respect of which arrangements best meet public needs, strongly associated with sensitivities over the rights of federal government to intervene in the detail of state arrangements. By contrast, in England, a national school certificate system was created in 1917, following the 1902 education act, although the school curriculum only became strongly centrally determined after the implementation of the Education Reform Act of 1987. This took control of curriculum and its specific assessment 5-14 away from Local Education Authorities, placed curriculum matters initially in the hands of the National Curriculum Council and placed new 5-14 national school assessment in the hands of the Secondary Examinations and Assessment Authority - roles now held by the Qualifications and Curriculum Authority (QCA). In the event, independent assessment organisations continued to provide public examinations at 16 and 18, although under more heavily-regulated arrangements. But this impression of central national government as the final arbiter of assessment policy for compulsory schooling and in respect of publicy-funded qualifications is tempered by the continuing impact of the legal requirements placed on all members of the European Union, with an attendant raft of public policy instruments such as the European Qualifications Framework.

2

Legal theory suggests that legislation both *regulates* social, political and economic processes but also *responds* to social, political and economic developments (Holdsworth W 1972; O'Brian 1997). There may be a considerable time lag in the legislature responding to such developments, and different forms of pressure need to be applied to effect development and change in legislation. Assessment agencies are close to realities in education and training and aim to innovate in assessment technology as well as respond to emerging needs – as they did particularly intensely during the 70's and 80's – as outlined above. Their role can be construed, therefore, as not merely complying with existing legislative requirements, but also contributing to – indeed *leading*, in certain circumstances – the process of effecting revision and development in learning and this then feeding into revised legislation relating to assessment and qualifications.

3

There are strongly competing models of accountability (Pollit C op cit), the more recent emerging primarily from the work of non-government organisations in respect of action in developing countries (Fumo C et al, 2000). In countries such as the UK, some of the radical models have intriguingly been imposed, by legislation or requirement, on public service providers as a means of increasing the responsiveness of their services (DHS 2005). Such models are highly sceptical of the extent to which certain forms of central government can, with sensitivity, represent and respond to community need, and that communities should themselves take action to ensure that their needs are both recognised and responded to.

Interest in accountability arrangements is strong in the public policy arena in the European Union, stimulated by the Commission's concerns to carefully balance action at a system level with empowerment and recognition of community groups within the Union. This key area of accountabilities is neglected in many analyses of the operational arrangements around organisations involved with assessment and qualifications (eg Bill L, 1991; Gipps C 1994), and is fraught with multiple interpretations of what constitutes adequate and appropriate accountability (Day & Klein 1987; Hirst P 1993). In this paper I wish to distinguish (i) formal mechanisms from (ii) informal accountability mechanisms. I argue that the latter are frequently overlooked in external analyses of the operation of assessment agencies, yet they can provide vital linkages between assessment organisations and the constituencies which they serve.

'formal mechanisms' include: governing councils and the documents which regulate them; advisory committees; policies and requirements regarding consultation; statements of mission and of standards relating to operation

'informal mechanisms' include the practices of staff in gathering evidence regarding that which will best meet the needs of communities; the involvement of outside professionals in key operations; the networks which individuals and the organisation maintain

Only by consideration of elements which fall into both categories can the precise nature and extent of the accountabilities of any assessment organisation be fully appraised. It is vital to understand the precise nature of informal mechanisms which are in a play in a given organisation at a specific time; otherwise it is easy to underestimate the extent to which the organisation is linked into the communities which it serves. The downside of informal mechanisms is that whilst they may be entrenched features of the organisation, they may also be subject to change which is not the outcome of specific deliberation.

Proximity to the community which is served is thus a vital feature of high accountability systems (Politt C op cit), with high levels of accountability coming from systems in which the community makes a direct input to decision-making. This is a very important feature of the operation of school examinations in England. Whilst the QCA has dismantled many of its subject committees and implemented severe cuts in the subject teams which it employs for development purposes – both being principal means of reaching into the school community – the awarding bodies in England continue to work with a large community of examiners, which are an intrinsic part of the design and operation of qualifications. To illustrate the scale of this, the OCR board, one of the three main awarding bodies and part of Cambridge Assessment, relies on a population of 16,000 examiners and principal examiners in the design, evaluation, and operation of its public examinations. This provides a large scale intrinsic, informal accountability function, which exists in addition to specific design programmes for entirely new qualifications and specific consultations on key issues. This has led to the assessment agencies being highly sensitive to the needs emerging from specific learner groups and subject communities. The importance of this informal exchange, a crucial and intrinsic element of the history of qualifications and the agencies, and built into the heart of the operation of the qualifications, cannot be overstated (Raban S ed, in press).

