

The regulation of examinations and qualifications:

An international study



Author: Andrew Boyle
Principal research officer, e-assessment

Date: April 2008

Ofqual/08/3736

Acknowledgements

I have been greatly assisted by several groups of people in this project – so much so that I wanted to put the acknowledgements section at the front of the document, rather than hidden away at the back.

So, many thanks to the following groups of people:

- the INCA and Eurydice country contacts for providing such rich and useful information upon which to base the findings of the research
- the staff at the International Information Unit (IIU) at the National Foundation for Educational Research (NFER) for helping with all aspects of the project
- research colleagues at QCA (and now Ofqual) for vetting drafts of this report

Of course, any errors in this document are my responsibility alone. The views expressed in this report do not necessarily represent those of the QCA, Ofqual, NFER or any other organisation.

Andrew Boyle, April 2008

Contents

1	Executive summary.....	1
2	Introduction	6
2.1	The English examinations and qualifications system	6
2.2	An international research project.....	7
2.3	A hypothesis-generating project.....	9
2.4	Scope of reported findings	10
2.5	Research questions.....	11
2.6	Structure of this report.....	11
3	Methodology	12
3.1	INCA and Eurydice.....	12
3.2	Data collection procedures.....	12
3.3	Analytical techniques.....	14
4	Data	15
4.1	Time of data collection	15
4.2	Amount of data collected.....	15
4.3	Quality of data collected	15
5	Background issues.....	17
5.1	Some definitions of regulation	17
5.2	The range of factors that affect regulation.....	18
5.3	Histories of examination arrangements.....	18
5.4	Histories of regulation.....	19
5.5	International understandings of regulation	20
5.6	The quality aspect of regulation	21
5.7	Regulation in the future	21
5.8	Summary of this section	23
6	Findings	24
6.1	Types of qualifications.....	24
6.2	Facets of the examinations and qualifications process	25
6.3	Vesting of exams and qualifications powers.....	37
6.4	Aims for regulation	44
6.5	Use and regulation of e-assessment.....	51
7	Implications of the findings.....	57
7.1	Purpose of this section	57
7.2	Framework for later phases.....	57
7.3	Low-hanging fruit.....	61
8	Appendices	63
8.1	Project questionnaire.....	63
8.2	Bibliography.....	68

List of tables

Table 1: Interpretation of OECD data showing assessment methods for level 3 general qualifications	28
Table 2: Categorisation of jurisdictions as using 'mainly internal' or 'mainly external' assessment	30
Table 3: Properties of assessment designs and score data that jurisdictions expressed as aims for regulation	47
Table 4: System-level aims for the regulation of exams and qualifications.....	48
Table 5: Jurisdictions' use of e-assessment.....	52
Table 6: Issues affecting the future of regulation	61

List of figures

Figure 1: Possible steps in the convergence process.....	22
Figure 2: Certified assessment at the end of general upper secondary education, 2002/03	29
Figure 3: Continuum showing centralisation of techniques to assure public confidence in the examinations process	35

1 Executive summary

This is a report of an international research project in which the data were collected in the autumn of 2007. The research was designed to be hypothesis generating; to give an initial understanding of how international findings could provide insight into English qualifications and examinations regulation, and inform future work.

The research questions were:

1. What insights can international findings provide to either support or challenge English approaches to the regulation of examinations and qualifications?
2. To what extent are English approaches to the regulation of qualifications and examinations mirrored by practices in other international jurisdictions?
 2. a. When international approaches to regulation differ from those of England, what insights can the English regulator take from international practice?

The project gathered data via the INCA and Eurydice networks, and the two main data collection procedures were the completion of a constructed response questionnaire and the collation and validation of country-specific summaries derived from network archives. These collection procedures were supported by a literature review.

Questionnaires were completed and returned by 15 jurisdictions, supplementing 46 summaries from the network archives. The data were judged to be sufficient in quantity and quality to sustain the findings reported.

Before expressing the new findings from this research, relevant background issues were described. These are summarised below:

- 'Regulation' at its broadest means actions taken on behalf of governments in the public interest to steer events and behaviour, rather than to provide or distribute goods or services.
- Regulation increased in significance as an area of enquiry following the privatisation of public utilities during the 1980s. This applies to the United Kingdom and to other countries in various parts of the world (such as Europe and Australia).
- The provision and regulation of qualifications in England has a distinct history, which is different both from that of qualifications in other jurisdictions and from the post-1980s experience in other UK public utilities.
- Understanding regulation requires an understanding of quality in educational assessment. Quality as understood in the English regulatory context includes: validity, reliability, fairness, manageability and comparability.

- Regulation may change in the near future. Such changes arise from at least two directions: new political arrangements (for example the European Union – EU) and the rise of technology.

The research findings were divided into five areas: types of qualifications, facets of the examinations and qualifications process, vesting of 'governmental' powers for exams and qualifications with different bodies, aims for exams and qualifications regulation and use and regulation of e-assessment. Each area of findings is summarised below:

Types of qualifications

- This research confirmed earlier research in finding that most jurisdictions in the sample issued a single school leaving diploma or certificate to students on leaving post-compulsory education.
- Although not unique, England's practice of awarding separate qualifications for individual subjects within a qualifications framework was less common.

Facets of the examinations and qualifications process

- All jurisdictions ensure public confidence in the exams and qualifications process.
- Jurisdictions' techniques for defending public interest typically contain a mixture of centralising and de-centralising features.
- Centralising and de-centralising features can occur at different points in examinations and qualifications processes (for example defining qualifications standards, assessment, awarding certificates).
- Jurisdictions can change their orientations from centralised to de-centralised (or vice versa) over time.
- It is possible to construct and justify a continuum showing centralisation and de-centralisation as used to defend public interest in examinations and qualifications.
- England would be near the centralised end of such a continuum, but not absolutely at the extreme.

Vesting of powers relating to exams and qualifications

- Definitions of regulation as 'governments steering events' make a statement about the types of organisations that provide public services; that is, that there are multiple organisations involved and that providers would typically be private-sector bodies.
- This part of the research findings looked at ways in which (parts of) responsibilities and powers relating to examinations and qualifications were vested with government and/or other organisations.

- It was found that various of the responsibilities and powers were vested with schools and regional authorities in some jurisdictions in this study.
- It was also found that seven jurisdictions in the study had set up national bodies performing technical tasks concerning exams and qualifications which were independent from (or 'at arm's length from') political control.
- Such independent bodies had substantial provision and distribution functions and so were not regulators, but each of them carried out one or more tasks that would be carried out by the regulator in England.
- In some ways, setting up an independent body to carry out exams tasks is a de-centralising measure. However, five of the seven jurisdictions with such arm's-length organisations also mainly used external assessment – a centralising tendency.

Aims for exams and qualifications regulation

- The issues in this part of the findings were the most complex to be addressed in the research. As such, the interpretation of this set of findings should be limited to the pointers they provide for further research.
- The sub-section related to the aims that jurisdictions had for exams and qualifications *regulation*, not to their aims for examinations and qualifications systems *per se*.
- The notion of aims is complex, and is related to other concepts such as 'values' and 'purposes'.
- Aims were expressed at two levels of generality: the properties of assessment designs and the system-level aims for exams and qualifications.
- Using findings regarding those levels of generality it was possible to do comparative and contrastive analyses of jurisdictions' aims for their systems.
- Such analyses suggested that understanding how jurisdictions balance standardising and localising measures is a useful way to understand the regulation of qualifications and examinations in an international context. However, comparative analyses challenged the view from the literature that it was possible to associate aims with particular assessment designs incontrovertibly.

Use and regulation of e-assessment

- For the majority of jurisdictions responding to this questionnaire, there was either no use of e-assessment (eight jurisdictions) or initial use (three jurisdictions) for high-stakes qualifications or exams taken at the end of post-compulsory schooling.
- The two jurisdictions in this sample with the most use were *Queensland* and *the Netherlands*. Their use of technology for assessment reflected their overall approaches – *Queensland* used technology to support teacher assessment, whereas *the Netherlands* had various approaches to running e-tests.

- There are many reports in the literature of e-assessment programmes. But typically, these are of 'low-stakes' assessment applications and/or pilot projects. The exception to this general picture is the USA, which has several large-scale high-stakes e-assessment programmes.
- Two jurisdictions replied to the questionnaire that they had regulations pertaining to e-assessment. *Queensland* said that they had regulations in respect of information privacy, whilst *Latvia* responded by describing its plan to develop a bank of state examination tasks by 2013.
- There were several sets of quasi-regulatory standards or guidelines documents in the English-speaking world. Such documents varied along the following dimensions: scope, audience, enforceability, wording, granularity of detail, relationship with other sets of standards and vision of education quality.
- In addition to such quasi-regulatory documents there were several sets of guidance documents for centres providing e-assessment to candidates.

The report concludes by suggesting implications from the findings. Firstly, it suggests a framework for future research work on this topic. That framework consists of the following parts:

- defining exams and qualifications regulation as the ways in which and the extent to which jurisdictions ensure that the public have confidence in their examinations systems
- taking the notion of (de)-centralisation as a key one for understanding regulation, but doing further work to define different meanings and applications of de-centralisation
- doing further work to investigate arm's-length technical agencies
- following up to find out international respondents' views on the future of regulation

The second part of the implications section suggests six key insights that the English regulator can take away from this research **now** – without waiting for the follow-up work:

- Regulation means something in the international context if it is about how jurisdictions help the public to maintain confidence in their exams and qualifications systems. If it is merely about the respective and mutual duties of a quasi-public sector technical body (a regulator), and private sector providers (awarding bodies), then it does not.
- Centralisation and de-centralisation can both help to ensure public confidence in exams and qualifications systems. In England, qualifications regulation has often been synonymous with centralising measures.
- Centralisation versus de-centralisation could be seen – stereotypically – as a bald choice between external examinations and teacher assessment. In fact, there are many ways in which de-centralising measures can be incorporated into qualifications processes – they

can be incorporated at many stages of the process: defining standards, assessment itself, awarding qualifications, and so on.

- The international jurisdictions in this sample did not have regulators as such, but several of them did have technical bodies set up at arm's length from government to perform examinations tasks.
- An analysis of examinations purposes, following the work of Paul Newton, should form part of what regulation is about in the international context.
- Challenges to regulatory practice in future will come both from the impact of new technology **and** the fall-out of globalisation in general and specifically from the existence of supra-national political entities (for example the EU) and practices (for example international comparative surveys and credit transfer).

2 Introduction

2.1 The English examinations and qualifications system

The English examinations and qualifications system has specific characteristics. Typically (at the time of this study), 16-year-old students are studying around 10 subjects – most often they are attempting to gain General Certificate in Secondary Education (GCSE) qualifications. GCSE qualifications are – by default – taken at the end of year 11 of secondary education, the end of compulsory education. In post-compulsory education many students take General Certificate of Education (GCE) Advanced Subsidiary (AS) and Advanced (A) level qualifications. In most cases students complete their A levels after two years of post-compulsory study.

There is a strand of vocational qualifications available to students in England. These reside within the National Qualifications Framework (as indeed do GCSEs and GCEs – general qualifications). As such, vocational qualifications can serve a range of purposes in different sectors and at different levels, varying in size, level and method of assessment.

There are increasing numbers and types of qualifications available to young people in England. For example, there are several 'portmanteau' or 'umbrella' qualifications, which are constituted from various elements including some other freestanding qualifications. Most prominent amongst these portmanteau qualifications is the new suite of Diplomas. As well as being made up of several elements, Diplomas seek to 'bridge the gap' between general and vocational qualifications; being highly applied, relevant to specific employment sectors and yet retaining academic rigour (DCSF, 2007a).

Qualifications are owned, developed, administered, marked and awarded (or 'certified') by 'awarding bodies'. Awarding bodies are independent organisations, with varying statuses – some are charities, others are linked to professional bodies. Yet others are commercial companies.

Qualifications taken in schools and colleges are regulated by the Qualifications and Curriculum Authority (QCA). Awarding bodies apply to QCA to have qualifications accredited, and then the qualifications are monitored whilst they are in use. QCA may also recommend that a qualification be approved for use by students of particular ages (for instance those under 18 or under 16). Then, schools and students can receive public funds to take the qualification. (See Nisbet & Greig (2007) for more detailed description of the functions of awarding bodies and the QCA in England.)

QCA is a statutory body, sponsored by the Department for Children, Schools and Families (DCSF), with powers, duties and responsibilities under the Education Act 1997. The duties and responsibilities of the QCA are many and broad. Parts of the Authority develop the national curriculum and new qualifications for life and work, and QCA also subsumes a large executive agency (the National Assessment Agency) that delivers national curriculum tests and has a large modernisation and delivery role for qualifications used in schools and colleges.

The regulatory duties of the QCA have been largely implemented by the QCA's Regulation and Standards division. During the period in which this research has been conducted, the UK government announced its intention to 'float off' the regulator, and to reconstitute it with independence from the DCSF (see DCSF, 2007b). The new regulator known as the Office of the Qualifications and Examinations Regulator, or Ofqual (DCSF, 2007c) has been constituted on an interim basis during the time that this research was being conducted.

Much about the new regulatory body is still being finalised at the time of writing, and significant parts of this research (especially the writing and distribution of the central instrument, a questionnaire) were conducted before the announcement of the split.

2.2 An international research project

QCA has conducted a range of research into the regulation of qualifications and assessment (see Nisbet & Greig, 2007). Most of this work has been based on information gathered from sources that analyse the situation in England. Given that limitation of work done so far, it was decided to investigate international practice in the regulation of examinations and qualifications.

There is a potential objection to running an international project to investigate this issue that must be immediately addressed. That is, the researcher strongly believes, before starting the project, that the English set up of awarding bodies providing services under the supervision of an independent regulator is pretty uncommon in other countries, if not unique (cf. Nisbet & Greig, 2007, p. 50). This scenario is unlike others that might pertain in an international research project – for example, in a study of early years education such as that of Bertram & Pascal (2002), the researchers would be more or less certain on starting the work that all the countries to be studied *did* provide some form of early years education. A slightly different scenario would be if the researchers started out their enquiry not knowing whether a phenomenon existed or not overseas; the point of the research being to find out whether it did and to gain insight relevant to the English context from the (dis)-similarity of international experience. An example of such an enquiry would be research into dips in students' performance in the lower secondary phase of education (Whitby et al, 2006).

So, an international study of regulation, when the researcher starts from the understanding that most other countries do not regulate as England does may be a dubious notion. However, the proposition can in fact be defended. Whilst other countries might not separate provision and supervision as England does, they do have examinations and/or qualifications systems. Underpinning those examination systems are conceptions of educational quality. Also, examination systems in all political entities will be governed; that is, central authorities will exercise responsibilities over examinations and qualifications, and the authorities – in turn – will be held to account by the public for the extent to which the examination systems are perceived to deliver the aims for which they were set up. In the light of these facts, the researcher suggests that it is a useful and potentially insight-generating exercise to find out the extent to which different systems for running examinations and qualifications systems give rise to issues that are recognisable to English eyes, and – potentially – to find ways of doing things that might enhance English practice¹.

Although doing an international research project that explores regulation is defensible, the justification set out above does give rise to some consequences for the research that need to be managed. Essentially, these consequences flow from the fact that the conceptualisation of international research adopted in this project means that the notion of 'regulation' has to be broadly defined. This runs the risk that the research might lack focus, and that it might 'mean different things to different people'.

In response to this risk, two points can be made; the first is a mitigating step, and the second amounts to an assertion of the nature of the project. To mitigate the risk that an over-general, and therefore unclear, definition of regulation will impair the research, a section has been added to this report to cite definitions of regulation and to enumerate intellectual strands that must be considered in coming to an understanding of the subject (see pp. 17ff, below). Secondly, the sub-section that immediately follows will describe how this is a 'hypothesis generating project'. In the context of such a project it is a good thing to have a relatively general definition of the subject in hand, allowing for focusing in the confirmatory stage that follows.

¹ There is justification for this approach in the international comparative research literature. Le Métails (2002, pp. 1 – 2) lists four potential justifications for/approaches to doing international research: '*quick fix*' searches, *understanding other systems*, *exploring the causes* of one's own country's relative position, and the final justification, to which this research most closely corresponds, that is:

Informed self-review, using international indicators, relative success or weakness and contributory causes to evaluate progress from a national perspective, [which] may result in a set of ideas and priorities for action.

2.3 A hypothesis-generating project

This document reports a first pass at examining the regulation of examinations and qualifications in an international context. It is intended that this research will be followed up by a focus group of international delegates. It is the function of the current research to facilitate an understanding of an area of enquiry that can be tested in detail in the follow-up research exercise. As such, the current research can be described as hypothesis generating. This hypothesis-generating research is designed to have the following characteristics:

- The research starts by being *descriptive* and *definitional*. To some extent this will reproduce work that has been done on related issues (for instance by categorising examination systems), but this is acceptable because it relates this new research to existing fields of enquiry.
- The reporting of research findings should be *cumulative*; that is, as new types of findings are reported, they are linked back to what has already been discovered to build up an increasingly rich understanding of the area of study.
- Once the totality of research findings has been reported, there should be a reasonable understanding of the *limits of the domain* being studied. In other words, the research should tell the reader what an understanding of the regulation of examinations and qualifications based on international evidence is about (and what it is not about).
- In defining the limits of the domain, the reporting of findings in this report has been *re-organised to permit a coherent narrative* (it does not necessarily follow the order of questions in the research instrument). Also, where the research has not returned strong and clear findings, *parts of the questionnaire used are not reported*.
- As part of that definition of 'what this subject is about', the research findings should *suggest connections* between different concepts within the domain. However, it is a limitation of this kind of research that, although such associations and relationships may be noted, it may not be easy to unambiguously establish causality. This may need to be confirmed in follow-up research.
- It should be possible, even though these are 'initial' or 'generative' findings, to suggest *implications* from the findings *for English regulatory practice*. Such implications may be tentative and it is intended that they should be tested in subsequent phases of this work to enhance their robustness.
- Finally, one may state an 'acceptance criterion' that might be applied to the outputs from this kind of research. For findings of hypothesis-generating research to be acceptable, they should be *plausible* to a reasonable reader. It would fall to a follow-up research exercise to confirm whether one or more (potentially conflicting and yet still plausible) strands in the findings was in fact true.

