

Predictability in the Irish Leaving Certificate Examination

Working Paper 1: Review of the Literature

Jo-Anne Baird, Therese Hopfenbeck, Ayesha Ahmed, Jannette Elwood, Christine Paget and Natalie Usher

This research was sponsored by the State Examinations Commission (SEC) of Ireland. Ruairí Quinn, Minister for Education and Skills in Ireland, announced this project and his commitment to tackle any problematic predictability in the Leaving Certificate examinations.¹

Contents

Introduction	2
What counts as valuable learning?	3
Implications of examination predictability	5
Washback effects of predictability upon learning and teaching	8
Beneficial and problematic effects of washback	8
Deep and surface learning	9
Empirical studies of washback.....	10
What students need in the 21 st century	11
Methodologies for investigating examination predictability.....	12
References	14
Appendix A: Irish media analysis on allegations of predictability in 2012.....	19
Appendix B: Preliminary analysis of media coverage 2013	26

¹ Department of Education and Skills (2013) *Supporting a better transition from second level to Higher Education: Key directions and next steps*. 27 March. (<http://www.education.ie/en/The-Department/Re-use-of-Public-Sector-Information/Library/Announcements/Supporting-a-Better-Transition-from-Second-Level-to-Higher-Education.html>)

Introduction

Assessment in many countries is no longer the secret garden that it once was. In society more broadly, there has been more questioning of experts (O'Neill, 2002) and it is in this context that examining bodies have gradually become more open about the design of their question papers and mark schemes. Openness and transparency about important examinations is in the interests of the public good, as examination outcomes should be less dependent upon access to the knowledge of a select group of examiners. After all, if we want students to learn, why not be explicit about what it is we want them to know and how they should demonstrate their skills and knowledge? For these and other reasons, criterion-referencing assessment gained popularity over norm-referenced assessment in many countries (eg Black, 1994; Hambleton, 1994; Wikström, 2005). Norm-referenced assessment is a process in which students are graded according to the rank of their scores on a test.² Glaser (1963) is credited with the first definition of criterion-referenced assessment:

the determination of the characteristics of student performance with respect to specified standards (p519)

Thus, there was a shift towards being more open and explicit about the achievement standards expected of students. Even if examinations such as the Leaving Certificate continue to be influenced by norm-referenced considerations, the criterion-referencing *zeitgeist* is likely to have affected the way of thinking about examining in Ireland, as well as in other countries.³ Thus, in a review of standards in public examinations between 1975 and 1995 in England, the School Curriculum and Assessment Authority (SCAA) and Office for Standards in Education (Ofsted) commented that

Subject coverage and assessment requirements are set out in syllabuses in far more detail now than was the case 20 years ago. Clearer examination papers and the publication of mark schemes have enabled teachers and candidates to see more precisely than before what is expected. (1996, p15)

Additionally, they pointed out that since 1990 the objectives, aims and philosophy of the examinations had been increasingly more explicit and available in the public domain (ibid, p17).

In more recent years, the rise of formative assessment (Black & Wiliam, 1998) has been a dominant shift in the assessment research literature. In part, this builds upon the criterion-referencing literature, as one aspect of successful formative assessment is sharing assessment criteria with students. This has been found to have positive effects upon student learning and is related to the capacity to self-assess (McDonald & Boud, 2003).

In a special issue of the journal *Assessment in Education*, Torrance (2007) drew together the above trends and noted how the transparency they have produced resulted in instrumentalism related to qualification attainment. He named this shift 'assessment *as* learning' because complying with criteria can displace learning more broadly. Data presented in the special issue (see Torrance, 2007) related to a range of assessment types, including traditional academic examinations, vocational assessments, formative assessments, coursework and portfolios. The work of understanding the

² Although Wiliam (1996) correctly pointed out that this is typically cohort-referencing as it usually applied to each year group rather than compared with a sample population.

³ West and Crighton (1999) also observed a trend towards competence-based assessment in Central and Eastern Europe.

requirements of assessment was scaffolded to a large extent by tutors, who broke them down into digestible chunks, or 'atomised teaching and assessment', in Sadler's (2007) terms. Torrance (2007, p291) argued that transparency undermined the development of learner autonomy and wrote that

Making learning objectives and instructional processes more explicit calls into question the validity and worthwhileness of the outcomes achieved. (p389)

Sadler (2007) reacted to the findings of those articles with despair, writing:

In my heart of hearts, I believe that a blinkered conceptualisation of curriculum, the strong tendency towards fine-grained prescription, atomised assessment, the accumulation of little 'credits' like grains of sand, and intensive coaching towards short-term objectives, are a long call from the production of truly integrated knowledge and skill. (p391)

McNeil and Valenzuela (2001) vividly depicted the effects of a target-driven education and testing system upon drilling of students in Texas. They showed how repetitive and superficial the learning and teaching became and how even the students who got their high school diplomas felt let down by the system and in one case staged a sit-in protest at the graduation ceremony. However, it is possible that the learning, teaching and assessment situations described in the collection of papers referred to by Torrance (2007) or by McNeil and Valenzuela (2001) may be very different from the situation in Ireland.