But just as the market relations which Kingdon emphasizes as being previously in balance have been disturbed by an extension of regulation and a cycle of continued – and in many cases contradictory – revisions (Kingdon M, op cit), there are emerging some tendencies which may compromise these vital informal accountability elements:

tendency #1 – increased scope of state regulation

While for the majority of public examinations there remains a high level of involvement, of educational professionals from the schools community, in the marking and awarding process, there is increasingly tight regulation of the form, content and operation of qualifications by central government, through qualification approval, qualification criteria and codes of practice. These have increasingly operated not only at a general level in terms of ensuring comparability between qualifications, but have increasingly imposed upon highly specific design features of qualifications – eg in 2000, specifying that all advanced level general qualifications would consist of six units (modules) (QCA 2001), changing this to four units (QCA 2007e) for implementation from 2008.

tendency #2 – pressures for technological transformation of processes
The technologically-based revision of assessment has a tendency to diminish the proximity and amount of contact between the assessment agencies and the educational community. Substituting machine marking for human marking, deriving tests from pretested item-banks, and related innovations which reduce or remove the role of subject professionals in marking and awarding (determining cut scores for grades) can change substantially the nature and quantity of contact. While professionals remain involved in the creation of item banks, this is frequently a very specific groups of educationalists – it is only a small subset of the substantial, broad population of professionals who are involved in marking and awarding national general qualifications. In addition, such contact between the assessment agency and the item authors is very 'front loaded' – once an item is admitted to the bank and validation data builds, there is no need for continuing contact between the professionals and the assessment agency, in contrast to the sustained contact with markers and principal examiners which is a feature of 'conventional' public examinations.

In conclusion, in order to understand with precision the structural positioning of independent assessment agencies and the extent to which they are empowered to discharge clear public goods, it is vital to attend not only to the mission of organizations and the legal basis for that mission, but to the specific 'space' for assessment activity created by public policy strategy in each country in which the organizations operate. This paper suggests that the structural location of assessment agencies in the English setting is highly specific to the particular historical conjunction of control factors in that setting but that using these control factors as a framework for analysis can help us to understand how other systems -

such as Singapore - place independent assessment agencies in a different structural position. For example, there, Government deliberately maintains closer accountability-based partnership and liaison links with schools, using third party assessment agencies as highly specialist support for achieving the aims of negotiated national strategy. Thus, what I suggest is that it is vital to comprehend the importance of the direct relationship between agencies and the communities they serve, and not underestimate the importance of these in the face of the state-assessment agency relations existing in the country setting in which the agency is located. In other words, the 'noisy' overt political relations around the state can obscure 'quiet', less apparent developmental and operational relations, which nonetheless represent substantial accountability linkages. It is also vital to recognize that the technological transformation of assessment presents important shifts in informal relationships, and where these shifts represent a dilution of accountability, agencies would be wise to attend to alternative arrangements which ensure continuation of the existing high levels of accountability.

References

ATL 2006 *The Rose Review – ATL response* Association of Teachers and Lecturers

Broadfoot P 1979 *Assessment, schools and society* Methuen

Behringer F & Coles M 2003 *The role of national qualifications systems in promoting lifelong learning* OECD

Bjornavold J 2004 *Discussion note on new and emerging models in vocational education and training and higher education* European Commission

Bovens M 2005 *Public accountability. A framework for the analysis and assessment of accountability arrangements in the public domain.* Connex Belfast conference.

Bovens M et al, 2006 *Multilevel governance and public accountability in Europe: which institutions, which practices, which deficit?* Utrecht University

Coles M & Oates T 2004 *European reference levels for education and training*

CEDEFOP Day & Klein 1987 *Accountabilities: five public services* Tavistock

Dearing R 1996 *The Dearing Review of 16-19 qualifications* HMSO

DS 2005 *Delivering better public services – progress* Department of Health

Dougherty J 2001 *Can Elizabeth Dole fill Helms' shoes* Worldnet Daily Nov 2 2001

Earl L et al 2001 *Watching and learning 2* Ontario Institute for Studies in Education

Fumo C et al 2000 *Social Fund: an effective instrument to support local action for poverty reduction?* Department for International Development

Gipps C & Murphy P *A fair test – assessment achievement and equity* Open University Press

Graham D & Tytler D 1993 *A lesson for us all* Routledge

Green A 1990 *Education and state formation* Macmillan

Green A 1997 *Education, globalisation and the nation state* Macmillan

Hamp-Lyons L & Kroll B 1997 *TOEFL 2000 – Writing composition community and assessment* Educational Testing Service

Hayward G 2005 *Where do all the learners go? Efficiency and equity in the provision of 16-19 learning opportunities* SKOPE issues paper #8 Oxford University Department of Educational Studies

Hayward G & Sturdy, 2005 *Turning up for training but who pays the price?* SKOPE issues paper #7 Oxford University Department of Educational Studies

Hirst P 1993 *Associative democracy* Polity Press

Holdsworth W 1972 *A history of English law* Methuen

IES 2007 *NAEP Overview* National Center for Educational Statistics

Keep E & Mayhew K *2001Globalisation, models of competitive advantage and skills SKOPE Research Paper No. 22,* SKOPE, University of Warwick