2.4 Scope of reported findings

In order to facilitate a manageable project, certain areas were excluded from the scope of the research from the outset. Firstly, this research focused on national systems², rather than day-to-day classroom practice.

Also, 'in honour' of the newly named English regulator (see above at page 7), this report refers to 'qualifications and examinations'. In a report on the situation in England, Price Waterhouse Coopers (PWC) defined 'qualification' as follows:

A qualification is an award made by an awarding body, which signifies that learners have acquired the knowledge, skills or competence associated with the particular subject or occupation, or completed a programme of learning associated with these skills. (PWC, 2005, pp. 2 – 5)

As will be seen below, this notion needs to be treated cautiously in an international context. The English approach of awarding separate qualifications for different subjects is fairly unusual in international terms (although not unique). In rather more countries students leaving school can gain a unified 'school-leaving'/'graduating' certificate or diploma. Nonetheless, the term 'qualification' remains the best word to describe the award that can be made to a student, as defined above.

'Qualifications' as a concept can be contrasted with 'examinations', which are a method of assessment that may provide evidence that a particular candidate is eligible to be awarded a qualification³. Examinations may be also considered as entities in their own rights – for example, university entrance examinations may confer a right to enter university, but not a qualification per se (see, also: Eurydice, 2004, p. 12).

It may be inferred from the preceding paragraphs that issues concerning qualifications, examinations and cognate concepts are difficult to grapple with in an international context. In order to make such grappling manageable in this report, findings are reported in terms of the qualifications and examinations that students typically take at the end of upper secondary education (aged around 18, generally). Excluded are examinations taken at the end of compulsory education (typically, around 16 years old) or other assessment systems – for example, those used throughout schooling for monitoring national standards, or to ensure schools' accountability to parents and other stakeholders.

² Or jurisdiction-wide issues where the responding jurisdiction was a state, province, and so on.

³ They are not the only such method – there are many forms of 'teacher' or 'continuous' assessment that may provide evidence for qualifications to be awarded.

Once again, with the aim of simplifying a complex international situation, this report refers by default to general ('academic') qualifications. There are places in the report where it adds to things to refer to the situation in vocational or professional qualifications. When this is done, it is done explicitly. In other cases, the inference should be that the report is referring to general qualifications.

2.5 Research questions

Given the description in this introduction, it is now possible to state the questions that this research will address:

1. What insights can international findings provide to either support or challenge English approaches to the regulation of examinations and qualifications?
2. To what extent are English approaches to the regulation of qualifications and examinations mirrored by practices in other international jurisdictions?
 2. a. When international approaches to regulation differ from those of England, what insights can the English regulator take from international practice?

2.6 Structure of this report

The rest of this report is organised as follows. Firstly, there is a short methodology section. That section describes the data collection and analytical techniques used in the research. After the methodology section comes the data section. That section describes the types and quantities of data that were gathered. It also evaluates whether the data collected are sufficient in quantity and quality to support the inferences that will be drawn from them in this research.

Following the data section comes a section elaborating what regulation is about. This section seeks a definition of regulation and points out the different strands that must be considered in order to understand regulation in the international context.

Next comes the findings section. That section has five main sub-sections, corresponding to the main strands in the findings. Each sub-section of the findings starts by setting out 'what is already known' with respect to that strand, and/or by describing specific analytical tools that are used to interpret the findings in each particular strand.

The report concludes by discussing the implications of the international findings for the English regulator. It does so by saying why each strand of findings is important for the English regulator, and by suggesting how issues in each strand might be further investigated in follow-up research activities.

3 Methodology

The research was designed to answer the research questions set out above (section 2.5, at p. 11). The project was expected to generate hypotheses for further investigation, as defined in section 2.3, above at p. 9.

Project data were collected via the INCA and Eurydice networks. More detail on these networks is given below.

3.1 INCA and Eurydice

QCA sponsors the International review of Curriculum and Assessment frameworks internet archive (INCA) which is supported and managed by the International Information Unit (IIU) at the National Foundation for Educational Research (NFER) on QCA's behalf. The IIU is also the EU's national Eurydice Unit for England, Wales and Northern Ireland and QCA also has access to this network and its information.

INCA provides descriptions of government policy on education in Australia, Canada, England, France, Germany, Hungary, Ireland, Italy, Japan, Korea, the Netherlands, New Zealand, Northern Ireland, Scotland, Singapore, Spain, Sweden, Switzerland, the USA and Wales. It focuses on curriculum, assessment and initial teacher training frameworks for pre-school, primary, lower secondary and upper secondary education in schools (3-19 age range) (INCA, undated).

Eurydice is the European Union's information network on education in Europe. The Network produces:

- databases and reports on the organisation and structure of education systems
- indicators relating to the organisation and operation of education systems
- comparative studies of topics of current interest in education in Europe
- themed glossaries on educational terminology.

The network currently has national units in more than 30 European countries, including all the Member States of the European Union (Eurydice, undated).

3.2 Data collection procedures

Two main data collection procedures were implemented. These were:

- summary documents constructed by the IIU at the NFER in England from the INCA and Eurydice Eurybase databases
- responses to a constructed-response questionnaire provided by contacts in the INCA and Eurydice networks

These two main instruments for collecting data relevant to this research were supplemented by a review of relevant literature in the English language. This review was used to supplement the first two major data collections – for instance, to enhance the strength of findings from this research by confirming that they were supported by existing literature, or – in a few cases – by filling gaps in the data collected by the main instruments.

More information on the two main data collection instruments is given below.

3.2.1 Summary documents

Information officers in the International Information Unit (IIU) at the NFER searched the INCA and Eurydice Eurybase databases for relevant information on the regulation of assessment and e-assessment. They were looking for responses to such questions as:

- Which agency or agencies are responsible for the regulation of assessment?
- What exactly are these agencies responsible for?
- How is regulation organised?
- How are product standards established and regulated?

This information was used to produce summary documents for each of the countries represented in the INCA and Eurydice Eurybase databases.

On the basis of these summary documents, 30 countries and states were selected for further research by questionnaire.

3.2.2 Questionnaire

A questionnaire was constructed and sent to INCA and Eurydice contacts in the 30 countries and states selected following production of the summary documents. The questionnaire was designed to be constructed response, since – at this initial stage of investigation – it was felt better to permit respondents to use their own words to describe their situation, rather than being constrained by a set of (perhaps inappropriate) options from which to select.

The instrument had questions on the following themes:

- qualifications
- organisations
- aims for regulation
- documents
- use of e-assessment
- regulation of e-assessment

The responses that England would have given to the questions were inserted in a left-hand column to the questionnaire, in order to give international contacts further guidance as to what was intended by the questions. The questionnaire is shown in an Appendix – below at p. 63.

3.3 Analytical techniques

This report has already described some analytical issues in earlier sections (see: 'A hypothesis-generating project', above at p. 9). For current purposes, it should be stated that – in analysing a constructed response questionnaire and other text documents – this is a piece of qualitative research. However, the entire programme of research, including the current and subsequently planned phases, involves the provisional positing of initial ideas, followed by gathering data, which then gives rise to revised and updated theories. As such, this work may be part of the grounded theory paradigm.

Although the central instrument of this research is a questionnaire, this is not a comprehensive survey of a well-known domain. Rather, the research instrument and the associated data collection define in more detail the issues to be studied in this work; they do not give a comprehensive picture of the international situation with respect to some already-known definition.

Findings are presented in several ways; for example through bulleted lists and tables. Also, the position of England in relation to the findings varies throughout the report. When it is sensible to do so, England is included alongside the international jurisdictions. On other occasions, the English situation is kept separate. There is no fixed rule for this, rather the most convenient approach is taken on each occasion.

4 Data

4.1 Time of data collection

The summary documents were generated in summer 2007, and the questionnaires were administered in the same autumn. The quantities of data returned from the various collections are set out below.

4.2 Amount of data collected

4.2.1 Summary documents

There were 22 and 24 summary documents from the INCA and Eurydice Eurybase databases, respectively. Respondents who completed the questionnaire were also asked to check and revise (if necessary) the summary documents produced from the databases. Six contacts revised the summaries about their jurisdictions.

4.2.2 Completed questionnaires

Completed questionnaires were received from 15 jurisdictions. These were:

Denmark, Estonia, Greece, Ireland, Japan, Korea, Latvia, Netherlands, Poland, Queensland (Australia), Slovenia, Spain, Sweden, Switzerland and Turkey

4.2.3 Literature

The supporting literature review encompassed over 70 papers on specific jurisdictions. Much of this literature amounted to 'background information' (for example descriptions of national examinations and qualifications systems) rather than explicit studies of the question of regulation in examinations systems. Country studies described examination systems in Europe, Australasia and North America, and several papers described approaches in eastern Asia (Japan, Korea, China, and so on). There was a small body of literature describing issues in other parts of the world (Africa, the Middle East, South America, and so on).

4.3 Quality of data collected

The receipt of completed questionnaires from 15 jurisdictions represents a reasonable return for this international project. All the questionnaire responses were complete and informative. Most of the responding countries were in Europe, but there were also Asian respondents (*Japan* and *Korea*) and one Australian state (*Queensland*).

It is important to note that the questionnaire responses came from countries and one state that had diverse political arrangements. Several of the countries in the sample were members of the supra-national entity the EU, which does have regulatory implications – see p. 22, below). Some of the responding countries were either constitutionally federal states (*Switzerland*) or ones in which regional authorities were guaranteed a role in education policy

(*Spain and Poland*). *Queensland* is a state within the nation of Australia. The *Queensland* findings are presented in their own right, rather than as a proxy for Australian approaches to the regulation of qualifications and examinations. To reflect the diverse constitutional statuses of the entities on whose behalf questionnaires were completed, the report refers to 'jurisdictions' as a generic term⁴, rather than 'countries', 'nations' or 'states'.

It is unfortunate that – although several were approached – no North American jurisdiction completed and returned a questionnaire.

The database summary documents were – by and large – slightly less informative than the completed questionnaires. Jurisdictions that revised the summaries to reflect their own understanding of the issues generally provided a wealth of relevant information. However, the archives do not set out to collect information on the regulation of qualifications and examinations systems and so archived information tended to address background issues (such as the structure of examination systems), rather than directly discussing regulation.

Like the summary documents, the literature review addressed background issues, rather than the research question directly. One useful facet of the review was that it did provide insight into practices in North America – compensating to some extent for the absence of questionnaire responses from that part of the world.

The overall evaluation is that there was sufficient information upon which to base this research project and with which to attempt to answer the research questions (see above at p. 11).

⁴ Jurisdiction is defined as: a 'political boundary within which officials have authority'.

5 Background issues

The purpose of this section is to define, describe and contextualise the English term 'regulation'. This is done so that the manner in which the English experience and the international findings interact can be contextualised. It is important to contextualise the term, as it was suspected – before this research was conducted – that international jurisdictions did not employ the concept of 'regulation' in the way that England does (see p. 9, above).

Finally, the need for this section of the report arises from the fact that developing an understanding of regulation involves synthesising evidence from very diverse practical and intellectual fields. These include regulation as:

- a sub-set of government (essentially a political science concept)
- an important element in the history of examinations and as a concept that arises from the histories of the several types of organisations that are involved in it
- a practice that implements an understanding of what constitutes quality in educational assessment
- something that will, in the future, be influenced by developments such as new political arrangements (for example the emergence of new supra- and sub-national forms of government) and the increasing use of technology in education.

5.1 Some definitions of regulation

Finding a convincing definition of regulation is surprisingly difficult. In the introductory article for a new journal on regulation and governance, the editors propose the following definition:

Governments and governance are about providing, distributing, and regulating. Regulation can be conceived as that large subset of governance that is about steering the flow of events and behavior, as opposed to providing and distributing. (Braithwaite et al, 2007, p. 3)

QCA has described its regulatory activity in the following terms:

QCA regulates awarding bodies, qualifications, examinations and national curriculum assessments to ensure that the qualifications market is fit for purpose, that qualifications are fair, standards are secure, public confidence is sustained and that QCA acts as the public champion of the learner. (QCA 'Key Result Area', quoted in Nisbet & Greig, 2007, p. 1)

In the questionnaire that formed the central research instrument for this project, the following was given as a definition of regulation:

Regulation refers to measures taken in the public interest to assure the quality of examinations and qualifications. It is a broad concept and can be interpreted to include other concepts, such as quality assurance.

Thus, for the purposes of this report, regulation can be said to have two faces: firstly, it is about the supervision of provision on behalf of the public and secondly, it contains within it a notion of quality – in this context that being a notion of educational quality. Also, regulation concerns the provision of public sector type services (education, transport, health, and so on) even if those services are in fact delivered by the private sector.

5.2 The range of factors that affect regulation

Oates (2005) asserts that four considerations are important to understand the regulation of qualifications in an international context:

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), and 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 vocational qualifications.

Oates (2007) states that it would be useful if economists, historians and political scientists wrote about assessment matters. This paper is written by an educationalist, but in the following sections it outlines some wider social features that may assist an understanding of the regulation of assessment and e-assessment.

5.3 Histories of examination arrangements

In order to understand how an exams and qualifications system is currently organised (including how it is regulated), it is important to understand its historical development.

The modern public examinations system in England has roots at least as far back as 1836 (Tattersall, 2007, p. 43). From that time, universities were involved in setting examinations for school leavers – often selection exams for scarce university places. Over time, universities have ceased to have a day-to-day role in the setting, administration and awarding of exams in schools (*ibid.* at p. 73) with only Cambridge University retaining ownership of an awarding body (exam board). In the present day, awarding bodies function as independent companies, forming a market sector with a few large organisations and a larger number of smaller bodies (PWC, 2005).

There is also a story of change in the regulation of qualifications in England. Thirty or so years ago, the English education system was more lightly regulated than its near neighbours (Meuret & Duru-Bellat, 2003). However, under successive governments, education has been a central political issue (see Callaghan, 1976). Government interest in education has been accompanied by a succession of regulatory bodies (Nisbet & Greig, 2007, p. 2) and – some argue – successively, regulators have increased their control on the operations of awarding bodies (Tattersall, 2007, p. 95).

The most recent trend has been for the regulator to embrace the UK government's approach to 'better regulation' – by implementing the maxim 'less is more' and by regulating at higher, more strategic levels and trusting awarding bodies to implement their own internal quality control practices (Nisbet, 2007).

This is a very brief outline of the history behind current examinations arrangements in England. It is self-evident that this history will be different from the histories of other nations that respond in an international survey, and that those differing experiences of qualifications and examinations will have an impact on current understandings (see: West & Crighton, 1999 and Lubisi & Murphy, 2002 describing how moves to democracy in Eastern Europe and South Africa, respectively, impact on current qualifications systems).

5.4 Histories of regulation

The history of regulation as understood in the modern sense can be traced to many roots. Spyrelli (2003, p. 8) quotes a legal case of 1877 in the US state of Illinois as establishing the rights of governments to regulate businesses. The appeal court judgement in that case outlined the important 'public utility principle' underpinning the state's right to regulate certain activities:

When, therefore, one devotes his property to a use in which the public has an interest, he, in effect, grants to the public an interest in that use, and must submit to be controlled by the public for the common good, to the extent of the interest he has thus created.

Braithwaite et al (2007, p. 1) cite measures enacted by Woodrow Wilson and Franklin Roosevelt in the first half of the twentieth century in the United States as early examples of regulation. Similarly, developments in regulation were encouraged by scandals in the 1970s under Nixon (*ibid.* at p. 2).

However, it has been claimed that an important stimulus for the growth of the 'modern regulatory state' was the widespread privatisation of previously nationally owned utilities (a

phenomenon that occurred in a range of developed countries, including EU member states and Australia, in the 1980s). Privatisation was accompanied by the institution of 'arm's length' 'independent regulators', which inherited powers previously exercised by central government to defend the wider interests of the general public, in the face of private sector providers with primary obligations to increase share value.

Grant (undated) explains the logic behind this process:

[Arm's length regulators] were favoured precisely because they were free from elected political influence. The argument is that delegation improves electoral legitimacy and is essentially a response to general public cynicism with traditional modes of governance. In the absence of political flexibility, citizens and investors are likely to be more assured of optimal policy outcomes than if the process was left to politicians' discretion. (Grant, undated, p. 5)

It should be noted that this context for thinking about regulation – and the consequent association with economic matters – is different from the history of examinations in England sketched above. It is true that the regulation of qualifications has received increased prominence in England since the 1980s, contemporaneously with the development of 'arm's length' regulators of privatised utilities. However, the distinct roots of the regulation of qualifications in England means that analyses of post-1980s regulation can only be imported to the qualifications context cautiously.