The Leaving Certificate is not a modular assessment, nor are assessment criteria presented as checklists (see Wolf, 2002). Nonetheless, as we outline in our media analysis (Appendix A and Appendix B), there have been significant public expressions of concern about the learning associated with them. Sadler (2007) argued that the focus upon credentials in modern society is a trap and it may be this factor that is key in the Leaving Certificate, as it is a high-stakes examination. In Section 4 we explore the effects of high-stakes testing upon teaching and learning. For now, we simply note that unfortunately, assessment is subject to Goodhart's Law, which is that

When a measure becomes a target, it ceases to be a good measure. (Strathern, 1997, p308)

The importance of qualifications in modern society makes them a target for students, even if there are no league tables for schools in Ireland. Instrumentalism is a negative view of behaviour in relation to these targets, but students are also highly strategic in their approaches to learning for high-stakes qualifications (Daly, Baird, Chamberlain & Meadows, 2012).

What counts as valuable learning?

Underlying the debates about the predictability of assessment are tensions relating to what counts as valuable learning. However, the model of learning underpinning assessment is rarely articulated (Murphy, 1999). Recently, there have been attempts in the literature to make more explicit links between theories of learning and assessment (Black, 1999; Pryor & Crossouard, 2008; Elwood, 2008; James, 2006; Moss et al, 2008; Murphy, 1999; Sfard, 1998). A number of writers have pointed to the fact that our views of learning can be at odds with assessment design. Added to this, learning itself

can be mediated by assessment design because assessment is so powerful in the process of defining curriculum and learning. Theories of learning are contested and the prominence of certain theories has changed over time. Black (1999) and James (2006) have related developments in assessment formats to changes in emphases given to various theories.

The word 'examination' conjures a schema that we are all familiar with from our own educational experiences. Students are sat in rows in a hall, in silence, with no access to their notes or books and are tested on their understanding of the curriculum. The examination paper itself relates to the topics they have learned, although there could be an element of surprise about which topics come up or the focus of the questions. Our schema for school examinations involves these elements of controlled conditions, reacting to questions with an element of unpredictability, but not so much unpredictability that they are not recognisably related to the curriculum.

For the first half of the 20th century, a behaviourist theory of learning had precedence. In this conceptualisation of learning, mental states are not of much interest and the main focus is upon how people behave in relation to their environments. Black (1999) and James (2006) argued that traditional forms of testing, with knowledge reproduced under controlled conditions, as outlined in the familiar schema above, reflect this behaviourist model of learning. Nowadays we are not impressed with a reproduction of facts without understanding.

Cognitive constructivist views of learning became more accepted in the 1960s and the thinking skills in learning were emphasised. As such, partial credit for thinking processes rather than all of the credit for a correct answer became a more frequently used scoring mechanism, and extended responses were considered useful in allowing students the opportunity to explain their thinking and evaluation of materials to which they had been exposed. A third view of learning stems from sociocultural theory. Here, learning is not seen as something acquired by the learner from the teacher, but as something that is jointly created between the learner and their social environment (see Sfard, 1998). According to this theory, students can create new knowledge and therefore assessment needs to capture this by utilising a variety of techniques that are closely tied to the learning situation (eg learning portfolios or coursework). Learners are seen as active participants in the process. Where they learn and from whom they learn are key elements in defining learning in this perspective (Elwood, 2008).

Tracing the implications of these high-level theories for assessment is not easy because as with most societal trends, it is difficult to show a causal connection. Furthermore, the different theories contest how learning should be defined in terms of the outcome *and* the process. Our point here is to indicate that examination commentators will have different views about what counts as valuable learning – the extent to which knowledge or skills should be emphasised and which knowledge and skills are valued in each subject. Education is fundamentally concerned with cultural transmission, so we can expect societal debate over whether the right signals are being sent by our assessment systems.

By making things more transparent, opportunities for credit-worthy performances become more accessible and the proportion of students who attain qualifications can rise. Two key functions of education, selection and the facilitation of learning are thus also tensions that play out through assessment systems, with some stakeholders believing that it is the performances demonstrated that are most important and others believing that only a certain proportion of learners can attain

the highest standards (Baird, 2007). In a test for student teachers in Sweden, Jonsson (2010) showed that making the assessment criteria, scoring scheme and exemplar work available had a very large impact upon performances. For some stakeholders, creating assessment conditions in which too many can achieve the required performances is not acceptable, yet for others, fostering educational attainment for all is highly desirable. National assessments therefore have a very difficult balance to strike if they are to be broadly accepted as a useful societal currency. The Leaving Certificate needs to strike the right balance for each subject assessed for a wide range of requirements, not all of which are under the direct control of those setting the examinations.