Keep E 2004 *The multiple dimensions of performance* working paper 23 Nuffield review of 14-19 education and training SKOPE issues paper #22 Universities of Warwick and Oxford

Kingdon M 1991 *The reform of advanced level* Hodder & Stoughton

Lerner J & Tetlock P 1999 *Accounting for the effects of accountability* Psychological Bulletin 125 2 255-275

McCandless H, 2001 *A citizen's guide to accountability* Trafford

McGaw B 2006 *Assessment fit for purpose* Paper to International Association for Educational Assessment Singapore

Mansell W 2007 *Education by numbers: the tyranny of testing* Politico's

Mulgan R 2003 *Holding power to account: accountability in modern democracies* Pelgrave

O'Brian D 1997 *Constitutional law and politics vol2 civil rights and civil liberties* WW Norton

Oates T 1999 *Framework for comparison and evaluation of national classification systems of vocational standards – the English case* Qualifications and Curriculum Authority

Oates T 2000 *Progress on the implementation of the national qualifications framework (NQF)* Qualifications and Curriculum Authority

Oates T 2004 *The role of outcomes-based qualifications in the development of an effective vocational education and training system: the case of England and Wales* Policy Futures in Education vol2 #1

Oates T 2005 *Understanding how regulation works in different national settings* Qualifications and Curriculum Authority

Oates T 2007a in press *Commentary* in Newton P, Baird J, Goldstein H, Patrick H and Oates T 2007b in press *Protecting the innocent – the need for ethical frameworks within mass educational innovation* in Saunders, L. ed 2007 *Educational research and policy-making: exploring the border country between research and policy*, Routledge

Oates T 2007c in press *Going round in circles: temporal discontinuity as a gross impediment to effective innovation in education and training* in *Knowledge transformation and impact* special edition of the Cambridge Journal of Education

Ofsted 2003 *The national literacy and numeracy strategies and the primary curriculum* Office for Standards in Education

Payne J 2007 *Skills in context: what can the UK learn from Australia's skill ecosystem projects?* Skope research paper #70 SKOPE Cardiff University

Pollit C 2003 *The essential public manager* Open University Press / Mc Graw Hill QCA 2001 *Curriculum 2000 review phase 1 report* Qualifications and Curriculum Authority

QCA 2007a *Minutes of the fifty fourth meeting of the board of the qualifications and curriculum authority* Qualifications and Curriculum Authority
QCA 2007b *Withdrawing a qualifications* Qualifications and Curriculum Authority

QCA 2007c *The National Qualifications Framework – factsheet* Qualifications and Curriculum Authority.

QCA 2007d *http://www.accreditedqualifications.org.uk* Qualifications and Curriculum Authority.

QCA 2007e http://www.Qca.org.uk/qca_4076.aspx Gce AS/A levels: overview Qualifications and Curriculum Authority

Raban S ed in press *Examining the world* Cambridge University Press

Raffe D 2005 *Learning from 'home international' comparisons: 14-19 curriculum and qualifications reform in England, Scotland and Wales* Centre for Educational Sociology University of Edinburgh Select Committee for Education and Skills 2005 *Education and skills – eighth report* session 2004-05 House of Commons

Rawling E 1999 *Geography in England 1988-89 – costs and benefits of national curriculum change* International research in geography and environmental education 8 (3)

Schools Council 1965 *Change and response – the first year's work oct 1964-sept 1965* Her Majesty's Stationery Office

Stenhouse L 1975 *An introduction to curriculum research and development* Heinemann

Strom K 2000 *Delegation and accountability in parliamentary democracies* European Journal of Political Research 37

Sylva K et al 2004 *Technical paper 12 – the final report* DfES/IoE London

Tognolini J 2005 Meeting the challenge of assessing in a standards based education system Curriculum Council

Tomlinson M 2004 14-19 *Curriculum and qualifications reform – final report of the working group on 14-19 reform* Department for Education and Skills

Tymms P 2007 *Techniques for monitoring the comparability of examination standards.* Osborn M et al 2000 *What teachers do – changing policy and practice in primary education* Continuum

US Dept of Education 2007 *Statement by US Secretary of Education Margaret Spellings on the No Child Left Behind Act of 2007* US Dept of Education

Warrington M & Younger M 2006 *Raising boys' achievement in primary schools: an holistic approach* Open University Press

Weiss C 1998 *Evaluation* Prentice Hall

West J & Steeman H 2003 *Finding our way: vocational education in England* Centre for Economic Performance

Wong J 2006 *Varieties of capitalism in Asia: health industry development in Korea, Singapore and Taiwan* Paper presented at the annual meeting of the American Political Science Association Pennsylvania Convention Center, Philadelphia PA Aug 31 2006 retrieved from http://www.allacademic .com/meta/p151124_index.html

Yong-Shik L 2006 *Reclaiming Development in the World Trading System* Journal of World Trade

Zadja J et al 2005 *PISA in Germany: a search for causes and evolving answers* Springer