Similarly, whilst a large number of developed nations have shifted the paradigm for the provision of public utilities from state- to private-sector control, there has not been a corresponding privatisation of examination provision in European countries (and, in England, exams provision has always had the element of non-state ownership). Once again, this suggests that analyses of privatised utilities regulation should only be used cautiously in the context of examinations.

5.5 International understandings of regulation

Different nations and regions have different histories, each of which impacts on our understanding of current policies and practices with respect to the regulation of assessment. It is also relevant, and not surprising, that parts of the world with differing political traditions have different approaches to regulation.

Falch (2006) has argued that different regions put varying emphases on the role of the state in controlling markets. He asserts that EU jurisdictions have put more emphasis on public authorities acting as guarantors of the public good, whereas in the US (and to some extent

Japan) the trend has been for the private sector to self regulate, and to solve any ensuing disputes through civil actions in courts (Falch, 2006, p. 139).

5.6 The quality aspect of regulation

A central concept in understanding the quality of an educational assessment is validity. Moss et al (2006) provide a recent and comprehensive history of the idea of validity.

Perhaps the best known recent definition of validity is that of Messick:

Validity is an integrative evaluative judgement of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment. (Messick, 1989, p. 13, emphasis in original.)

QCA, working with other UK regulators has produced a series of documents that set out procedures that must be followed in the design, development, administration, marking and reporting of qualifications (QCA et al, 2004; QCA et al, 2007a). The UK regulators' documents include a specific set of regulatory principles for e-assessment (QCA et al, 2007b).

In regulating qualifications, the QCA and its sister regulators seek to implement the following facets of a definition of educational quality: validity, reliability, fairness, manageability and providing learners with a clear means of progression (Charlton & Higson, 2008). It has also been observed that a key concern for examinations and qualifications in England is comparability – between different awarding bodies' qualifications, between subjects, between years and so on (see Newton et al, 2007).

5.7 Regulation in the future

New technologies have blurred the distinction between media. For example, one can now watch TV or browse the internet on a mobile phone, receive interactive services via a TV or listen to a radio programme via an internet-connected computer. It is said that such convergence of previously distinct media is likely to be followed by convergence of previously distinct regulators. For example, Henten et al (2002, p. 21) have developed the following figure to show the current and future convergence of regulators:

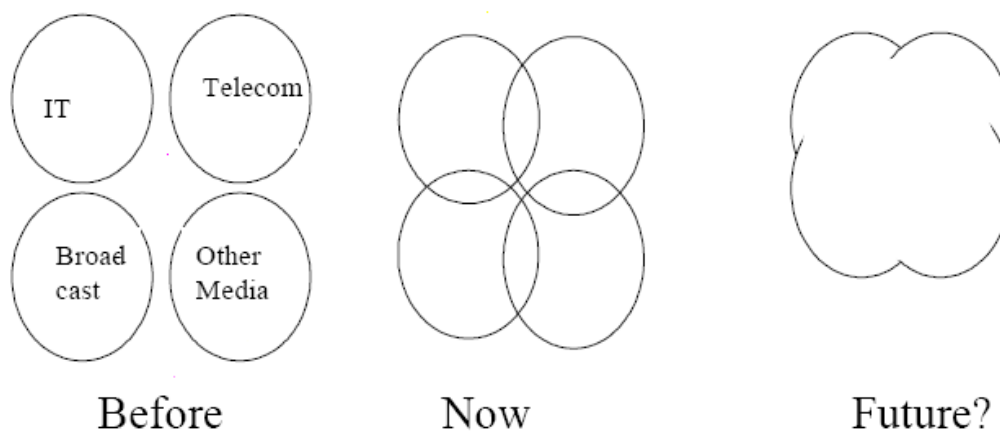


Figure 1: Possible steps in the convergence process

Technology has been predicted to have an impact on the practice of examinations as well. Some have predicted that the internet will cause testing to become increasingly internationalised. Bartram has given this hypothetical illustration of an assessment future in which 'location, for all practical purposes, no longer exists':

An Italian job applicant is assessed at a test centre in France using an English language test. The test was developed in Australia by an international test developer and publisher, but is running from an ISP location in Germany. The testing is being carried out for a Dutch-based subsidiary of a multi-national. The position the person is applying for is as a manager in the Dutch company's Tokyo office. The report on the test results, which are held on the multi-national's Intranet server in the US, are sent to the applicant's potential line-manager in Japan having first been interpreted by the company's out-sourced HR consultancy in Belgium. (Bartram, 2006, p. 123)

Political developments may also impact on regulation; causing the discipline to be more internationalised. The existence of supra-national organisations (such as the EU) impacts on national policies and practices. The EU's 'Open Method of Co-ordination' (OMC) (Régent, 2002; Chalmers & Lodge, 2003; Jacobsson, 2004) is a 'soft form of regulation' that could have an impact on examinations policy. Also, Oates (2007, p. 12) cites Germany's development of federal testing systems over and above state or provincial mechanisms in response to poor results on the Programme for International Student Assessment (PISA) (see also: Ertl, 2006; Kuiper et al, 2007, p. 7). Thus, whilst PISA is not a direct regulatory measure, it can have profound impacts on the organisation of examinations systems.

5.8 Summary of this section

This section has described a diverse set of issues. Therefore, this summary expresses the key points from the section succinctly, before moving on to the research findings.

- 'Regulation', has many definitions, but at its broadest it means actions taken by governments in the public interest to steer events and behaviour.
- Understanding regulation involves understanding elements of political theory – in particular, the role of governments in steering (or supervising) activities, rather than providing or distributing.
- Regulation increased in significance as an area of enquiry following the privatisation of public utilities during the 1980s. This applies to the United Kingdom and to other countries in various parts of the world (such as Europe and Australia).
- The provision and (allegedly increasing) regulation of qualifications in England has a distinct history, which is different both from that of qualifications in other jurisdictions and from the post-1980s experience in other UK public utilities.
- Understanding regulation requires an understanding of the notion of quality in educational assessment. Aspects of assessment quality as understood in the English regulatory context include: validity, reliability, fairness, manageability and comparability.
- Some commentators have predicted that the practice of regulation will change in the near future. Such changes are of two types: convergence of regulators and dealing with internationalisation. They also arise from at least two directions: new political arrangements (for example the advent of the EU) and the rise of technology.

6 Findings

6.1 Types of qualifications

Le Métais (2002, p. 24) suggested that most countries make available upper school diplomas that a student can receive on satisfactory attainment of a block of core subjects. An alternative model (used in England, New Zealand and some Australian states) was for students to be able to combine different courses leading to individual accredited qualifications, which resided within an overall framework. Sweden and the USA were given as examples of jurisdictions in which courses provide set numbers of points that count towards high-school graduation.

The findings of this new research confirmed Le Métais' supposition. Amongst the sample of jurisdictions that responded to questionnaires only students in *Denmark, England* and *Queensland* may be awarded discrete qualifications in particular subjects. In all the other surveyed jurisdictions, students work towards achieving a school leaving or graduating certificate or diploma.

Students who had attained a school-leaving certificate in different jurisdictions in the sample would have different paths (and rights) if they were contemplating entering Higher Education (HE). In several jurisdictions (such as *Poland* and *Turkey*), possession of an upper-secondary school diploma gives the holder an automatic right of entry into HE.

In other countries there are university entrance examinations in addition to the school leaving diploma. In *Japan*, there is a standardised national university entrance examination, whilst in *Korea* admission to HE is based on high school record, involvement in extra-mural activities and scores on nationwide HE entrance examinations. *Sweden* also has a national university examination. However, in that country the examination is not a mandatory component of HE admission for the majority of students, as in *Japan* and *Korea*. Rather, the SweSAT is generally taken by a small percentage of (typically mature) students who had not gained the necessary credits in school to access HE.

Unfortunately, no United States jurisdictions responded to this research's questionnaire. However it is possible to say something about the qualifications system of the USA, based on existing literature.

Black & William (2005, p. 256) state:

The most important feature of the education system in the United States is that there isn't one.

Isaacs (2001, p. 392) states that the USA does not have a national diploma or qualifications framework in the way that might be understood in many European jurisdictions. Isaacs identifies three assessment mechanisms that are extremely important to young people in the USA making the transition between school and Higher Education. These are:

- A school-produced transcript listing courses taken, grades achieved and so forth. These grades are then aggregated using a range of formulae.
- The SAT (formerly the Scholastic Assessment Test⁵): a test made up of three verbal and three mathematical sections, traditionally constituted mainly from multiple-choice items, but now also containing an essay section (College Board, undated). The score that a student achieves on SAT is extremely important to determine the university that they will have access to.
- Advanced Placement (AP) program examinations. Described by Isaacs (2001, p. 403) as 'the *de facto* national honours curriculum in the US'. This suite of subject-based, criterion-referenced examinations are taken by increasing numbers of young Americans to assist them to gain entrance to university.

6.2 Facets of the examinations and qualifications process

The following section looks at a set of activities that are related to the conduct of examinations and the awarding of qualifications. These include: defining standards for assessment, the assessment methods used and awarding qualifications once assessment has been conducted. International findings in respect of these processes are set out below. The processes are described in the order in which they would occur in practice.

In analysing such processes, one needs a way of grouping international approaches. It is proposed to consider the extent to which processes suggest 'centralisation' or 'de-centralisation'. This technique follows the work of Kuiper et al (2007), who categorised jurisdictions' approaches to curriculum policy and practice along a 'centralised-to-decentralised' continuum. Kuiper and his colleagues do not give formal definitions of 'centralisation' and 'de-centralisation' in their paper. However, their examples of the respective tendencies give readers a reasonable idea of what the concepts mean. De-centralised type features of curriculum policy and practice include: targets being less detailed and only covering a portion of the curriculum; and schools and teachers having opportunity for site-specific curriculum choices. In contrast, centralised type features in curriculums

⁵ Many students in the Mid-West of the USA take the American College Test.

include: centrally formulated prescriptive standards; centralised high-stakes assessment and rigid external evaluations (*ibid.* at p. 1).

Kuiper et al characterised the respective ends of the centralisation cline as emphasising prescription or professionalism (of teachers) in curriculum use. Kuiper and his colleagues' use of the terms 'prescription' and 'professionalism' gives one a clear indication of their preference for de-centralisation in curriculum policy and practice. In thinking about the regulation of examinations and qualifications, the centralised to de-centralised continuum is adopted without the same value judgement. It is quite possible for a jurisdiction to adopt either a centralising or a de-centralising measure as a way of re-assuring the public about the quality of examinations and qualifications.

6.2.1 Defining examination and qualification standards

This first part of the examinations process relates to the definition of what amount (or type) of knowledge, skills or abilities will be sufficient for candidates to be awarded qualifications. Typically, the output in this type of activity would be a document outlining what needed to be learnt in order for a candidate to be awarded a qualification. Such documents can go under a number of names, such as assessment or subject criteria, subject specifications, or even programmes of work.

For clarity's sake, it is emphasised that the findings in this sub-section relate to that definitional work that goes on long before an assessment process (for example an examination) is actually run. It is distinct and different from standards setting in the sense of setting a pass/fail type cut score after an examination or moderating teacher assessment judgements to ensure comparability.

Findings for different jurisdictions are grouped around cognate concepts (rather than listing findings by jurisdiction). Although there is no absolute single dimension of centralisation to de-centralisation, those findings that suggest more centralisation are reported first.

In several jurisdictions, the central education ministry has a role that has not been limited or delegated in defining examination and qualifications standards. Examples of such jurisdictions include *Greece*, where the Ministry of National Education and Religious Affairs has responsibility for all aspects of examination and qualifications regulation. Also, in *Turkey* the Ministry of National Education is responsible for all aspects of examinations regulation (that is in general qualifications – private institutions and companies are involved in the development of vocational qualifications).

In some jurisdictions, a separate body is responsible for defining examinations and qualifications standards. Such bodies may be public sector in nature, but are often formally separate from the education ministry (see sub-section 6.3.3, below at pp. 39ff for more on this issue). The interaction of such separate bodies with other organisations in defining examinations standards can have several facets. These are listed and exemplified below:

- There may be a single body separate from the education ministry that defines examination standards. Jurisdictions with such a set up include: *Latvia, Estonia, Queensland* and *Sweden*.
- Other jurisdictions have more than one non-ministerial body involved in examinations and qualifications, and the organisation that defines the examinations and qualifications standards has interactions with separate organisations that perform other tasks (such as developing and/or running examinations). Such jurisdictions include: *Ireland* and *the Netherlands*.
- A slightly different way for a non-ministerial organisation definer of examination standards to interact with a third-party organisation is exemplified by *Slovenia*. In that country, the National Council of Experts and the education ministry develop qualifications. However, National and Subject Examination Committees also provide advice – for example in relation to their subject expertise.

Another set of issues that show how jurisdictions can introduce de-centralised aspects into the definition of examinations and qualifications standards concerns the involvement of regions in the process. For example, in *Poland*, the Central Examination Commission develops standards of attainment, but is supported in its work by eight Regional Examination Commissions. *Danish* education legislation enshrines central oversight of educational standards, but gives significant powers to municipalities to 'lay down the targets and frameworks of activities in schools' (Egelund, 2005, p. 207). *Switzerland* is the jurisdiction with the greatest vesting of authority with regions in this sample, being a formally federal state. In *Switzerland*, cantons are sovereign, and central authorities can only pass regulations when lower authorities are not in a position to do so.

By this point, one has reached the more de-centralised end of the continuum with respect to the definition of examination and qualifications standards. There are two approaches that suggest considerable de-centralisation in this area. In *Queensland*, schools are able to propose work plans for the central Queensland Studies Authority (QSA)'s approval. *Japan* and *Korea* do not have nationally standardised qualifications systems for the end of post-compulsory education. Rather, such qualifications are awarded on the basis of unmoderated teacher assessment. Caution should be exercised in interpreting this finding – both *Japan* and *Korea* do have university entrance examinations and – certainly in the case of *Korea* –

could be characterised overall as centralised education systems (Government of WA, EAD, 2006, p. 3; INCA, 2005, p. 2).

6.2.2 Assessment methods

The centralised versus de-centralised metaphor set out above can be applied to this part of the examinations and qualifications process. Conceptualising assessment methods in such terms, one would perceive external examinations as being a centralised procedure (confidential question papers written at the behest of a central agency, administered under controlled conditions, marked under the supervision of a central body, and so on), whereas assessment methods over which schools and teachers have more control (for example continuous, teacher assessment) are a de-centralised assessment mechanism.

There is already a body of information regarding the relative preponderance of use of 'centralised' and 'de-centralised' assessment methods in international jurisdictions. Such work is summarised below.

The Organisation for Economic Co-operation and Development (OECD) has recently published a major description of international education systems (OECD, 2006). In that publication, the OECD lists assessment methods for general and vocational qualifications at International Standard Classification of Education – 1997 version (ISCED-97) level 3 (OECD, 2006, table X1.3, p. 425). The following table is a summary of a larger table in the OECD publication, intended to show the relative frequency of occurrence of internal and external assessment models in general qualifications at ISCED-97 level 3:

	Final examination	Series of examinations during programme	Specified number of course hours AND examination	Specified number of course hours only
No	9	10	11	25
Yes	15	18	17	2
'Yes and No'	5	1	1	1
Information not available	1	1	1	3
Total	30	30	30	31 ⁶

Table 1: Interpretation of OECD data showing assessment methods for level 3 general qualifications

⁶ This table summarises the OECD table, upon which many caveats and riders are expressed (see OECD, *ibid.* for these exceptions).

The column totals differ by one because the USA is considered to be able to be categorised as 'yes' for 'Specified number of course hours only' whereas for all the other columns, mention is made of the different arrangements enacted by individual states.

Eurydice (2005) has developed the following figure to show assessment methods in the nations within the Eurydice network:

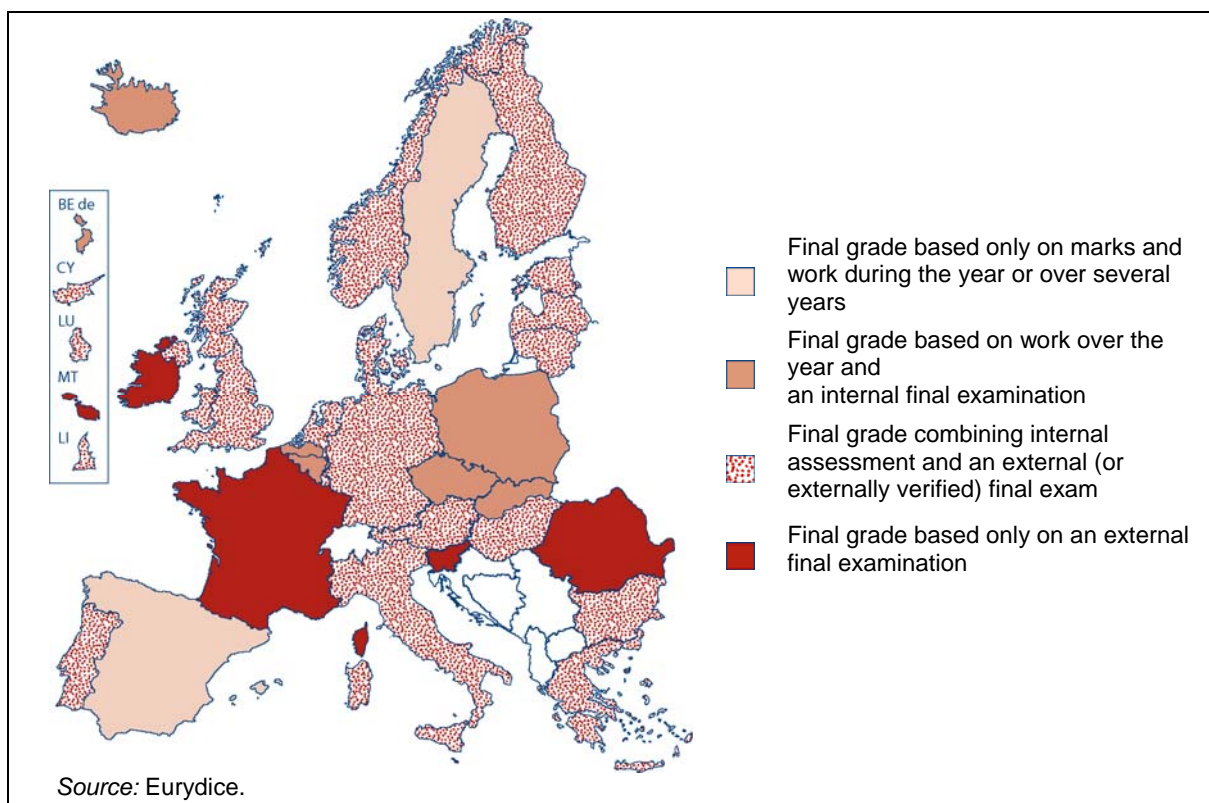


Figure 2: Certified assessment at the end of general upper secondary education, 2002/03

Phelps' study (2000) suggests that there is an extension of 'standardized testing' methods worldwide. His research found that most recent innovations worldwide in large-scale assessment had been to introduce formal testing, rather than to remove it or reduce its use. Morrison (1994) also estimated that centralised, formal examining was becoming the assessment method of choice. However, Phelps and Morrison make diametrically opposing interpretations of the desirability of the trend that they both discern.