Implications of examination predictability

Ultimately, debates over the predictability of examinations relate to the extent to which assessments allow students to demonstrate whether they really understand a broad range of material in a given subject area. Learning a narrow range of material in a superficial way, which may soon be forgotten and is disconnected from the broader subject field, is clearly undesirable. If we were only interested in the regurgitation of facts, in line with a behaviourist tradition, we would not balk at this. Facts are necessary to learning, but additionally we want people to connect them with other information, to be able to analyse and evaluate the material they are engaging with.

Some elements of predictability for teachers and learners are necessary to ensure that a structured educational experience can be provided. After all, the Leaving Certificate is a curriculum-related examination, unlike, for example, Programme for International Student Assessment (PISA) tests, which claim to be curriculum independent. To be able to learn, practise and perform, students need to know the topics to study, what performances are expected, how they will be assessed and what counts as progress in a subject. Examinations are not intended to be entirely unpredictable because we want students to be able to prepare for them.

Yet, if these principles were taken too far, automatons understanding nothing would be able to attain highly in examinations. This deeply philosophical question has plagued the field of artificial intelligence – how do we know when someone (or something) understands, *really*? In practice, our answers to this have been to require a broad range of performances (different topics and assessment formats), in-depth explanations in extended writing, and assessment objectives requiring analysis and evaluation. Our practices for evidencing understanding involve repeated observations of performances, with different requirements, and assessment of higher order thinking skills as well as factual knowledge. Unfortunately, people can be drilled into performing well on these kinds of assessments in undesirable ways. Gordon and Reese (1997) reported that higher order thinking skills were being rote learned for an examination in Texas. Nevertheless, examining is a business of finding practical solutions that are acceptable to society for difficult educational and philosophical problems.

We have already seen that making the assessment criteria transparent can increase outcomes in assessments (Jonsson, 2010), but we have also seen that it can foster rote and superficial learning, as well as drilling and instrumentalism. On the positive side, students might be less anxious taking predictable examinations because they can prepare for them and there are few surprises. A predictable examination is also likely to have the technical quality of reliability, as it will assess

consistently. Consistently assessing the wrong thing is not very useful though – we need our examinations to have the technical property of validity too.

Using an experimental design and structural equation modelling, Kleinmann, Kuptsch and Köller (1996) found that people's performances were explained by ability to a larger extent when the scoring criteria were made transparent in an occupational assessment centre. Thus, making the criteria transparent increased the construct validity of the assessments. However, it has been seen above that there is a balance to be struck here, as too much transparency led to undesirable implications for teaching and learning in England (Daly et al, 2012; Torrance, 2007) and in Texas (McNeil & Valenzuela, 2001).

In another experimental study, Crisp et al (2008) manipulated the design of a number of GCSE questions and tested them with 339 sixteen year olds in England. Here, the authors were interested in the effects of students' expectations regarding the meaning of the questions, rather than predictability per se. Findings indicated that students had expectations related to the subject (eg the kind of questions and answers expected in science), the questions (eg difficulty, relevance of resources, space allotted for a question) and the language used (eg expecting questions to be phrased positively). The authors argued that the place for contradicting students' expectations is in the classroom and not in the examination. They concluded:

By avoiding misunderstandings as much as possible we can ensure that more students are trying to do the things we want them to show us they can do, and so increase the construct validity of examination questions. (p113)

There are different views about the extent to which the examination questions should meet student expectations. Using unpredictable, novel questions in examinations makes it difficult for students to use the strategy of recalling a response that they have already prepared. As such, using novel questions and materials is one technique for assessing higher order thinking skills rather than rote learning (Brookhart, 2010, p26).

In Table 1 we depict aspects of examinations that produce predictability and draw out some of the implications. As Ofqual (2008) noted, these features can interact with each other, so that particular combinations may be problematical, while other combinations strike an acceptable balance. However, unpredictable question formats could result in students' cognitive resources being used to decipher the question rather than tackle the demands of the task.

Table 1. Elements of examination predictability

Feature	Predictable	Unpredictable
Curriculum coverage	<ul style="list-style-type: none"> • Know the topics that will be assessed • May not need to study the breadth of material intended • Teachers may narrow the taught curriculum 	<ul style="list-style-type: none"> • Do not know how to prepare for the exam • Performance based upon luck of studied curriculum–exam match • Teachers must judge which aspects of the syllabus to teach
Test conditions	<ul style="list-style-type: none"> • Known in advance and can be practised 	<ul style="list-style-type: none"> • Vary and therefore students’ capacity to adapt is part of the assessment
Question format	<ul style="list-style-type: none"> • A set format for questions, perhaps even related to specific topic areas, is known in advance • The phrasing and structure of questions can be explained to students in advance and they can be taught test wisdom 	<ul style="list-style-type: none"> • Novel question styles are used frequently • Teachers can prepare students to think about what is required to respond to different question styles • The assessed construct changes
Performance format	<ul style="list-style-type: none"> • How students will be required to respond is known in advance • Teachers can school students on how to produce the kinds of performances required 	<ul style="list-style-type: none"> • Changes are made to the ways in which students’ knowledge and skills are demonstrated • Match between performance required and student skills will affect results • The assessed construct changes
Scoring	<ul style="list-style-type: none"> • How performances are credited is known openly • Detail is known regarding scoring • It is possible to learn the scoring criteria rather than the syllabus materials in an extreme case 	<ul style="list-style-type: none"> • Information on rubrics is not available • What is credit-worthy may change in relation to changes in the assessment
Examination support materials	<ul style="list-style-type: none"> • Past papers available publicly • Model answers accessible • Advisory materials from the examination board, such as examiners’ reports, available • Textbooks closely aligned with examination questions • Other publicly available information, such as newspaper coverage of the examinations, teacher publications etc 	<ul style="list-style-type: none"> • Little information publicly available relating to the examination questions