In the current study, jurisdictions have been categorised as 'mainly using external (centralised) assessment', or 'mainly using internal (de-centralised) assessment'. The distribution of countries into those two categories is shown in the following table:

Mainly using external assessment	Mainly using internal assessment
Denmark Estonia Greece Ireland Latvia Netherlands Poland Slovenia England	Japan Korea Queensland Spain Sweden Switzerland Turkey

Table 2: Categorisation of jurisdictions as using 'mainly internal' or 'mainly external' assessment

This juxtaposition is defended as a useful and meaningful analysis of jurisdictions' approach to assessment. For example, one can juxtapose features of the Japanese assessment system

There are no formal national systems of assessment for students in compulsory education in Japan. ...

The grade [for the 'Certificate of Graduation from Upper Secondary Education'] is based largely on teacher assessment; there is no external moderation.

with this feature of the Irish system:

There is no school-based assessment component towards the qualifications; all assessment is external.

However, Japan and Ireland would be at opposite ends of an imagined continuum of assessment methods for school-leaving qualifications. Most jurisdictions could be conceived of either as moderately favouring internal or external assessment, or of having mixed systems. In the following paragraphs, the nature of that mixing is described in more detail.

The clearest indicator of centralisation in assessment methods is the use of external examinations. All the jurisdictions listed in **Table 2** as 'mainly using external assessment' (that is nine from 16 jurisdictions) used external examinations as a method of assessment for academic post-compulsory school-leaving examinations. *Ireland*, as illustrated by the quote on the previous page, was unique in that all assessment was external. The other 'mainly external' jurisdictions had some internal facets of their examination process which mitigated the 'purity' of the external terminal examinations model.

- In some jurisdictions (for example *Denmark*, *Greece* and *Poland*) written external exams are supplemented by oral components, which are conducted and graded by candidates' teachers.

- In other jurisdictions, the externally set examinations may be supplemented by an internal test or exam. *Estonia*, *Latvia* and *the Netherlands* are jurisdictions in which such internal measures are used.
- In *the Netherlands* teachers mark the external examination, as well as constructing, setting and marking the internal assessment components.
- *Slovenia* permits a wide variety of assessment methods, in addition to external examinations. This variety can be expressed as follows:

Exams can be written; written and oral; written and practical; written, oral and practical; only practical or can take the form of a demonstration.

A minority of jurisdictions in this sample (7 from 16) have been categorised as favouring internal assessment. This tendency has several facets – stereotypically it is about assessment being conducted locally, by teachers. Often this is synonymous with continuous assessment. Coursework is a preferred method in some jurisdictions (for example *Korea*). 'Favouring internal assessment' can be associated with teachers and schools having a choice as to the assessment methods to be used (*Spain*), or with teachers marking the tests used for the *Matura* examination (*Switzerland*). In *Sweden* the preference for local decision-making is partly operationalised by deliberately expressing grade criteria in relatively general terms, and thus making sure that a large portion of decision-making power remains with the teacher (Ramstedt, 2005, p. 15). The final point to illustrate how jurisdictions keep decision-making de-centralised is the disinclination of some (for example *Japan* and *Sweden*) to moderate or 'normalise' grades set by teacher assessment.

Whilst 'mainly favouring internal assessment' is a meaningful and distinct category, no jurisdiction in this sample is absolutely 'pure' in favouring internal assessment. All jurisdictions have one or more measures in place that provide a degree of centralisation (jurisdictions might well prefer to use a term such as 'standardisation') to counterbalance the prevailing tendency towards localisation.

- Even though they allow teachers a large say in determining students' grades, several jurisdictions put controls in place with the aim of increasing standardisation. For example, in *Queensland* the central Queensland Studies Authority (QSA) has considerable powers to develop syllabuses and other materials, and to accredit the same developed by third parties (State of Queensland, 2002). Similarly, in teacher-assessed *Turkey* central authorities issue regulations affecting the local conduct of assessment.
- Another 'standardising' type feature is that present in *Sweden*, where mandatory, standardised tests exist in some core subjects, which must be administered, although with a function limited as follows:

The national tests system shall not:

- influence the choice of teaching content and teaching methods (since this shall be determined by teachers and students) ...
- function as final examination tests (as teachers should award grades based on the assessment of the student's accumulated work or demonstrated knowledge and skills). (Ramstedt, 2005, p. 19)

- In *Switzerland* central control is balanced against local autonomy by allowing teachers to set their own tests for the *Matura* exam, but requiring them to get the completed tests approved by experts appointed by the cantonal authorities.
- Local assessment can be controlled or moderated in several ways, to increase standardisation. In *Switzerland*, teachers' marked exam scripts are sent to an expert for approval, whilst *Queensland* has a sophisticated system for moderating teacher-assessed grades to ensure comparability of State-wide standards.
- Some ostensibly internal-assessment-using jurisdictions back up the local assessment for school-leaving qualifications with external examinations for university entrance. In *Japan* and *Korea* all students have to sit some form of university entrance examination, which can take on very high stakes and thus have negative consequences, such as being stressful for candidates and distorting the conduct of middle and high school education (MOEHRD, 2004, p. 13). Somewhat different scenarios exist in *Queensland* and *Sweden*. *Queensland's* questionnaire response states that:

The QSA maintains external examinations at the end of Year 12 in a limited range of Authority subjects as a safety net for those few students who are unable to be assessed at a school.

The *SweSAT* examination is also mainly intended for students accessing university through non-standard routes, although it has also been quite popular with those coming to higher education via academic routes (Wikström, 2006, p. 122). Ramstedt (2005, p. 18) describes selection rules for Swedish HE which state that at least one third of entrants should be selected on the basis of their performance on the *SweSAT*.

6.2.3 Awarding certificates

There is a further way in which a researcher might understand the extent to which central jurisdictions retain or delegate authority to issue examination and qualifications certificates. On the face of it, if a central education ministry retains the power to issue certificates this could be taken as a centralised approach. In contrast, if the authority to issue certificates had been delegated, this might be taken as evidence of de-centralisation. The following are a set of the main types of arrangements that have been discovered in the research data for the issuing of examinations and qualifications certificates:

- The central (typically the nation state) ministry of education issues examinations and qualifications certificates in its own right. This arrangement exists in *Greece* and *Turkey*.
- A non- or quasi-governmental, 'arm's-length technical body' (see page 39, below for a fuller discussion of these bodies) has responsibility for awarding certificates. This approach is used in *Estonia*, *Ireland* and *Queensland*.
- Responsibility/authority for issuing certificates is shared between the nation state government (education ministry) and that of regions. This arrangement can have several variants – for example it may be combined with the involvement of 'arm's-length technical bodies' (see previous bullet point), either at the national and/or regional level – *Poland* and *Switzerland* both have variants of this model.
- Regional authorities issue examinations or qualifications certificates in their own rights – in *Spain* the education departments of autonomous communities issue examinations and qualifications certificates.
- Schools have a role in issuing certificates. This can be jointly with the central education ministry, as in *Denmark*, where the certificates are issued by the national education ministry, but also signed by the school headteacher (Eurydice/Eurybase, 2007, p. 65). Alternatively, in perhaps the most de-centralised model described in this list, the each individual school principal can issue a qualification certificate in his/her own right – this is the model in *Japan*.

6.2.4 Other possible areas of interest

The purpose of this work is to investigate the issue of regulation while seeking insights from different jurisdictions worldwide, not to categorise all aspects of examinations and qualifications processes along a centralisation continuum for all jurisdictions. However, there is a set of other issues that might help a researcher to gain an understanding of how jurisdictions balance centralising and de-centralising approaches to guaranteeing that examinations and qualifications retain high public esteem. Such issues might be further investigated in follow-up research work. The issues include: techniques used in setting cut scores/pass marks and, more broadly, in maintaining the standards that examinations and qualifications embody; the extent to which qualifications systems are made up of compulsory and optional elements and the nature of what is reported on certificates (for example grades, simple pass/fail or a numerical value, such as a 'grade point average').

6.2.5 Exceptions, limitations and trends

The following paragraphs make explicit three important features of the findings of this subsection that may have been only implicit in the text above. Firstly, most of the jurisdictions use both standardising and localising techniques to ensure that the public continue to hold

examinations and qualifications systems in high esteem. When jurisdictions are categorised as 'mainly centralised' or 'mainly decentralised', it is essential to retain the qualifying adjective; no jurisdiction maintains a *wholly* centralised or de-centralised system.

Secondly, these research findings relate to the ways in which jurisdictions ensure the public interest with respect to general examinations and qualifications taken at the end of the post-compulsory phase of education. They can stand as a proxy to some extent for techniques used in other exams and qualifications (such as those taken at the end of compulsory education⁷) or – to a somewhat lesser extent – to assessment approaches used in vocational qualifications or in national assessment systems used throughout schooling. However, it would be inappropriate to use findings on the centralisation of techniques to re-assure the public in 18-plus examinations to generalise to statements about the centralisation of the entire education system (cf. example of *Korea* at p. 32, above).

Thirdly, and finally in this sub-section, it should be noted that jurisdictions may well change the ways in which they prioritise centralising or de-centralising techniques in maintaining confidence about examinations and qualifications over time⁸.

These three points should be borne in mind when considering the continuum diagram that follows.

6.2.6 Continuum of centralisation for examination processes

Kuiper et al (2007, p. 3) developed a continuum diagram to show the extent of centralisation and de-centralisation in curriculum policy and practice. That endeavour was the inspiration to create a similar continuum showing the extent to which jurisdictions in this study value centralising or de-centralising measures to re-assure the public about examinations and qualifications.

The metaphor of a continuum of centralisation was adopted despite its potential limitations. These were: firstly, that the continuum was not based on a formal scoring procedure (as a continuum, such as an ability scale, would generally be in reporting the scores from an educational or psychological test). Secondly, it was understood that this single continuum conflated more than one dimension in which jurisdictions might take more or less centralised

⁷ There are exceptions even here – for example, *Slovenia* takes a relatively less centralised approach to end of compulsory schooling examinations than to end of post-compulsory schooling.

⁸ See for example, Looney (2006, p. 351) discussing the possibility that *Irish* approaches will become less centralised in the near future; Government of WA EAD (2006, p. 1) and MOEHRD (2004) discussing the six changes to assessment and curriculum regimes that have occurred in *Korea* since 1955; and Wikström (2006) and Ramstedt (2005) describing the major shift to a less centralised system that occurred in *Sweden* in the 1990s.

approaches (for example a jurisdiction might have centralised awarding, but de-centralised assessment methods – as *Queensland* does).

These limitations are acknowledged to reduce the validity of the continuum as a 'scientific instrument'. However, the continuum is presented both as a useful summarising tool for policy makers and as a spur to further phases of research. It is believed that it is a useful tool for those purposes. Some justifications for the placing of jurisdictions on the continuum are given below the figure.

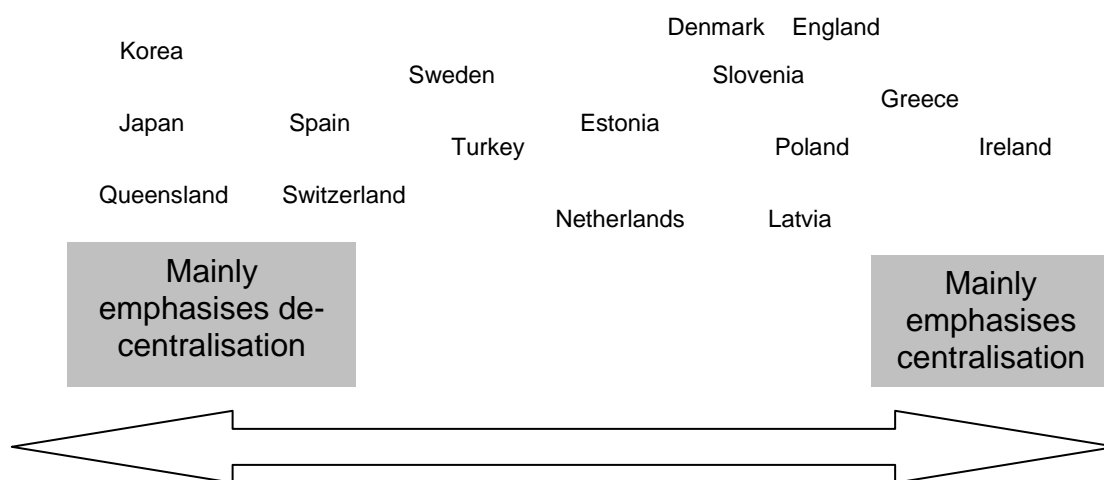


Figure 3: Continuum showing centralisation of techniques to assure public confidence in the examinations process

One can discuss the positioning of jurisdictions by 'working in from the edges' – taking the strongly de-centralised or centralised jurisdictions first. *Korea*, *Japan* and *Queensland* have been located at the 'de-centralised' end of the scale. This is despite the fact that they have very different approaches to re-assuring the public about examinations and qualifications. The qualifications gained by students in *Japan* and *Korea* are assessed by unmoderated teacher assessment, which is (in *Japan*) certified by the school principal⁹. *Queensland*, in contrast, has a fairly strong set of central controls through its systems for publishing and accrediting subject syllabuses, and for moderating teachers' assessments. However, it has a strong commitment to teacher assessment.

Spain and *Switzerland* also use teacher judgement as the prime assessment method, and both – for their own historical reasons – give substantial powers to regions in the setting of standards, and the awarding of qualifications. There are standardising tendencies, however,

⁹ Both these jurisdictions also have university entrance examinations, which have not been considered for the purposes of constructing this continuum.

that balance the localisation. For example, in *Switzerland* teachers are required to get the tests that they construct approved by cantonal authorities before using them for assessment.

Ireland has been nominated as the jurisdiction that places most emphasis on standardised approaches to maintaining public confidence in the examinations and qualifications process. In its general qualifications, *Ireland* uses external assessment methods only, and correspondingly, standards definition, certificate awarding and other aspects of the examinations process are conducted externally to schools.

Greece is placed next at the centralised end of the continuum. The *Greek* ministry of education has not delegated its functions to regions or an arm's-length technical body, and it also favours external examinations as an assessment method. However, its qualifications do reflect students' participation in daily classroom work and their overall activity at school, their performance in written tests and their composite creative projects.

England has been placed as having the third most centralised approach to maintaining public confidence in examinations and qualifications. Standards for examinations and qualifications in England are centrally defined, and external examinations are the main method by which students are assessed at 18. De-centralising tendencies do exist, however. Many A level qualifications have contained coursework elements (QCA, 2005); an internally assessed, de-centralised measure. Also, assessment for A levels is now modular, spread throughout the two-year course rather than being purely terminal. As such, students, parents and teachers now have more involvement in the process by deciding on entry to modules, re-sits, and so on (see: Matthews & Pepper, 2007).

Many of the jurisdictions in the centre of the continuum are closely bunched and their positions may well overlap. For example, *Sweden's* centralisation and de-centralisation is balanced between its preference for teacher assessment and its insistence on teachers' use of external tests as a controlling mechanism for their judgement. *Turkey* uses internal assessment, too, but its ministry of education has not delegated powers to regions or an external technical body. Similarly, on the 'slightly centralised' side of the continuum, one can observe *the Netherlands* and *Estonia*, both of which employ external examinations, but both of which also use internal components to contribute to the grade given in qualifications. Similar comparisons and contrasts can be made for all the jurisdictions on the continuum.

6.2.7 Summary of this sub-section

The following bullet points present a summary of the findings from the current sub-section.

- All jurisdictions guarantee the public interest in the exams and qualifications process.