We have the following working definition of predictability.

A problematically predictable examination is one in which teachers and students can anticipate the test-taking conditions, performances required, question formats and topics and scoring to the extent that undesirable backwash upon the educational process is pervasive. These include narrowing of the taught curriculum, superficial, rote learning, drilling on test content, and failure to develop a broad and deep understanding of a subject. Such an examination lacks validity because it measures test preparation narrowly rather than the intended assessment objectives. Predictability is not a necessary or sufficient condition for reliable assessment, but it is likely to foster reliability. Reliability may be won at the expense of validity.

Washback effects of predictability upon learning and teaching

Washback or **backwash** are terms for the influences of testing upon teaching, attitudes, students' approaches to learning, and behaviours (Alderson & Wall, 1993). In their seminal work, Alderson and Wall (1993) argued that tests will have washback effects on some teachers and learners, but not for others, and it is likely that high-stakes tests will have more influence than tests that do not have any important consequences for the learner or teacher. Messick (1996) described washback in language testing as

The extent to which the test influences language teachers and learners to do things that they would not necessarily otherwise do. (p243)

In Messick's definition, there is a focus upon behaviour, and the way teachers and students approach learning based upon what they know about the examination. Messick argued that the move from learning exercises to test exercises should be seamless, and positive washback effects would therefore involve little difference between activities involving learning and activities preparing for the test (Messick, 1996, p242).

Madaus (1988) stated that

It is testing, not the "official" stated curriculum, that is increasingly determining what is taught, how it is taught, what is learned, and how it is learned. (p83)

In other words, those who control testing control the real curriculum in school, and therefore learning. Looking at empirical research we will see that this is not always the case and the picture is more complex when we look at (1) what is taught and how (which involves syllabus content and the teacher), and (2) what is learned and how (which involves the student).

Beneficial and problematic effects of washback

Pearson (1988) claimed that a good test would enhance good teaching, and would be good for teaching and learning activities, while other studies show both problematic and beneficial sides of washback (Buck, 1988). In the late 20th century, the measurement-driven instruction movement in the US promulgated the notion that tests should drive the instruction through clear goals for students, which could be tested on a high-stakes test (Popham, 1987). The process of matching the

content of the teaching with the items on the tests is often referred to as curriculum alignment (Shepard, 1990, 1993). There are conflicting views of this form of backwash of tests. Curriculum alignment has been alleged to be unethical and to undermine validity (Haladyna, Nolen & Haas, 1991). Cheng and Curtis (2004) reported that curriculum alignment is favoured in a number of countries, for instance in Hong Kong. In the US, measurement-driven instruction has been viewed by some as positive reinforcement for the students, but there is no consensus (Madaus, Russell & Higgins, 2009).

The National Commission on Testing and Public Policy warned that high-stakes testing could drive teaching away from 'instructional practices that would help to produce critical thinkers and active learners' and narrow the curriculum (Madaus et al, 2009). More than half a century ago, Vernon (1956) claimed that teachers tended to ignore topics that were not on the exam, and warned that tests could 'distort the curriculum' (ibid, p166).

It has been found that states and districts in the US that rely primarily on test-based accountability have not been successful in supporting their most vulnerable students (Darling-Hammond, 2010). Instead, teaching was primarily focusing upon rote learning, memorising pieces of information and low-level learning, which was particularly evident in schools with disadvantaged students (Darling-Hammond & Rustique-Forrester, 2005). In England, high-achieving students reported being taught test techniques at the expense of deep learning in the classroom in a research study investigating the introduction of A-level exams in 2008 (Daly et al, 2012).

A recent review of the effects of testing on student achievement between 1910 and 2010 found positive effects of testing both in qualitative and quantitative studies, with moderate to fairly large effect sizes across hundreds of studies (Phelps, 2012). Rote learning and coaching for tests, and predictable tests, can increase scores, but there is less understanding of how much deep knowledge students have or how prepared they are for future learning.