- Ensuring public confidence in examinations and qualifications processes can be thought of (with some reservations) as a part of the regulation of the examinations and qualifications process.
- Jurisdictions' techniques for defending public interest with respect to examinations and qualifications processes typically contain a mixture of centralising and de-centralising features.
- Centralising and de-centralising features can occur at different points in examinations and qualifications processes (for example defining qualifications standards, assessment and awarding certificates, amongst others).
- Jurisdictions can change their orientations from centralised to de-centralised (or vice versa) over time.
- Although it inevitably conflates several dimensions (and thus potentially under- or even misrepresents the data upon which it is based), it is possible to construct and justify a continuum showing centralisation and de-centralisation as they are used to defend the public interest in examinations and qualifications.
- England would be near the centralised end of such a continuum, but not absolutely at the extreme.

6.3 Vesting of exams and qualifications powers

The background section of this report defined regulation as a facet of governance. Specifically, regulation was that tranche of activity that related to steering (and/or supervising) events, rather than providing or distributing (see p. 17, above). Implicit in this definition is a comment about institutional structures. That is, by saying that regulation is about steering and not providing, one assumes that more than one organisation delivers (some aspect of) the public service concerned¹⁰. This supposition was confirmed in the background section, when it was pointed out that, following privatisations in the 1980s, technical organisations had been set up at 'arm's length' from government, to regulate private sector providers (p. 20, above).

Before commencing this research, it was known that England had a set up for delivering and regulating examinations and qualifications for 18-year-olds that could (in part¹¹) be described by the definitions referred to in the previous paragraph. There is an education department, and a stratum of independent awarding bodies. Between these two layers sits an

¹⁰ In this model of privatised public service provision, government nonetheless retains important roles. Whilst government may not directly provide a public service in the privatised model, it may well retain political responsibility for the quality of the service. Similarly, government may provide for a sanctions regime in the case that a privatised utility provider does not deliver an aspect of a public service effectively. However, the power to impose sanctions may be vested in a regulator acting at arm's length from political control.

¹¹ There was no privatisation of examinations provision in the 1980s.

independent regulator (see p. 6, above). However, the researcher believed, before starting this work, that this division between provision and regulation was not employed by other jurisdictions for examinations and qualifications (see p. 7, above).

So, the preceding paragraphs suggest both encouragement and discouragement to study the nature of organisations involved with examinations services in international jurisdictions. The importance of the separation between providers and regulators suggests that this issue should be studied, whereas, the pre-supposition that independent regulators will not be found suggests that this is not a fruitful avenue of investigation. This dilemma can be resolved by reference back to the justification for doing this international research project (above, at p. 7). Reference will be made to international organisational structures under a fairly broad definition – it will not be necessary for an organisation to have the word 'regulator' on its brass name plaque for it to be considered relevant to this research. Rather, the organisation findings will be considered with respect to the extent to which powers relating to examinations and qualifications are vested with governments or with other bodies. This location of powers at the central or non-central level can have two types of history. For some jurisdictions powers would have been initially exercised from the centre, only to be subsequently divested from government. In other jurisdictions powers would never have been accrued by government in the first place.

This variety of locations of powers in relation to exams and qualifications will be shown insofar as power is vested with three types of bodies: schools, regions and 'arm's-length technical bodies'. The sub-sections on schools and regions to some extent repeat findings from the previous section, and thus will be kept very brief, whereas the sub-section concerning 'arm's-length technical bodies' will be the most substantial element of this part of the report.

Finally, in the introduction to this sub-section, it is recalled that the previous 'Facets of the examinations and qualifications process' sub-section (pp. 25ff) suggested that there were many stages in an examinations and qualifications process, including: defining assessment standards, assessment itself and awarding certificates, amongst other things. This section will refer to all those phases of the examinations and qualifications process, although – so as not to repeat substantial chunks of the previous section – not exhaustively.

6.3.1 Vesting of powers with schools

There are several ways, internationally, in which powers relating to examinations and qualifications are vested with schools. The principal involvement of schools is in assessment; in 15 of the 16 jurisdictions in this study (including *England*) there was some form of teacher assessment. The sole exception was *Ireland*.

However, schools' involvement was not limited to assessment. In some jurisdictions schools had responsibilities for defining standards (see p. 27, above), also in some jurisdictions schools were involved in the awarding of certificates – either as a counter-signatory or as the sole signatory (see p. 33, above).

6.3.2 Vesting of powers in regions

Powers are vested in regions in various jurisdictions in this study, with respect to various stages in the examinations process and in various ways. In some cases regions hold the majority of power with respect to the education system. *Switzerland* is the prime example of this phenomenon, and expresses its approach as follows:

There is no Ministry for Education at the national level. The principal responsibility for education lies with the cantons.

The overarching characteristics of the educational and political system are:

- its federalism – the sovereignty of the cantons
- its decentralisation – the importance of the responsibilities of the cantons and municipalities
- the subsidiarity of state/national measures; that is, the principle that superior levels of authority, such as the Confederation or the cantons, can only pass regulations or similar in situations where any subordinate levels of authority are not in a position to do so
- its semi-direct democracy – national referenda, ballot initiatives, and so on.

In other jurisdictions, regional authorities may interact either directly with the national education ministry (*Spain*) or with a national technical agency that provides services at arm's length from the government (*Poland*). It should also be remembered that *Queensland* is a state within the nation of Australia and that *England* is a constituent country within the United Kingdom. Thus, when talking about these jurisdictions one is talking about the context in which responsibility for education policy has already been devolved to states or constituent nations¹².

6.3.3 Vesting of powers with arm's-length technical bodies

A particular method for central governments to delegate their powers with respect to examinations and qualifications is to constitute one or more technical agencies with responsibilities for examinations and qualifications. Typically, such agencies are national bodies, which are publicly funded (and which in some jurisdictions are staffed by civil

¹² The situation in the UK is complex. Education policy for *England* is set by the *United Kingdom* government, there is no government of England as such. Responsibility for education policy in the other constituent nations of the UK (Scotland, Wales and Northern Ireland) has been devolved to those nations (see Leeke et al, 2003, pp. 11ff).

servants – see: SEC (2005, p. 6) on *Ireland*). However, the agencies are also typically constituted to be independent from central government, and thus from political interference.

Examples of such agencies can be found in several of the jurisdictions in this sample:

- the National Examinations and Qualifications Centre (NEQC) in *Estonia*
- the State Examinations Commission (SEC) in *Ireland*
- the Centre for Curriculum Development and Examinations (CCDE) in *Latvia*
- CEVO – the Central Examinations Commission and CITO – the National Institute for Educational Measurement In *the Netherlands*
- the *Queensland Studies Authority (QSA)*
- the National Examinations Committee and the National Councils of Experts in *Slovenia*
- *Skolverket* – the Swedish National Agency for Education

The view that such arm's-length technical bodies are a flourishing phenomenon is backed up by observation of several bodies in jurisdictions that did not provide questionnaire responses in the current research, including:

- the Education Quality and Accountability Office (EQAO) in Ontario, Canada (EQAO, 2005)
- the New Zealand Qualifications Authority (NZQA) (Crooks, 2002)
- the Scottish Qualifications Authority (SQA) (SQA, undated)
- The Victoria (Australia) Curriculum and Assessment Authority (VCAA, undated)

Also, in the United States, the SAT and AP examinations (see above at p. 25) are owned by the College Board, which works with the Educational Testing Service (ETS) to develop and administer the programs, and to provide a wide range of associated service. Both the College Board and ETS are private, not-for-profit organisations. There is not regulation of these two organisations by federal or state governments in the US, in a way that an English observer would understand as regulation (see Bennett, 2007).

The arm's-length agencies whose jurisdictions responded to the current research's questionnaire typically provide a complete service relating to examinations (including provision (for example test development) and distribution elements (for example of exam papers)). Also, generally, although not exclusively, there is a single arm's-length agency providing examinations and qualifications services for a jurisdiction. These two facts suggest that none of these agencies are directly analogous to the English regulator of examinations and qualifications.

Perhaps the jurisdiction that comes closest to the English separation between regulation and delivery is *the Netherlands*. Alberts (2001, p. 355) describes the respective roles of the two agencies CEVO and CITO as follows:

The national exams are constructed by CITO by order of CEVO. CEVO is appointed by the Ministry of Education, Culture and Science. The secretary is a seconded officer from that Ministry. The chairperson is directly appointed by the Ministry; the committee members for the different subjects are appointed by the minister but nominated by educational stakeholders, such as universities, association of headmasters, trade unions, subject associations, etc.

The Minister is directly responsible for the examinations. He may even be interrogated on them by the parliament. However, he has delegated the execution of this responsibility to CEVO. It is CEVO's task to make sure that exams meet the requirements laid down in the law, by means of attainment targets and regulations for administration and marking. CEVO exercises its task by preparing specifications for the content of exams, to be used by the contractor, which is CITO. CEVO communicates these to the field by way of syllabus specifications, it is CEVO that checks and ratifies them, gives the green light for printing and is legally responsible for the content.

Despite the differences (from the English set-up) in the way that these agencies are constituted overseas, several of them carry out tasks that in England would be carried out by the regulator of qualifications and examinations.

- In *Estonia*, the NEQC issues subject specifications for examinations subjects.
- In *Ireland*, the SEC's mission is expressed as follows:

The SEC's ... mission is to provide a high quality state examination and assessment system incorporating the highest standards of openness, fairness and accountability.

This mission has much in common with these (slightly selectively chosen) elements of the QCA's regulatory mission as expressed in its Key Result Area (KRA):

QCA regulates ... to ensure that ... qualifications are fair, standards are secure, public confidence is sustained and that QCA acts as the public champion of the learner. (Edited version of full quote, above at p. 17)

Amongst its many functions, the SEC:

- accredits secondary-level examinations in Ireland
- works with Independent Appeals Scrutineers in a manner that appears similar to the way in which the QCA has worked with the independent Examinations Appeals Board (SEC, 2005, p. 10; EAB, 2008, p. 1).

- In *Latvia*, the CCDE's responsibilities include:
 - Develop[ing] educational standards and teaching and learning programs models (CCDE, undated)
- In *The Netherlands*, CEVO and CITO's functions include those described in the citation from Alberts above.
- In *Queensland*, the QSA:
 - develops and/or purchases and approves syllabuses for Years 1-12 and preparatory guidelines
 - accredits syllabuses for Years 1-12 and preparatory guidelines developed by entities other than the QSA for implementation in schools.
- In *Slovenia* several agencies conduct tasks that in England would be conducted by a regulator. The National Councils of Experts are responsible for specifying syllabuses and curriculums, and subject and examination knowledge standards. They also review standards in the curriculums/programmes provided by private schools and formulate initiatives and proposals for the minister of education. National and Subject Examination Committees also monitor development, develop guidelines, and formulate various initiatives and proposals.
- In *Sweden*, *Skolverket* has developed the relatively loosely-drafted 'content strands'. These have a similar function to 'content standards' or 'assessment criteria', but they are deliberately drafted to allow teachers considerable latitude in assessment judgements, to reflect the post-1994-reformed Swedish system's principle of subsidiarity in assessment matters to municipalities and individual teachers (Ramstedt, 2005, p. 15).

Skolverket also conducts a range of research into assessment standards. In particular, it looks at how grades from national tests and final grades made by teacher assessment match up, and seeks to make statements about how fair and reliable the national grading system is (Ramstedt, 2005, p. 28). Although there are many contextual differences between teacher-assessed *Sweden* and examinations-driven *England*, it can be argued that this work has much in common with work led by the English regulator into the comparability of examinations standards (see Newton et al, 2007).

It is instructive at this point to relate the findings regarding 'arm's-length technical bodies' back to the earlier findings from the 'Facets of the examinations and qualifications

process' sub-section (pp. 25ff). In particular, that sub-section developed the notion that jurisdictions could ensure public confidence in qualifications and examinations by using (often a combination of) two trends: centralisation (which could also be called 'standardisation') and de-centralisation (or 'localisation'). On the face of it, one might suppose that setting up an agency at arm's length from central government was a de-centralised way of ensuring public confidence. However, it is suggested that such an interpretation is not the best for current purposes.

Seven jurisdictions are listed above as having arm's-length agencies. Of these, five were listed in **Table 2** as mainly using external exams (the centralising tendency), whereas only two mainly used de-centralising teacher assessment. One should not make a simple equation 'arm's-length body = external exams'; after all, the two jurisdictions that both have arm's-length agencies and mainly use teacher assessment (*Queensland* and *Sweden*) have robust approaches to ensuring the primacy of teachers in assessment. Further, several of the 'mainly externally assessed' jurisdictions also contain substantial elements of local engagement (for example *the Netherlands*, which has internal examination components and internal marking of externally developed exams). Nonetheless, one can take a general trend that the constitution and use of arm's-length technical agencies as a method of organising examinations provision seems to be associated with the preference for standardisation as the preferred method for assuring the public about examinations.

6.3.4 Summary of this sub-section

The following bullet points present a summary of the findings from the current sub-section.

- Definitions of regulation as 'governments steering events, rather than providing or distributing' make an implicit statement about the types of organisations that provide public services.
- This part of the research findings looked at ways in which (parts of) responsibilities and powers relating to examinations and qualifications were vested with government and/or other organisations.
- It was found that various responsibilities and powers were vested with schools and regional authorities.
- It was also found that seven jurisdictions in the study had set up national bodies performing technical tasks relating to exams and qualifications which were independent from (or 'at arm's length from') political control.
- Such independent bodies had substantial provision and distribution functions and so were not regulators, but each of them carried out one or more tasks that would be carried out by the regulator in England.

- In some ways, setting up an independent body to carry out exams tasks is a de-centralising measure. However, five of the seven jurisdictions with such arm's-length organisations also mainly used external assessment – a centralising tendency.

6.4 Aims for regulation

The issues in this part of the data collection are the most complex and embedded to be addressed. Their complexity and highly contextual nature make them particularly suitable to be explored in a longer, more discursive research exercise, such as the focus group that is planned to follow the current exercise. As such, the findings expressed in this part of the report are put forward cautiously – mainly being pointers for further thinking.

6.4.1 Background to aims

Question three of the questionnaire asked:

What does your country aim to achieve through the regulation of examinations and qualifications?

This question is complex: having at least two elements. A point that can be made clearer by reference to the exemplification that accompanied the question:

*In **England**, QCA aims to ensure that:*

- *standards for qualifications and assessments are maintained*
- *the qualifications market is fit for purpose*
- *public confidence in qualifications and assessment is improved and maintained*
- *customer concerns and enquiries are dealt with in a fair, timely, helpful and responsive manner.*

It aims to put into practice the following five principles:

- *proportionality (interventions are related to risk)*
- *accountability (the public has a right to see what QCA does)*
- *consistency (in judgements made; in data requested; in criteria used)*
- *targeting (measures taken related to purpose)*
- *transparency (open and visible).*

The two elements of the question can be re-phrased as: 'what aims does your jurisdiction seek to achieve by having a regulated examinations and qualifications system?' and, additionally, (and the English exemplification makes this plain) 'what amounts to good regulation in your jurisdiction?' or 'what is your jurisdiction trying to do when it regulates its exams and qualifications system?'

The English answer to the second meaning within this question derives from a UK approach to making regulation more effective and intelligent across government sectors, not simply education (BRC, 2006).

This nuance in the question (it being about regulation *per se*, not about the exams system as such) may not have been sufficiently clear for some of the respondents, as five jurisdictions (*Japan, Poland, Slovenia, Sweden and Switzerland*) gave answers that essentially related to the aims of their examinations and qualifications systems.¹³ Although one would not have expected any jurisdictions to respond in terms of post-privatisation of public utilities regulatory theory¹⁴, it was open to jurisdictions to say what aims – for example – an education ministry had in controlling the exams and qualifications system.

The aims for regulation that follow from the first meaning of the question need not necessarily be purely educational aims; for example, aims of examination systems may include political ends such as integrating different regions of a previously diverse nation state or economic imperatives such as ensuring that a jurisdiction's economy can benefit from a qualified workforce. Indeed, it is arguable that there is no such thing as 'educational aims' that are independent of political, economic and wider social concerns.

It is now proposed to look at some definitions of aims and related concepts from the literature, which touch on some of the issues set out above. Le Métais (1997) defined 'aims' as a concept within an international study. For Le Métais, aims are related to values, where the latter are 'beliefs which need not rely upon facts or evidence' (1997, p. 1). Le Métais defines educational aims as follows:

[They] may also be expressed as objectives, goals and targets. Educational aims may be intrinsic (for example to contribute to lifelong education, to develop knowledge, skills and understanding for the individual and society) or instrumental (preparing young people for work and contributing to the national economy). (1997, pp. 8 – 9)

Newton has written extensively about the purposes that educational assessment systems address (for example Newton, 2007). Newton argues that the purposes that educational assessment address have often been unhelpfully imprecisely expressed (*ibid.* at p. 159). He

¹³ Greece's answer was a hybrid; partly relating to the exams system and partly to its regulation, whilst Ireland's answer pertained only to the vocational qualifications system.

¹⁴ In fact, a better source of evidence on the 'principles of good regulation' in a qualifications context is Higher Education research. For example, Blackmur (2006) evaluates the approach taken by the South African Council on Higher Education (CHE) when it withdrew accreditation of 15 Master of Business Administration (MBA) degree qualifications (which were provided by overseas awarding bodies). In a stinging critique, Blackmur states that the CHE's action did not meet the principles of best practice in regulation because it did not consult with affected organisations (*ibid.* at p. 84), did not investigate a non-regulatory alternative to the regulatory action (*ibid.*), operated a 'command and control', 'one-size-fits-all' approach (*ibid.* at p. 85) and lacked consistency, credibility and clarity (pp. 86ff).

also argues that the imprecise expression of assessment purposes has been associated with an unhelpful accumulation of very many (often mutually contradictory) purposes for some high-stakes assessment systems – so much so that some such systems end up not being sufficiently fit for any of their purposes (*ibid.* at p. 167).

Newton concludes by suggesting (*ibid.* at p. 168) that specific assessment purposes ought to give rise to specific assessment designs. For example a selection examination may be designed to separate candidates reliably with respect to the selection criterion – making an examination with highly discriminating items a likely candidate for such a purpose. This is in contrast to an assessment system designed to support a formative purpose – presumably assessment procedures that are compatible with classroom practice, yield rich information to suggest future learning trajectories, and so on would be best for that purpose.

This proposed association between purpose (or aim in the terms of the question) and the design of the system heralds a profitable way in which to organise the findings for the rest of this sub-section. Firstly, jurisdictions' description of the 'virtues' they aim for in their assessment procedures and associated score¹⁵ data are organised and presented. Then, the broader 'systemic' aims that jurisdictions profess for their examinations and qualifications systems are similarly treated. Thirdly, comparative and contrastive analysis of the international jurisdictions shows instances of how jurisdictions use similar or different assessment designs to fulfil the broader aims that they have for their examination systems.

6.4.2 Properties of assessment designs and score data

The properties of assessment designs and score data that respondents report as being aims for their jurisdictions are shown in the following table:

¹⁵ 'Score' is broadly intended here – it can mean the properties of grades achieved through a teacher assessment system, as much as the properties of examination scores.

Properties of assessment designs and score data	Jurisdiction
Assessment procedures are uniformly applied across the jurisdiction.	Denmark Spain
Assessment standards are maintained.	Estonia Korea Slovenia Turkey
Jurisdiction uses a variety of assessment forms and techniques.	Greece
Assessment (and/or certificates) is/are valid.	Greece Spain
Assessment is reliable.	Greece
Assessment is objective.	Greece Latvia
Assessment is accurate.	Greece
Assessment grades/score data enable comparability.	The Netherlands
The jurisdiction has concrete goals and grading criteria.	Sweden
The data from the assessment system facilitate the provision of profiles of individual learners.	Sweden
There is equity and fairness in grading.	Sweden

Table 3: Properties of assessment designs and score data that jurisdictions expressed as aims for regulation

The limitations of this part of the data collection have already been noted, but specific issues should be noted with respect to this table. Firstly, this table does not represent a comprehensive survey of the aims that jurisdictions have for their assessment designs and data. Jurisdictions expressed their aims in a constructed response item. If they had been presented with a comprehensive list of possible aims in a selected response item, they might have chosen more (or differently). Further, issues about the variable ways in which respondents answered this question have been discussed at length above. Finally, it may be that some of the expressed aims implicitly include others – for instance 'equity and fairness' may be a superordinate category of 'accurate'. Despite these weaknesses, it is suggested that this table can be used cautiously – both as the basis for the subsequent comparative analysis, and as a starting point for considering issues in subsequent phases of this research work.

6.4.3 System-level aims for exams and qualifications systems

The aims that jurisdictions allege their examinations and qualifications systems promote are presented in the following table:

System-level aims for exams and qualifications	Jurisdiction
Education provision is common for all students, regardless of the Autonomous Community within which they are educated.	Spain
The examinations and qualifications system support the educational aim of providing education to every citizen.	Korea
Higher education institutions can use results for their admissions.	Estonia Latvia Queensland Slovenia
Standards are maintained so as to assure public confidence.	Queensland
There is ongoing analysis of results.	Latvia
The exams and qualifications system facilitates ongoing monitoring and analysis of results.	Latvia
The exams and qualifications system facilitates evaluation of the overall position of the education system through publishing information.	The Netherlands Poland
The exams and qualifications system provides statistics for local and national analysis of educational achievement.	Sweden
The examinations and qualifications system contributes to the education system's aim of fostering students' self-knowledge.	Greece
Data from the exams and qualifications system provides teachers feedback on their teaching.	Greece Sweden
The design of the exams and qualifications system allows schools to facilitate quality learning outcomes.	Queensland
The operation of the examinations and qualifications system enhances achievement.	Sweden

Table 4: System-level aims for the regulation of exams and qualifications

The caveats expressed above about Table 3 apply equally to Table 4. Also, the aims in Table 4 could have been arranged in cognate groups so as to give a more coherent view of the data. This has not been done, so as to maintain a reasonably close match to the responses that the jurisdictions themselves gave (although they have been reworded as minimally as was compatible with maintaining ease of reading in the table). Despite these reservations, as with Table 3, it is submitted that these data are adequate to support the analyses that follow.

6.4.4 Comparative analysis of aims for regulation

As described above, this part of the sub-section presents some contrasts and comparisons with respect to aims for regulation. These comparisons and contrasts are of the following types:

- dissimilar jurisdictions that express similar aims for regulation
- jurisdictions that express similar aims and yet adopt differing assessment procedures

- jurisdictions that despite having a similar spread of aims, ostensibly place different emphases on particular aims

These three types of comparison and contrast will be addressed in turn.

Firstly, one may look at the two jurisdictions that, in **Table 3**, have been summarised as aiming for the following assessment design property:

Assessment procedures are uniformly applied across the jurisdiction.

Denmark is a northern European EU Member State, with a unified, non-federal governmental structure (see: Egelund, 2005). It expressed its aim of uniformity of assessment procedures in the following words in the questionnaire:

For the end of ... upper secondary education examinations, the Ministry of Education has established standard rules and regulations, with a view to ensuring uniformity throughout the country. For the same reason, the written examination questions are set and marked at central level.

Spain is another EU member state, yet it has a very different history to that of *Denmark*. In *Spain*, responsibility for education is constitutionally devolved to the autonomous communities (Le Métais, 1997, p. 1). Yet the aims that *Spain* expressed in relation to uniformity of assessment procedures bear comparison with those of *Denmark*:

By regulating the evaluation/assessment criteria and the certificates, the Ministry of Education aims to:

- ensure fundamental/basic common educational provision for all students regardless of the Autonomous Communities in which they live/attend school
- ...
- homogenise the assessment system by establishing some common guidelines for all Autonomous Communities.

This line of analysis can be taken further. *Denmark* and *the Netherlands* are – on the face of it – fairly similar jurisdictions (northern European, non-federal EU member states) with quite comparable examination systems used at the end of compulsory schooling. Yet, *Denmark* states in its questionnaire that examinations are centrally marked – in order to ensure that assessment procedures are uniform across the nation. In contrast, in *the Netherlands*, even the centrally-set exams are administered and marked in schools.

It is not the purpose of this sub-section to launch into discussion. However, it is apposite at this point to tie up some strands. Earlier in this paper, it has been advanced that standardisation versus localisation is a convenient metaphor with which to understand how jurisdictions ensure that the public has confidence in examinations. The examples above of 'nationwide homogeneity' and the differences in approaches to the centralisation of marking demonstrate possible routes into understanding what it means (in an international context) to ensure public confidence in examinations and qualifications.

Next, one can look at jurisdictions that express similar aims and yet adopt differing assessment procedures. The four jurisdictions that – in their questionnaire – stated that they had 'the maintenance of examination standards' as an aim were: *Estonia, Korea, Slovenia* and *Turkey*. Despite expressing this same aim in their questionnaire, two of these jurisdictions 'mainly favoured internal assessment' and two of them 'mainly favoured external assessment'. Exploring the reasons behind this same aim/different assessment approach situation might once again afford insight into what it means to reassure the public about examinations and qualifications. It might also challenge an implicit conclusion from Newton's work (summarised above at p. 46) that there were particular assessment designs that would be logically associated with particular assessment purposes.

This line of thinking may be taken forward by referring to the last type of analysis listed above; that is, 'jurisdictions that despite having a similar spread of aims, ostensibly place different emphases on particular aims'. Good examples of this bullet point are jurisdictions that favour internal assessment for their school leaving qualifications, such as (amongst others) *Queensland*. *Queensland's* aims for its school-leaving qualifications system include providing results that HE institutions can use for admissions¹⁶. Similarly, it states that it has system-level aims to assure standards. Despite this, it is still able to design a qualifications system with the aim of facilitating quality learning outcomes, and indeed, in general to run a fairly localised, teacher-centric approach to assessing school-leaving qualifications. As such, the further research planned might seek to confirm or refute the proposition that the choice of assessment method, and the balance between centralisation and de-centralisation as the main approach to ensuring public confidence stems from a variety of sources, rather than being a context-free 'logical' successor to the chosen assessment purpose.

¹⁶ The QSA produces a ranking of persons (the Tertiary Entrance Statement) to aid this function of its qualifications system.)

6.4.5 Summary of this sub-section

The following bullet points present a summary of the findings from the current sub-section.

- The issues in this part of the findings were the most complex to be addressed in the research. As such, the interpretation of this set of findings should be limited to the pointers they provide for further research.
- The sub-section related to the aims jurisdictions had for regulation, not to their aims for examinations and qualifications systems per se.
- The notion of aims is complex, and is related to other concepts such as 'values' and 'purposes'.
- Aims were expressed at two levels of generality: the properties of assessment designs and the system-level aims for exams and qualifications.
- Using findings regarding those levels of generality it was possible to do comparative and contrastive analyses of jurisdictions' aims for their systems.
- Such analyses reinforced the view expressed earlier in this report that understanding how jurisdictions balance standardising and localising measures to ensure public confidence is a useful way to understand the regulation of qualifications and examinations in an international context. However, the comparative analyses challenged the view from assessment literature that it was possible to incontrovertibly (and in a de-contextualised manner) associate particular aims (purposes) with particular assessment designs.

6.5 Use and regulation of e-assessment

The background section to this report suggested that new technologies would have profound implications; both for the practice of examinations and for its regulation (see above, at p. 22). Therefore, the questionnaire asked jurisdictions to describe their use of e-assessment, and – if they were using that assessment method – whether and how they might regulate it.

It should be recalled that this section, like the entire document, reports jurisdictions responses with respect to '18-plus', general qualifications for post-compulsory school leaving. It is possible that jurisdictions would be using e-assessment in national assessment programmes taken throughout schooling, low-stakes testing programmes, professional licensure examinations, and so on, but still answer that they were not using e-assessment to this questionnaire item.

6.5.1 Use of e-assessment

The questionnaire asked about the extent to which e-assessment was being used in each jurisdiction. The responses to this item in the questionnaire are collected in the following table:

Use of e-assessment	Jurisdiction
Not used	Japan Korea Spain Estonia Greece Poland Slovenia Switzerland
Initial, limited use	Sweden Denmark Netherlands ¹⁷
Used/available for vocational qualifications only	Sweden Ireland
Used for examinations processing (for example developing questions, marking, analysing item responses, reporting), but not for conducting tests	Turkey Latvia
As a support for internal assessment (for example via use of e-portfolios, digital recording of performances, email and web communication)	Queensland
Hybrid model used (for example questions on paper, responses captured either from paper or from on-screen entry)	Netherlands
Used to support moderation	Queensland
Used to register students for learning accounts	Queensland
Word processors allowed in examinations	Netherlands

Table 5: Jurisdictions' use of e-assessment

Thus, the majority of jurisdictions in this study (8 from 15) appear to not be using e-assessment yet. The two jurisdictions which, on the information from the questionnaires, appear to be using e-assessment most are *Queensland* and *the Netherlands*. These two jurisdictions have quite different approaches to assessment; the former emphasising internal, teacher assessment and the latter external exams¹⁸. These emphases carry over to the developments that both jurisdictions have made in e-assessment. The following are extracts from the two jurisdictions' questionnaire responses:

In **Queensland's** system of externally-moderated school-based assessment, e-assessment is used extensively:

- Students' portfolios frequently contain pieces of work done in electronic media.
- Some schools use intranets to conduct on-line assessment.
- Teachers record practical performances electronically.
- Submissions of student portfolios to moderation panels frequently contain CDs and DVDs as either assessment pieces or records of performance.
- Communication with moderation panels is often done via email and the web.

¹⁷ 'Limited use' when referring to an examination conducted entirely via an electronic medium. Wider use of the Dutch 'hybrid model' (see below).

¹⁸ Although with substantial local involvement.

- All samples of student work used to demonstrate standards are on the Queensland Study Authority's (QSA's) website.
- Moderation submissions are tracked electronically.
- Students in Year 10 or in the year before they turn 16 (whichever comes first) have to be registered with the QSA. This registration automatically opens students' web-based learning accounts, which are used to record what, where and when the different types of learning that attract credit for the QCE take place.
- Students' tertiary entrance results are made available to them by SMS (text messaging) and on the web.
- The Queensland Core Skills (QCS) Test (for tertiary entry) is developed electronically and in hard copy and the item bank is both electronic and hard copy. QCS Test markers are monitored electronically. The multiple-choice papers of the QCS Test are marked electronically. Aggregated results are reported on the QSA's website.

The Netherlands

Computers are used in various ways in the examinations which take place in general secondary education:

General/combined

Examination questions are set on paper and computers are used to calculate the responses/mark the responses either from answers provided on paper or filled in on computer. In some cases special applications have been developed which pupils can access and complete. These methods are currently in use in two of the learning pathways in pre-vocational secondary education (the theoretical and combined learning pathways of the VMBO for biology, economics and physics and chemistry); and in the two other forms of secondary education (HAVO for geography, biology, economics and physics; and VWO for biology, physics and mathematics). The use of word processors is allowed in all examinations.

Monitoring examinations

In this form of central examination the complete examination is on computer. Questions are on-screen and the pupil answers these at the computer. Essay answers are assessed by the teacher; the computer scores yes or no/multiple choice questions. All the examinations in this category are *in a pilot phase*¹⁹. These include, in pre-vocational secondary education (VMBO), examinations in French, German and general subjects; general senior secondary education (HAVO) examinations in culture and the arts, and German; and pre-university (VWO) qualifications in culture and the arts.

Central written and practical examinations (CSPE)

Practical examinations in pre-vocational secondary education (VMBO) include IT-based questions. Pupils perform certain actions and describe the results/or print them. Some parts of CSPE exams are also taken completely on computer.

Culture and the arts

Computers are used to enable a rich visual presentation of examination questions. Currently answers have to be provided on paper.

¹⁹ Emphasis added by current researcher.

There are very many academic articles describing e-assessment developments around the world. Many of these describe either 'low-stakes' assessment initiative (for example the Assessment Resource Bank (ARB) in *New Zealand* – see: Croft (2002)) and/or pilot or trial projects (such as: PASS-IT in *Scotland* – see: Ashton & Thomas (2006)). There are far fewer large-scale, high-stakes, well-established e-assessment programmes.

The USA is perhaps the spiritual home of 'computerized testing'. Several American assessment programs used electronic delivery for high-stakes testing programs. These include ETS's products: the Test of English as a Foreign Language (ToEFL) (ETS, undated-a) and Graduate Record Examinations (GRE) (ETS, undated-b); both high-stakes assessments taken by large numbers of students.

6.5.2 Regulation of e-assessment

Given the relatively light uptake of e-assessment, it is not surprising that there is not much evidence of jurisdictions regulating e-assessment. Twelve jurisdictions said that they did not regulate e-assessment. Two jurisdictions answered that they did have regulations concerning e-assessment, and one said 'yes and no'.

The two jurisdictions that did regulate for e-assessment were *Queensland* and *Latvia*. *Queensland* replied as follows:

In accordance with the Queensland Government's Information Privacy Standard IS42, learning accounts are password protected and only available to the student and his/her parents or carers.

Latvia described their plans as follows:

The document 'Basic Standpoints in Education Development 2007-2013' (adopted by the Cabinet of Ministers in September 2006) requires that a bank of state examination tasks is developed.

In addition, the Ministry of Education and Science has commissioned the development of an 'Information System for State Examinations'. This system will be tailored to the 'e-organisation' of examinations. The system will support the electronic registration of pupils for examinations; it will deliver data to the Centre for Curriculum Development and Examinations; and will ensure data control and the control of proper and legitimate procedures for examinations.

The Netherlands replied 'yes and no' to question six. Their response said:

Regulation of e-assessment will be included in the examination regulations, along with the technical conditions required to make e-assessment possible.

For some subjects e-assessment will be optional; for others it will be the only way examinations are presented.

Ireland stated that it would not regulate specifically for e-assessment as a separate assessment method. However, the Further Education and Training Awards Council (FETAC) *Guidelines for providers* [of vocational assessment] refer to e-assessment (FETAC, 2007, p. 35). This guidance emphasises that assessment providers must quality assure e-assessments so that they conform to the values of validity and reliability expressed elsewhere in the *Guidelines*. They also emphasise specific issues regarding e-assessment, such as ensuring that test takers have the necessary ICT skills (and familiarity) to interact with the test interface, and guarding against plagiarism when assessment work is submitted electronically.