Spratt (2005) summarised research showing that teachers varied in how they responded to exams, and therefore the real changes and washback on teaching also varied. With reference to the work by Alderson and Hamp-Lyons (1996) and Watanabe (1996, 2000), Spratt argued that there were large differences in the way teachers teach towards the exam, and that an exam cannot dictate what and how teachers teach and learners learn. The degrees of washback varied depending upon the teacher, and how the teacher decided to make the most of teaching towards exams, and these decisions are influenced by the ideological, historical, economic and political context in which teachers work. Some teachers used 'teaching to the test' and 'textbook slave' approaches, while others adopted more creative and independent approaches. In other words, it could be that the most influential variable was not the exam at all, but the teacher and how they responded to the exam. Alderson and Wall (1993) claimed that any test, good or bad, may result in beneficial or problematic washback, depending upon how it is used.

Deep and surface learning

As the SEC (2012) report underlined, problematic predictability can lead to undesirable teaching and learning behaviours, primarily focusing on the tests instead of broad, deep learning and higher order thinking. In their classic research on the quality of students' learning, Marton and Säljö (1976) found that students in higher education adapted their study strategies according to how they perceived

the test items and how they believed they would be assessed. A significant number of students used approaches to learning, or learning strategies, that they believed best corresponded to the demands of the learning tasks, and as a result of using these strategies, they increased their achievement outcomes (Marton & Morris, 2002; Marton & Säljö, 1976). Students' approaches to learning could be classified as either deep or surface approaches to learning. Deep approaches to learning were characterised by students' active search for meaning in texts they read while surface approaches to learning focused primarily upon rote learning and factual knowledge (Biggs, 2003; Diseth & Martinsen, 2003; Entwistle & Entwistle, 1991; Entwistle, 1988; Ramsden, 1988; Shepard et al, 2005).

One example of how students adapted their approaches to learning based upon their knowledge of how they are assessed was given in the longitudinal project *Effectiveness of Higher Education* in the Netherlands, which involved 565 first-year students in the University in Groningen (Bruinsma, 2004). During the first two academic years, students answered self-report questionnaires about their motivation, deep information processing approach and on student achievement. It was predicted that the use of a deep information processing approach would lead to higher achievement, but it did not. The researchers (Bruinsma, 2004) offered the following explanation:

...the relationship between a deep information processing approach and academic achievement might be mediated by other characteristics of the learning environment, eg the assessment procedure. The assessment procedure might focus on memorization knowledge rather than on a conceptual mastery of the subject matter...if students need to learn deep information processing strategies, they should be assessed on these strategies and not on surface learning strategies. (p564)

In other words, the way students are assessed can influence how they approach their learning tasks. This is also emphasised in the US National Research Council report *How People Learn*; if teachers stress the importance of understanding in their teaching, but give students tests that focus upon memorisation of procedures and facts, the latter will be what students focus upon. In fact, the report claims that most teachers over-emphasise memorisation in their classroom assessments, rather than depth of understanding (Bransford et al, 2000).

Empirical studies of washback

In 1982, Kellaghan, Madaus and Airasian published *The Effects of Standardized Testing*, a joint Irish and American investigation that examined the impact on Irish schools of introducing standardised tests. The study examined washback effects on schools, teachers, pupils and parents of Grade 2 and 6 students, and found no harmful effects of introducing standardised tests. However, this study has been criticised as an artificial situation, in which tests were introduced as part of the study experiment, with no currency in the Irish education system, and therefore it was no surprise that little negative impact was perceived (Alderson & Wall, 1993).

In a study of washback effects from the National Matriculation English Test (NMET) in China, Qi (2007) suggested that high-stakes tests were not an efficient agent for changing teaching practices. It was found that teachers neglected the communicative features of the NMET writing task, and no practice was observed in the schools that prepared the students for the task. It was concluded that language testing seems to influence teaching content, but has limited influence on teaching

methodology. In another study, Cheng (2003) examined the impact of changes in public examinations on secondary school English teaching in Hong Kong, where major changes were introduced into the Certificate of Education Examination in English with the intention of creative positive washback effects on classroom teaching. Data from teacher observations and interviews showed little change in classroom practice.

In an American study, Smith (1991) reported different ways teachers responded to tests, by studying their test preparation in the classroom, ending with categorising eight different responses:

1. no special preparation (teachers showing resistance to test)
2. teaching test-taking skills
3. exhortation, such as encouraging students to have a good night of sleep
4. teaching the content known to be covered by the test
5. teaching to the test, using material which would help students to understand the item format
6. stress inoculation, and boosting students' self-esteem before the test
7. practice on items from the test or parallel forms
8. cheating

In this American study, teachers defined cheating as providing students with extra time on the test, giving hints and rephrasing wording of items, providing correct answers or altering marks on answer sheets (Smith, 1991). The study raised the question over whether teachers' different responses to tests were ways of coping with a negative system of testing, where the test was an accountability measure that was used politically.

In England, Sturman (2003) reported a variety of responses from science teachers, arguing that teaching to the test can in some forms become a positive phenomenon, claiming that the science test already provided opportunities for students to apply their knowledge, not simply recall facts.

What students need in the 21st century

If the goal is to enhance students' critical thinking, problem-solving skills, and use deep approaches to learning, reasoning and reflection, tests need to have tasks that are rich and complex (Brookhart & Nitko, 2011). Changes in the job market following the use of ICT have led to a shift in educational demand, including a 60% increase in the demand for college educated labour between 1970 and 1998 (Autor, Levy & Murande, 2003). In Ireland, Freeney and O'Connell (2012) reported a study showing that almost 20% of students do not sit the Leaving Certificate. The opportunities available to these students are extremely limited, and they risk being excluded from the labour market. A gender gap in retention rates was also found, with 75.5% of males compared to 85.7% of females completing their Leaving Certificate examination (Byrne, McCoy & Watson, 2008). Furthermore, rapid changes in society call for an education system focusing upon 21st-century skills such as the four 'C's of creativity, critical thinking, communication and collaboration (OECD, 2012). Darling-Hammond has summarised these 21st-century skills for students as being able to find, evaluate, synthesise, and use knowledge in new contexts, and the ability to acquire well-developed thinking, problem-solving and communication skills (Darling-Hammond, 2010).

Examinations are generally thought to have a large effect upon behaviour in classrooms. Some authors have argued that it is teachers, rather than examinations, that change the learning experience. Aligning the curriculum and teaching with the examination can be seen as positive, but becomes problematic if it narrows the taught curriculum too far. Producing rote, superficial learning has also been seen as an undesirable feature of examination washback. Modern examinations need to produce positive washback upon learning to meet the needs of the 21st century.

Methodologies for investigating examination predictability

Three reports have been identified which outline methodologies for investigating examination predictability directly (Table 2). Two of these were draft internal reports that may be published at a later date in a more final form (Murphy, Stobart, Baird & Winkley, 2012; State Examinations Commission, 2012). Indeed, the Murphy et al report outlined strategies for investigating predictability, rather than describing a completed study. Importantly, they identified that undesirable forms of predictability could be caused by factors beyond the control of the examiners who wrote the question papers or even the examination boards who are responsible for them. Some of these have already been referred to in the literature review above, but others were not considered:

1. Policy and regulation
2. Examination board processes and cultures
3. Examining and resources
4. School level responses

None of the above causes are considered outside the scope of this study, but we do not intend to research the first two directly. Implications of policy and regulation or examination processes and cultures would have to be an emergent finding of the data collection on examination materials, perceptions of teachers and students in schools, indicated through the documentary materials related to the Leaving Certificate or through the review of our research and findings conducted by some of those working inside the system.

All three reports (Table 2) suggested the following methods: searching the literature, examination materials analysis and (two of them suggest) interviews. Expert identification of issues is also a noteworthy methodology from the Ofqual (2008) study. A range of possible research avenues were suggested in the Murphy et al (2012) paper, but those were intended as follow-up studies to a first phase of research in which the methodologies used in the other two studies were adopted.

Table 2. Studies of examination predictability

Examination	Method(s)	Details	Reference
England <i>GCSE & A-level</i>	<ul style="list-style-type: none"> • Expert review • Literature search • Exam materials analysis • Interviews 	<ul style="list-style-type: none"> • Identification of problematic areas by staff • Review of reports which identified predictability issues to select subjects for study • Structured review of examination materials by 2–3 subject consultants for each examination • Review of students’ examination booklets • Interviews with teachers (2–6 institutions for each subject) • Group interviews with students (up to 6 per group) 	Ofqual (2008)
Ireland <i>Leaving Certificate</i>	<ul style="list-style-type: none"> • Literature search • Exam materials analysis 	<ul style="list-style-type: none"> • 26 assessment managers completed a questionnaire on predictability in specific subject areas 	SEC (2012)
England <i>GCSE & A-level</i>	<ul style="list-style-type: none"> • Literature search • Exam materials analysis • Empirical study of expert predictions • Marker opinion survey • Examination policy research • Classroom practices • Interviews 	<ul style="list-style-type: none"> • Review of examination materials • Subject experts to predict the likely examination questions, subsequently compared with examination • Markers surveyed on student performance evidence that indicates predictable questions (eg formulaic responding from certain schools) • Longer term research on the development of examinations in specific subject areas related to factors which cause predictability (eg policy decisions that have affected predictability) • Classroom practices of teachers relating to textbook use, syllabus coverage, examination technique preparation etc • Interviews with experienced assessment experts on how public examinations have changed over a number of decades in relation to predictability 	Murphy et al (2012)

Research will be conducted in two further phases, as follows:

Analysis of examination materials

- conducted in 6 subjects
- participants are assessment researchers (6) and subject specialists (12)
- materials include 10 years' question papers and marking schemes

Research with teachers and learners

- 12 institutions included in fieldwork
- interviews with 6 subject teachers (72 in total)
- group interviews of approximately 6 learners in each institution (72 students involved in total)
- questionnaire survey of learners to 100 schools