Whilst there was little evidence from the current international survey of regulatory documents addressed specifically at e-assessment, there are some examples from English-speaking countries of documents embodying quality standards, as well as guidance documents for centres on implementing e-assessment. Such documents originated from a range of sources, such as: professional associations, cross-national organisations and quasi-governmental, but not specifically educational bodies. Examples of such documents include:

- International Standard ISO23988/British Standard 7988: *International code of practice for the use of information technology (IT) in the delivery of assessments* (ISO/BSI, 2007).
- International Testing Commission (ITC) *International Guidelines on Computer-Based and Internet-Delivered Testing* (ITC, 2006)
- Association of Test Publishers (ATP) *Guidelines for Computer-Based Testing*. (ATP, 2002)

These documents vary along a number of dimensions:

- the scope of the document (for example does it apply to the whole qualifications and examinations cycle, or just a specific part thereof, such as the conduct of an exam on screen?)
- the audience for (and/or objects of) the standards (for example awarding bodies, test users, test takers, and so on)
- the enforceability of the standards (those quoted above are currently voluntary codes)
- the wording of the document – especially, do the documents use the word ‘must’ or ‘should’? The ITC Guidelines are interesting, in that they use no ‘modal verb’ (such as ‘must’ or ‘should’).
- the granularity of detail in the standard
- the relationship of the e-assessment standards document to other sets of standards (for example the ATP CBT standards document styles itself explicitly as an adjunct to the

1999 AERA/NCME/APA *Standards for Educational and Psychological Testing* (AERA et al, 1999)).

- whether the e-assessment standards document implements a specific vision of educational quality (for example a view of validity).

In addition to standards documents, several guidance documents have been developed.

These include:

- UK regulators' *Guide to effective practice in e-assessment* (QCA et al, 2007c)
- the Scottish Qualifications Authority (SQA)'s *Guidelines on e-assessment for schools* (SQA, 2005) and *Guidelines on online assessment for Further Education* (SQA, 2003)

These documents have generally been aimed at schools, colleges and other learning providers and intend to provide them with information that will help them to run e-assessment sessions.

6.5.3 Summary of this sub-section

The following bullet points present a summary of the findings from the current sub-section.

- For the majority of jurisdictions responding to this questionnaire, in respect of the examination and qualification types considered in this report, there was reported to be either no use of e-assessment (eight jurisdictions) or initial, limited use (three jurisdictions).
- The two jurisdictions in this sample with the most use were *Queensland* and *the Netherlands*. Their use of technology for assessment reflected their overall approaches – *Queensland* used technology to support teacher assessment, whereas *the Netherlands* had various approaches to running e-tests.
- There are many reports in the literature of e-assessment programmes. But typically, these are either of 'low-stakes' assessment applications and/or pilot projects. The exception to this general picture is the USA, which has several large-scale high-stakes e-assessment programmes.
- Two jurisdictions replied to the questionnaire that they had regulations pertaining to e-assessment. *Queensland* said that it had regulations in respect of information privacy, whilst *Latvia* responded by describing its plan to develop a bank of state examination tasks by 2013.
- There were several sets of quasi-regulatory standards or guidelines documents in the English-speaking world. Such documents varied along the following dimensions: scope, audience, enforceability, wording, granularity of detail, relationship with other sets of standards and vision of education quality.
- In addition to such quasi-regulatory documents there were several sets of guidance documents for centres providing e-assessment to candidates.

7 Implications of the findings

7.1 Purpose of this section

It is worth re-iterating what this research has been trying to do. The research questions were:

1. What insights can international findings provide to either support or challenge English approaches to the regulation of examinations and qualifications?
2. To what extent are English approaches to the regulation of qualifications and examinations mirrored by practices in other international jurisdictions?
 2. a. When international approaches to regulation differ from those of England, what insights can the English regulator take from international practice?

Also, the section 'A hypothesis-generating project', above at p. 9, stated that the research would: describe and define; be cumulative; define the limits of the domain; suggest connections between different concepts within the domain; suggest implications from the findings for English regulatory practice and be plausible to a reasonable reader.

It is proposed that this concluding section will give two outputs: a(n intellectual) framework or structure within which to define subsequent phases of this work, and a concluding section of 'low-hanging fruit'; that is, insights that the English regulator can take away **now** to inform policy and practice – rather than waiting for confirmation in later research phases.

7.2 Framework for later phases

The first definition given was of regulation as governments steering events rather than providing or distributing goods and services (above, at p. 17). Using this definition to frame an international research project was always known to be problematic – other jurisdictions simply do not separate regulation and provision as England does. It might therefore seem to follow that one can assert that international jurisdictions do not regulate their exams and qualifications systems. However, such an assertion would be highly counter-intuitive. More sensible is a contention that all jurisdictions **do** endeavour to ensure that the public has confidence in exams and qualifications systems. This form of words seems a better way to define regulation in the international context; both because it is more defensible in simple terms and more profitable in that it allows one to conduct international studies into regulation.

One of the key tools for understanding how international jurisdictions ensure public confidence is their balancing of centralising (standardising) and de-centralising (localising) measures. Kuiper et al's (2007) paper has provided a useful stimulus to adopting the centralisation metaphor in this work. However, further work to examine the notion of (de)-

centralisation as it relates to international education policy would be likely to improve subsequent stages of this research.

Green et al (1999, p. 57) list six strategies that EU jurisdictions have used to divest themselves of responsibilities and powers with respect to education:

- dispersal of powers to 'social partners'
- regional devolution
- regional deconcentration
- localisation
- institutional autonomy
- marketisation

Green et al go on to elucidate the meanings and implications of each of these methods for delegating powers and duties relating to education. It is suggested that relating these categories to the practice of maintaining public confidence in examinations would be a useful way to organise findings in the next phase of the work.

A broader issue may underlie that of (de-)centralisation: globalisation, which is a term used to describe growing interdependence of people around the world. It has social, cultural and economic aspects (Green, 1997, pp. 151ff). Increasingly, the production of goods and services by multi-national corporations is ceasing to be linked to particular nation state, or even to particular regions of the world (*ibid.* at p. 158). An echo of this can be found in Bartram's graphic example of a multi-national application of testing (above, at p. 22). Under such an analysis, nation states have less control of economic activity than previously – having less to gain from the profitability and investment of nationally-registered corporations. Rather, the main capital over which nation states retain control is 'human capital'; in particular, highly prized problem solving, communication and inter-personal skills (Green, 1997, pp. 154). This strand of reasoning has been used by some English commentators as an explicator of government's allegedly increasing micro-management of education (Wolf, 2002, p. xiii). If it is true that education remains one of the few 'levers' over which governments retain control in the globalised world, then this may need to be considered in any analytical model of (de-)centralisation that is used in subsequent phases of this research.

However, the field of globalisation is complex and contested (Green, 1997, p. 168; see also: Dale, 2005), and the impact of globalisation on examinations and qualifications systems may pull in several directions. For example, trends towards within-nation-state regionalism may cause examinations responsibilities to be moved from central authorities. However, there

are also supra-national pressures; these include international comparative studies and cross-national qualifications frameworks (see: Coles, 2007).

The next point that will help to frame the subsequent research phase is that organisations **do** matter. Whether the central authorities in jurisdictions delegate powers to schools, regions or arm's-length technical bodies is likely to affect the ways in which examinations and qualifications systems are implemented, and correspondingly have an impact on how the public views such systems.

The investigation of the arm's-length technical agencies that were identified in this research will be an important task for subsequent phases. Essentially, it will be interesting to postulate whether England is at the head of an international trend and whether, therefore, other jurisdictions will be likely to institute comparable regulatory regimes over time – implementing values such as the 'principles of good regulation' as they are understood in the UK. In framing this part of the subsequent research, it will be useful to study histories of the privatisation of public utilities, both in the UK and internationally, and also including both supportive and critical readings of that political process.

Whilst one analysis is that England is potentially leading an international trend, the other possibility is that England could be following the trend of other jurisdictions. Many of those jurisdictions have a single, often national, technical examinations agency. In this context it is worth noting the trend towards consolidating English awarding bodies (Tattersall, 2007, p. 88).

The next main plank of subsequent research will be to investigate the relationships between purposes and the design of examinations systems. The questionnaire item in this research referred to 'aims' for regulation (see pp. 44ff). This was perhaps not the most appropriate wording. Newton has written extensively and to widespread approval about the purposes to which assessment systems are put. That work has been considered to date as assessment research, rather than as research into the regulation of qualifications and examinations as such. It is suggested that this international research into regulation should appropriate Newton's work as part of its understanding of regulation.

It may do so with several riders. Firstly, in an international study of regulation, one should consider the purposes of all elements of the examinations and qualifications process; how and the extent to which jurisdictions control the defining of examination standards and/or the awarding process is relevant to an understanding as well as whether the jurisdiction uses external examinations or teacher assessment. Secondly, when one is talking about

regulation, one ought to think about purposes as being mainly centralising or mainly de-centralising. This is a way of linking major strands in the understanding of the international findings.

The final way in which purposes can be investigated dovetails nicely with comparative and contrasting analysis. As at page 48 above, it is possible to set up archetypal comparative situations and use them to try to understand why jurisdictions' examinations and qualifications systems are organised as they are.

A good illustration of such analysis would be to compare *Queensland* and *England*. In both these jurisdictions a school-leaving qualifications system provides data with which to admit young people to university. And yet, *Queensland* jealously guards its teacher assessment system whereas *England* employs a largely centralised examination system. The study of international approaches to the regulation of examinations and qualifications should seek to account for that apparent dissonance between similar purposes and different assessment designs. Is the difference best explained by the presence or absence of other purposes (for example the purpose to hold teachers and schools accountable in England)? Or is it that other differences (for example in values, history, and so on) better explain the difference in assessment design than the purposes that those designs address? And, finally, since the research must maintain a focus on regulation rather than being assessment research *per se*, how does the interaction between purposes, values, history and so forth re-assure the public about the examination and qualifications systems in different jurisdictions?

The final aspect to be considered in subsequent research will be the future. This research returned relatively little evidence of jurisdictions using and regulating e-assessment for their school-leaving qualifications. However, this matter should be investigated further in follow-up work. It is suggested that this simple table will provide a useful framework for thinking about the future of regulation of qualifications:

Source of change	Nature of change
New supra-national entities (for example the EU) and practices (for example international assessments)	Nation states as objects, as well as originators of regulation
New technology	Regulators needing to deal with internationalisation/globalisation of exams and qualifications practice Convergence of regulators from different sectors Implementation of new technology in assessment requires researchers to re-theorise what assessment means by quality

Table 6: Issues affecting the future of regulation

Several points are plain from this table. Firstly, it is clear that the future of regulation is predicted to be influenced not just by technology, but also by the emergence of new supra-national entities and practices. Also, the future of regulation will be influenced by insights from regulation literature (for example convergence of sectors) and from assessment theory and practice (for example the fact that new forms of e-assessment will not be merely analogues of pencil-and-paper tests, but that they will be completely new forms of assessment and will require corresponding new approaches to understanding what makes them valid or invalid). Finally, it is worth noting the dissonance between what is predicted by theoretical literature and what was found in the current data collection. It is worth remembering that the majority of jurisdictions reported **not** using e-assessment in their examinations and qualifications for the end of post-compulsory education. Whether that simply means that predictions from the literature will be delayed, or whether the predictions should be discounted in the specific field of high-stakes school-leaving qualifications is a matter that could be further investigated.

7.3 Low-hanging fruit

The preceding sub-section has defined how follow-up work should be framed. The report now concludes by suggesting six key insights that the English regulator can take away from this research **now** – without waiting for the follow-up work:

- Regulation means something in the international context if it is about how jurisdictions help the public to maintain confidence in their exams and qualifications systems. If it is merely about the mutual duties of a quasi-public sector technical body, and private sector providers, then it does not.
- Centralisation and de-centralisation can both help to ensure public confidence in exams and qualifications systems. In England, regulation is often synonymous with centralising measures.

- Centralisation versus de-centralisation could be seen – stereotypically – as a bald choice between external examinations and teacher assessment. In fact, there are many ways in which de-centralising measures can be incorporated into qualifications process – they can be incorporated at many stages of the process: defining standards, assessment itself, awarding qualifications, and so on.
- The international jurisdictions in this sample did not have regulators as such, but several of them did have technical bodies set up at arm's length from government to perform examinations tasks.
- An analysis of examinations purposes, following the work of Paul Newton, should form part of what regulation is about in the international context.
- Challenges to regulatory practice in future will come both from the impact of new technology **and** the fall-out of globalisation in general and specifically from the existence of supra-national political entities (for example the EU) and practices (for example international comparative surveys and credit transfer).

8 Appendices

8.1 Project questionnaire

Qualifications

1. a) What types of general and vocational qualifications are available to students aged 16-18 in your country? 1. b) How are these qualifications assessed?	
<p>1. a) In England, assessment on completion of compulsory education (at age 16) usually involves the General Certificate of Secondary Education (GCSE) examination. Traditionally taken in academic subjects, and increasingly available in a range of applied or vocational subject areas, GCSE qualifications are usually required for access to further study.</p> <p>The main qualifications offered in post-compulsory education in schools are General Certificate of Secondary Education (GCE) Advanced (A) Levels, GCE Advanced Subsidiary qualifications and GCE 'A' Levels in Applied Subjects. Success in these qualifications is usually required for university entrance.</p> <p>A wide range of vocational qualifications is accredited into the National Qualifications Framework (NQF). Vocational qualifications serve a range of purposes in different sectors and at different levels, so they vary in size, level and assessment.</p>	1. a)
<p>1. b) GCSEs are mainly assessed via external, terminal examinations. However, some GCSEs contain coursework elements. Coursework has been replaced – in some cases – by a new form of 'controlled assessment'. Some GCSEs have modular options, with assessment at the end of each unit. An A-level consists of six modules. Modules are assessed by exam papers marked by national organisations (awarding bodies) and internally-assessed coursework. Vocational qualifications are assessed by a range of external and internal methods, such as: portfolios of evidence, observation, written examinations, interviews, and so on.</p>	1. b)

Organisations

<p>2. a) Which organisation(s) are responsible for general and vocational qualifications available to students aged 16-18 in your country?</p> <p>2. b) What is the nature of these organisation(s)?</p> <p>2. c) What tasks do(es) the organisation(s) carry out?</p>	
<p>2. a). In England, the Qualifications and Curriculum Authority (QCA) regulates qualifications offered in schools, colleges and workplaces.</p> <p>Qualifications are developed, administered and awarded by 'awarding bodies'.</p>	2. a)
<p>2. b) QCA is a statutory body, sponsored by the Department for Children, Schools and Families (DCSF).</p> <p>Awarding bodies are independent organisations; some of which are charitable trusts or linked to professional associations, while others are commercial companies.</p>	2. b)
<p>2. c) QCA maintains and develops the national curriculum and subject criteria for qualifications. It also accredits awarding bodies and qualifications and advises the DCSF on these matters.</p> <p>QCA oversees the work of awarding bodies in England (including administration, marking and awarding).</p> <p>Awarding bodies develop qualification specifications in line with subject criteria issued by QCA. They then develop examinations according to those specifications.</p> <p>Awarding bodies are responsible for administering and marking the examinations (or other assessments). Awarding bodies also award qualifications.</p>	2. c)

Aims for regulation

3. What does your country aim to achieve through the regulation of examinations and qualifications?

3. In **England**, QCA aims to ensure that:

- standards for qualifications and assessments are maintained
- the qualifications market is fit for purpose
- public confidence in qualifications and assessment is improved and maintained
- customer concerns and enquiries are dealt with in a fair, timely, helpful and responsive manner.

It aims to put into practice the following five principles:

- proportionality (interventions are related to risk)
- accountability (the public has a right to see what QCA does)
- consistency (in judgements made; in data requested; in criteria used)
- targeting (measures taken related to purpose)
- transparency (open and visible)

3.

Documents

4. a) What documents do you have that describe how examinations and qualifications are regulated?	
4. b) What values or quality standards do regulatory documents require from assessments?	
<i>4. a) In England, QCA has published regulatory criteria, codes of practice and regulatory principles.</i>	4. a)
<i>4. b) Regulatory documents require assessments to be:</i> <ul style="list-style-type: none">• <i>valid</i>• <i>reliable</i>• <i>free from bias</i>• <i>manageable</i>	4. b)

Use of e-assessment

5. To what extent is e-assessment, used in general and vocational qualifications, available to students aged 16-18 in your country?	
<i>5. In England, a small, but growing, number of qualifications use e-assessment. Most qualifications using e-assessment are vocationally related, although there are also a few general qualifications that are e-assessed.</i>	5.

Regulation of e-assessment

6. a) Are there any specific steps (either current or planned) in your country to regulate e-assessment?

6. b) If yes, please describe these steps, including linking to any available documents and summarising the rationale or aims of the steps.

<p><i>In England:</i> 6. a) Yes</p>	6. a)
<p>6. b) <i>In England, QCA has introduced 'regulatory principles for e-assessment'.</i></p> <p><i>The increasing use of technology for the development, delivery, administration, marking and reporting of assessment has lead to the conclusion that regulation is needed to take account of the specific nature of this mode of delivery.</i></p> <p><i>The 'Principles for the regulation of e-assessment' are intended to ensure flexibility and encourage innovation.</i></p>	6. b)

8.2 Bibliography

Alberts, R.V.J. (2001) Equating exams as a prerequisite for maintaining standards: experience with Dutch centralised secondary examinations. *Assessment in Education: Principles, Policy and Practice*, 8(3), 353 – 367.

American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement in Education (NCME) (1999) *Standards for Educational and Psychological Testing*. (Washington, DC: American Educational Research Association).

Ashton, H. & Thomas, R. (2006) Bridging the gap between assessment, learning and teaching, in: M. Danson (Ed) *Tenth international Computer Assisted Assessment (CAA) conference proceedings, Loughborough University, 4th and 5th July 2006*. Online: available: www.caaconference.com.

Association of Test Publishers (ATP) (2002) *Guidelines for computer-based testing*. (Washington, DC: ATP).

Bartram, D. (2006) The internationalization of testing and new models of test delivery on the Internet. *International Journal of Testing*, 6, 121-131.