References

- Alderson, J C & Hamp-Lyons, L (1996) TOEFL preparation courses: a study of washback. *Language Testing*, 13(3), 280–297
- Alderson, J C & Wall D (1993) Does washback exist? *Applied Linguistics*, 14(2), 115–129
- Autor, D H, Levy, F & Murnane, R (2003) The Skill Content of Recent Technological Change: an Empirical Exploration. *The Quarterly Journal of Economics*, 118(4), 1279 –1333.
- Baird, J-A (2007) Alternative conceptions of comparability. In P Newton, J Baird, H Goldstein, H Patrick & P Tymms (Eds), *Comparability of UK Public Examinations*. London: QCA book publication
- Biggs, J B (2003) *Teaching for Quality Learning at University* (2nd edn) Buckingham: Open University Press/Society for Research into Higher Education
- Black, P (1994) Performance assessment and accountability: the experience in England and Wales. *Educational Evaluation and Policy Analysis*, 16(2), 191–203
- Black, P (1999) Assessment, learning theories and testing systems. In P. Murphy (Ed), *Learners, Learning and Assessment*. London: Paul Chapman, 118–134
- Black, P & Wiliam, D (1998) Assessment and classroom learning. *Assessment in Education: principles, policy and practice*, 5(1), 7–75
- Bransford, J, Brown, A, Cocking, R, Donovan, M, & Pellegrino, J. (2000). *How People Learn: brain, mind, experience, and school*. Washington, DC: National Academy Press
- Brookhart, S (2010) *How to Assess Higher Order Thinking Skills in Your Classroom*. Alexandria, VA: ASCD
- Brookhart, S & Nitko, A (2011) Strategies for constructing assessments of higher order thinking skills. In G Schraw & D R Robinson (Eds), *Assessment of Higher Order Thinking Skills: a volume in current perspectives on cognition, learning and instruction*. Charlotte, NC: Information Age Publishing Inc
- Bruinsma, M (2004) Motivation, cognitive processing and achievement in higher education. *Learning and Instruction*, 14, 549–568

- Buck, G (1988) Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10, 15–42
- Byrne, D, McCoy, S, & Watson, D. (2008) *The School-Leavers Survey 2007*. Dublin
- Cheng, L (2003) Looking at the impact of a public examination change on secondary classroom teaching: a Hong Kong case study. *Journal of Classroom Interaction*, 38(1)
- Cheng, L & Curtis, A (2004) Washback or backwash: a review of the impact of testing on teaching and learning. In L Cheng, J Watanabe & A Curtis (Eds), *Washback in Language Testing: research contexts and methods*. New Jersey: Lawrence Erlbaum Associates Inc. Routledge, 3–18
- Crisp, V, Sweiry, E, Ahmed, A, & Pollitt, A (2008) Tales of the expected: the influence of students' expectations on question validity and implications for writing exam questions. *Educational Research*, 50(1), 95–115
- Daly, A L, Baird, J-A, Chamberlain, S, & Meadows, M (2012) Assessment reform: students' and teachers' responses to the introduction of stretch and challenge at A-level. *Curriculum Journal*, 23(2), 139–155
- Darling-Hammond, L (2010) *The Flat World and Education: how America's commitment to equity will determine our future*. New York: Teachers College Press
- Darling-Hammond, L. & Rustique-Forrester, E (2005) The consequences of student testing for teaching and teacher quality. In J Herman & E Haertel (Eds), *The Uses and Misuses of Data in Accountability Testing: the 104th yearbook of the National Society for the Study of Education, part II*. Malden, MA: Blackwell, 289–319
- Diseth, Å & Martinsen, Ø (2003) Approaches to learning, cognitive style, and motives as predictors of academic achievement. *Educational Psychology*, 23, 195–207
- Elwood, J. (2008). Gender issues in testing and assessment. In P Murphy & K Hall (Eds.), *Learning and Practice: agency and identities*. London: SAGE/Open University, 87–101
- Entwistle, N (1988) Motivational factors in students' approaches to learning. In R Schmeck (Ed), *Learning Strategies and Learning Styles*. New York: Plenum
- Entwistle, N & Entwistle, A (1991) Contrasting forms of understanding for degree examinations: the student experience and its implications. *Higher Education*, 22(3), 205–227
- Freeney, Y & O'Connell, M O (2012) The predictors of the intention to leave school early among a representative sample of Irish second-level students. *British Educational Research Journal*, 38(4), 557–574
- Glaser, R (1963) Instructional technology and the measurement of learning outcomes: some questions. *American Psychologist*, 18(8), 519–521
- Gordon, S P & Reese, M (1997) High-stakes testing: worth the price? *Journal of School Leadership*, 7(4), 345–368
- Haladyna, T M, Nolen, S B, & Haas, N S (1991) Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(5), 2–7
- Hambleton, R K (1994) The rise and fall of criterion-referenced measurement? *Educational Measurement: issues and practice*, 13(4 Winter), 21–26
- James, M (2006) Assessment, teaching and theories of learning. In J Gardner (Ed), *Assessment and Learning*. London: Sage, 47–60
- Jonsson, A (2010) The use of transparency in the "interactive examination" for student teachers. *Assessment in Education: principles, policy & practice*, 17(2), 183–197