Bennett, R.E. (2007) *What does it mean to be a non-profit educational measurement organization in the 21st Century?* A paper presented at the International Association for Educational Assessment Annual Conference, Baku, Azerbaijan, September 2007.

Bertram, T. & Pascal, C. (2002) *Early years education: an international perspective*. Online: available: www.inca.org.uk/pdf/early_years.pdf.

Better Regulation Commission (BRC) (2006) *Risk, responsibility and regulation – whose risk is it anyway?* Online: available:

http://archive.cabinetoffice.gov.uk/brc/upload/assets/www.brc.gov.uk/risk_res_reg.pdf.

Black, P. & Wiliam, D. (2005) Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal*, 16(2), 249 – 261.

Blackmur, D. (2006) The South African MBA review and the principles of efficient regulation. *Quality in Higher Education*, 12, (1), 81 – 93.

Braithwaite, J., Coglianese, C. & Levi-Faur, D. (2007) Can regulation and governance make a difference? *Regulation & Governance*, 1(1), 1-7.

Callaghan, J. (1976) *Towards a national debate*. Speech by the prime minister, at a foundation stone-laying ceremony at Ruskin College, Oxford, October 18 1976. Online: available: <http://education.guardian.co.uk/thegreatdebate/story/0,9860,574645,00.html>.

Centre for Curriculum Development and Examinations (CCDE) (undated) *Centre for Curriculum Development and Examinations*. Online: available: <http://isec.gov.lv/en/index.shtml>.

Chalmers, D. & Lodge, M. (2003) *The Open Method of Co-ordination and the European welfare state*, ESRC Centre for Analysis of Risk and Regulation Discussion Paper no. 11, London School of Economics, June. Online: available: www.lse.ac.uk/collections/CARR/pdf/Disspaper11.pdf.

Charlton, J. & Higson, T. (2008) *Reviewing our regulatory arrangements – developing common regulatory requirements*. Unpublished QCA presentation.

Coles, M. (2007) *Qualifications frameworks in Europe: platforms for collaboration, integration and reform*. A paper for the conference: Making the European learning area a reality, 3-5 June 2007, Munich. Online: available: www2.vhs-bayern.de/contentserv/4.2/bvv.de/data/media/_stories/372/Expertise_1.pdf.

College Board (undated) *SAT Reasoning Test*. Online: available: www.ets.org/portal/site/ets/menuitem.c988ba0e5dd572bada20bc47c3921509/?vgnnextoid=178daf5e44df4010VgnVCM10000022f95190RCRD&vgnnextchannel=e809197a484f4010VgnVCM10000022f95190RCRD.

Croft, C. (2002) *The Assessment Resource Banks and other approaches to school-based assessment in New Zealand*. Paper presented at the Association of Commonwealth Examination and Accreditation Bodies Conference, Malta, March 18–22. Online: available: www.nzcer.org.nz/pdfs/11562.pdf.

Crooks, T. J. (2002) Educational assessment in New Zealand schools. *Assessment in Education: Principles, Policy and Practice*, 9(2), 238 – 253.

Dale, R. (2005) Globalisation, knowledge economy and comparative education. *Comparative Education*, 41(2), 117 – 149.

Department for Children, Schools and Families (DCSF) (2007a) *Diplomas*. Online: available: www.dcsf.gov.uk/14-19/index.cfm?sid=3&pid=224&ctype=None&ptype=Contents.

Department for Children, Schools and the Family (DCSF) (2007b) *New independent regulator for the exam and qualifications system*. Online: available: www.dfes.gov.uk/pns/DisplayPN.cgi?pn_id=2007_0170.

Department for Children, Schools and Families (DCSF) (2007c) *Confidence in Standards: regulating and developing qualifications and assessment: e-consultation*. Online: available: www.dfes.gov.uk/consultations/conDetails.cfm?consultationId=1519.

Education Quality and Accountability Office (EQAO) (2005) *Annual Report 2004 – 2005: working together to improve student learning*. Online: available: www.eqao.com/pdf_e/06/06P013e.pdf.

Educational Testing Services (ETS) (undated-a) *The TOEFL® Test – Test of English as a Foreign Language*. Online: available: www.ets.org/portal/site/ets/menuitem.fab2360b1645a1de9b3a0779f1751509/?vgnnextoid=69c0197a484f4010VgnVCM10000022f95190RCRD.

Educational Testing Services (ETS) (undated-b) GRE® – Graduate Record Examinations®. Online: available: www.ets.org/portal/site/ets/menuitem.fab2360b1645a1de9b3a0779f1751509/?vgnnextoid=b195e3b5f64f4010VgnVCM10000022f95190RCRD.

Egelund, N. (2005) Educational assessment in Danish schools. *Assessment in Education: Principles, Policy and Practice*, 12(2), 203–212.

Ertl, H. (2006) Educational standards and the changing discourse on education: the reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32(5), 619 – 634

Eurydice (2004) *European glossary on education volume 1 – second edition: examinations, qualifications and titles*. Online: available: www.eurydice.org/ressources/eurydice/pdf/0_integral/046EN.pdf.

Eurydice (2005) Certified assessment at the end of general upper secondary education, 2002/03 in Eurydice *Key data on education in Europe 2005*. Online: available: www.eurydice.org/portal/page/portal/Eurydice/FigurePage?pubid=052EN&figCode=052E28.

Eurydice (undated) *Eurydice network*. Online: available: www.nfer.ac.uk/eurydice/about/eurydice-network.cfm.

Eurydice/Eurybase (2007) *The education system in Denmark*. Online: available: www.eurydice.org/ressources/eurydice/eurybase/pdf/0_integral/DK_EN.pdf.

Examinations Appeals Board (EAB) (2008) *Examinations Appeals Board: annual report 2007*. Online: available: www.theeab.org.uk/assets/eab_annual_report_2007.pdf.

Falch, M. (2006) ICT and the future conditions for democratic governance. *Telematics and Informatics*, 23(2),134-156.

Further Education and Training Awards Council (FETAC) (2007) *Quality assuring assessment guidelines for providers*. Online: available: www.fetac.ie/assessment/Quality_Assuring_Assessment_Guidelines_for_Providers_May_2007.pdf.

Government of Western Australia (WA), Evaluation and Accountability Directorate (EAD) (2006) *South Korea (Republic of Korea): school accountability framework review*. Online: available: www.det.wa.edu.au/education/accountability/Docs/SOUTH%20KOREA.pdf.

Grant, R. (undated) *Structure and tension in public and private sector governance: observations from the UK and Australian experience*. National Europe Centre Paper no. 133. Online: available: www.anu.edu.au/NEC/Archive/regulation.pdf.

Green, A. (1997) *Education, globalization and the nation state*. (London: Macmillan).

Green, A., Wolf, A. & Leney, T. (1999) *Convergence and divergence in European education and training systems*. (London: Institute of Education).

Henten, A., Falch, M. & Tadayoni, R. (2002) *Some implications for regulation of ICT and media convergence*. WDR discussion paper #0202. Online: available: www.regulateonline.org/2003/pdf/wdr0202.pdf.

International Review of Curriculum and Assessment Frameworks Internet Archive (INCA) (undated) *International Review of Curriculum and Assessment Frameworks Internet Archive (INCA)*. Online: available: www.inca.org.uk/.

International Review of Curriculum and Assessment Frameworks Internet Archive (INCA) (2005) *Korea : organisation/control of education system*. Online: available: www.inca.org.uk/korea-organisation-mainstream.html.

International Standards Organisation (ISO)/British Standards Institution (BSI) (2007) *Information Technology – a code of practice for the use of Information technology (IT) in the delivery of assessments*. BS ISO/IEC 23988:2007.

International Test Commission (ITC) (2006) International guidelines on computer-based and internet-delivered testing. *International Journal of Testing*, 6(2), 143 – 171.

Isaacs, T. (2001) Entry to University in the United States: the role of SATs and Advanced Placement in a competitive sector. *Assessment in Education: Principles, Policy and Practice*, 8(3), 391 – 406.

Jacobsson, K. (2004) Soft regulation and the subtle transformation of states: the case of EU employment policy. *Journal of European Social Policy*, 14(4), 355-370.

Kuiper, W., van den Akker, J., Letschert, J. & Hooghoff, H. (2007) *Balancing prescription and professionalism in curriculum policy and practices*. Paper presented at ECER 2007 conference, Ghent, Belgium, 19 – 22nd September.

Le Métais, J. (1997) *Values and aims in curriculum and assessment frameworks*. Online: available: www.inca.org.uk/pdf/values_no_intro_97.pdf.

Le Métais, J. (2002) *International developments in upper secondary education: Context, provision and issues*: INCA Thematic Study No. 8. Online: available: www.inca.org.uk/pdf/cav_final_report.pdf.

Leeke, M., Sear, C. & Gay, O. (2003) *An introduction to devolution in the UK*. Parliament and Constitution Centre, House of Commons Library. Research paper 03/84. Online: available: www.parliament.uk/commons/lib/research/rp2003/rp03-084.pdf.

Looney, A. (2006) Assessment in the Republic of Ireland. *Assessment in Education: Principles, Policy and Practice*, 13(3), 345–353.

Lubisi, R.C. & Murphy, R.J.L. (2002) Assessment in South African schools. *Assessment in Education: Principles, Policy and Practice*, 9(2), 258 – 268.

Matthews, A. & Pepper, D. (2007) *Evaluation of participation in GCE mathematics: final report*. Online: available:
www.qca.org.uk/libraryAssets/media/QCA_3388_Maths_GCE_eval_report.pdf.

Messick, S. (1989) Validity in Linn, R.L. (Ed.) *Educational Measurement*, 3rd edn (New York, NY: American Council on Education).

Meuret, D. & Duru-Bellat, M. (2003) English and French modes of regulation of the education system: a comparison. *Comparative Education*, 39(4), 463–477.

Ministry of Education and Human Resources Development (MOEHRD) (2004) *Quality education for all young people: challenges, trends and priorities*. National report of the republic of Korea. Online: available:
www.ibe.unesco.org/International/ICE47/English/Natreps/reports/korearep.pdf.

Morrison, K. (1994) Uniformity and diversity in assessment: an international perspective agenda. *Compare: a Journal of Comparative Education*, 24(1).

Moss, P.A., Girard, B.J. & Haniford, L.C. (2006) Validity in educational assessment. *Review of Research in Education*, 30, 109-162.

Newton, P. (2007) Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy and Practice*, 14(2), 149 – 170.

Newton, P., Baird, J-A., Goldstein, H., Patrick, H. & Tymms, P. (Eds.) (2007) *Techniques for monitoring the comparability of examination standards*. (London: QCA).

Nisbet, I. & Greig, A. (2007) Educational qualifications regulation in Vass, P. (Ed) *Centre for the study of Regulated Industries (CRI) regulatory review 2006/2007* (Bath: CRI). Online: available: www.bath.ac.uk/cri/pubpdf/regulatory_reviews/2006-2007.pdf.

Nisbet, I. (2007) *Intelligent regulation: trust and risk*. Paper presented at invitational conference. New Connaught rooms, London, 8th November 2007.

Oates, T. (2005) *Understanding how regulation works in different national settings*. Qualifications and Curriculum Authority unpublished internal paper.

Oates, T. (2007) *The constraints on delivering public goods – a response to Randy Bennett's 'What does it mean to be a non-profit educational measurement organization in the 21st Century?'* A paper presented at the International Association for Educational Assessment Annual Conference, Baku, Azerbaijan, September 2007.

Organisation for Economic Co-operation and Development (OECD) (2006) *Education at a glance: OECD Indicators 2006*. (Paris: OECD).

Phelps, R. P. (2000) Trends in large-scale testing outside the United States. *Educational Measurement: Issues and Practice*, 19(1), 11–21.

Price Waterhouse Coopers (PWC) (2005) *The market for qualifications in the UK: final report to: Qualifications and Curriculum Authority, Qualifications, Curriculum and Assessment Authority for Wales, Council for the Curriculum, Examinations and Assessment for Northern Ireland*. Online: available: www.qca.org.uk/libraryAssets/media/pdf_05_2239full_report.pdf.

Qualifications and Curriculum Authority (QCA) (2005) *Authenticating coursework: a teacher's guide*. Online: available: www.qca.org.uk/qca_4160.aspx.

Qualifications and Curriculum Authority (QCA), Qualifications and Curriculum Group, Department for Education, Lifelong Learning and Skills (DELLS) & Council for the Curriculum, Examinations and Assessment (CCEA) (2007a) *GCSE, GCE, GNVQ and AEA code of practice*. Online: available: www.qca.org.uk/libraryAssets/media/qca-07-3082_Unified_Code_Practice_web_April07.pdf.

Qualifications and Curriculum Authority (QCA), Department for Children, Education, Lifelong Learning and Skills (DCELLS), Council for the Curriculum, Examinations and Assessments (CCEA) & Scottish Qualifications Authority (SQA) (2007b) *Regulatory principles for e-assessment*. Online: available: www.qca.org.uk/libraryAssets/media/Final_regulatory_principles_document_-_PRINTED.pdf.

Qualifications and Curriculum Authority (QCA), Department for Children, Education, Lifelong Learning and Skills (DCELLS), Council for the Curriculum, Examinations and Assessments (CCEA) & Scottish Qualifications Authority (SQA) (2007c) *e-Assessment: guide to effective practice*. Online: available: www.efutures.org/docs/guide.pdf.

Qualifications and Curriculum Authority (QCA), the Qualifications, Curriculum and Assessment Authority for Wales/Awdurdod Cymwysterau, Cwricwlwm ac Asesu Cymru (ACCAC), & the Council for the Curriculum, Examinations and Assessment (CCEA) (2004) *The statutory regulation of external qualifications in England, Wales and Northern Ireland*. Online: available: [www.qca.org.uk/libraryAssets/media/6944_regulatory_criteria_04\(1\).pdf](http://www.qca.org.uk/libraryAssets/media/6944_regulatory_criteria_04(1).pdf).

Ramstedt, K. (2005) *National assessment and grading in the Swedish school system*. Online: available: www.skolverket.se/publikationer?id=1524.

Régent, S. (2002) *The open method of co-ordination: a supranational form of governance?* Decent Work Research programme discussion paper, No. 137, International Institute for Labour Studies, Geneva. Online: available: www.ilo.org/public/english/bureau/inst/download/dp13702.pdf.

Scottish Qualifications Authority (SQA) (2003) *SQA Guidelines on online assessment for Further Education*. Online: available: [www.sqa.org.uk/files_ccc/GuidelinesForOnlineAssessment\(Web\).pdf](http://www.sqa.org.uk/files_ccc/GuidelinesForOnlineAssessment(Web).pdf).

Scottish Qualifications Authority (SQA) (2005) *SQA guidelines on e-assessment for schools*. Online: available: www.sqa.org.uk/files_ccc/SQA_Guidelines_on_e-assessment_Schools_June05.pdf.

Scottish Qualifications Authority (SQA) (undated) *About the Scottish Qualifications Authority (SQA)*. Online: available: www.sqa.org.uk/sqa/5656.html.

Spyrelli, C. (2003) Regulating the regulators? An assessment of institutional structures and procedural rules of national regulatory authorities. *International Journal of Communications Law and Policy*, 8.

State Examinations Commission (SEC) (2005) *Annual report*. Online: available: www.examinations.ie/about/060306Finalpdf.pdf.

State of Queensland (2002) *Education (Queensland Studies Authority) Act 2002*. Online: available: www.legislation.qld.gov.au/LEGISLTN/CURRENT/E/EducQStAuA02.pdf.

Tattersall, K. (2007) A brief history of policies, practices and issues relating to comparability in Newton, P., Baird, J-A., Goldstein, H., Patrick, H. & Tymms, P. (Eds.) *Techniques for monitoring the comparability of examination standards*. (London: QCA).

Victoria Curriculum and Assessment Authority (VCAA) (undated) *About us*. Online: available: www.vcaa.vic.edu.au/aboutus/index.html.

West, R. & Crighton, J. (1999) Examination reform in Central and Eastern Europe: issues and trends. *Assessment in Education: Principles, Policy and Practice*, 6(2), 271 – 289.

Whitby, K. & Lord, P. with O'Donnell, S. & Grayson, H. (2006) *Dips in performance and motivation: a purely English perception?* Online: available: www.inca.org.uk/pdf/Final_Dip_Report.pdf.

Wikström, C. (2006) Education and assessment in Sweden. *Assessment in Education: Principles, Policy and Practice*, 13(1), 113–128.

Wolf, A. (2002) *Does Education Matter? Myths about education and economic growth*. (London: Penguin).

All web references were live as at 13th March 2008.

Ofqual wishes to make its publications widely accessible. Please contact us if you have any specific accessibility requirements.

First published by The Office of the Qualifications and Examinations Regulator in 2008.

© Qualifications and Curriculum Authority 2008

Ofqual is part of the Qualifications and Curriculum Authority (QCA). QCA is an exempt charity under Schedule 2 of the Charities Act 1993.

Reproduction, storage or translation, in any form or by any means, of this publication is prohibited without prior written permission of the publisher, unless within the terms of the Copyright Licensing Agency. Excerpts may be reproduced for the purpose of research, private study, criticism or review, or by educational institutions solely for education purposes, without permission, provided full acknowledgement is given.

Office of the Qualifications and Examinations Regulator
Spring Place
Coventry Business Park
Herald Avenue
Coventry CV5 6UB

Telephone 0300 303 3344
Textphone 0300 303 3345
Helpline 0300 303 3346

www.ofqual.gov.uk