- Kellaghan, T, Madaus, G F, & Airasian, P W (1982) *The Effects of Standardized Testing*. London: Kluwen Nijhoff Publishing
- Kleinmann, M, Kuptsch, C, & Köller, O (1996) Transparency: a necessary requirement for the construct validity of assessment centres. *Applied Psychology: an international review*, 45(1), 67–84
- Madaus, G F (1988) The influence of testing on the curriculum. In L Tanner (Ed), *Critical Issues in Curriculum: eighty-seventh yearbook of the National Society for the Study of Education (part 1)*. Chicago: University of Chicago Press
- Madaus, G F, Russell, M K, & Higgins, J (2009) *The Paradoxes of High Stakes Testing: how they affect students, their parents, teachers, principals, schools, and society*. Charlotte, NC: Information Age Publishing Inc
- Marton, F & Morris, P (2002) What matters? Discovering critical conditions of classroom learning. In F Marton & P Morris (Eds), *What Matters? Discovering Critical Conditions of Classroom Learning. Volume 181*. Göteborg: Acta Universitatis Gothoburgensis
- Marton, F & Säljö, R (1976) On qualitative differences in learning 2: outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46, 115–127
- McDonald, B & Boud, D (2003) The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. *Assessment in Education: principles, policy & practice*, 10(2), 209–220
- McNeil, L M & Valenzuela, A (2001) The harmful impact of the TAAS system of testing in Texas: beneath the accountability rhetoric. In G Orfield & M Kornhaber (Eds), *Raising Standards or Raising Barriers? Inequality and High-Stakes Testing in Public Education*. New York: Century Foundation Press, 127–150
- Messick, S (1996) Validity and washback in language testing. *Language Testing*, 13(3), 241–256
- Moss, P, Pullin, D, Gee, J P, Haertel, E, & Jones Young, L (Eds) (2008) *Assessment and Opportunity to Learn*. New York: Cambridge University Press
- Murphy, P (Ed) (1999) *Learners, Learning and Assessment*. London/Buckingham: Paul Chapman/Open University Press
- Murphy, R, Stobart, G, Baird, J-A, & Winkley, J (2012) *Investigating the Predictability of GCSE Examinations*
- O'Neill, O (2002) A question of trust. *The BBC Reith Lectures*. Retrieved from <http://www.bbc.co.uk/radio4/reith2002/lectures.shtml>
- OECD (2012) *Preparing Teachers and Developing School Leaders for the 21st Century: lessons from around the world*. Paris
- Ofqual (2008) *Predictability Studies Report on GCSE and GCE Level Examinations*. Coventry
- Pearson, I (1988) Tests as levers for change. In D Chamberlain & R Baumgardner (Eds), *ESP in the Classroom: practice and evaluation*. Modern English Publication, 98–128
- Phelps, R (2012) The effect of testing on student achievement, 1910–2010. *International Journal of Testing*, 12, 21–43
- Popham, J (1987) Two-plus decades of educational objectives. *International Journal of Educational Research*, 11(1), 31–41
- Pryor J & Crossouard, B (2008) A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34(1), 1–20

- Qi, L (2007) Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: principles, policy & practice*, 14(1), 51–74
- Ramsden, P (1988) *Improving Learning: new perspectives*. New York: Nichols Pub Co
- Sadler, D R (2007) Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: principles, policy & practice*, 14(3), 387–392
- SCAA/Ofsted (1996) *Standards in public examinations 1975 to 1995*. A report on English, mathematics and chemistry examinations over time. London: School Curriculum and Assessment Authority, ISBN 1 85838 215 7
- Sfard, A (1998) On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(4), 4–13
- Sfard, A (2008) Participationist discourse on mathematics learning. In P. Murphy and K. Hall (Eds) *Learning and Practice: agency and identities*, London: SAGE, 120–132.
- Shepard, L (1990) Inflated test score gains: is it old norms or teaching the test? *Educational Measurement: issues and practices*, 9(3), 15–22
- Shepard, L (1993) The place of testing reform in educational reform: a reply to CiZek. *Educational Researcher*, 22, 10–14
- Shepard, L, Hammerness, K, Darling-Hammond, L, Rust, F, Baratz Snowden, J, Gordon, E, Gutierrez, C, et al (2005) Assessment. In L Darling-Hammond & J Bransford (Eds), *Preparing Teachers for a Changing World: what teachers should learn and be able to do*. San Francisco: John Wiley & Sons Inc, 275–321
- Smith, M (1991) Meaning of test preparations. *American Educational Research Journal*, 28(3), 521–542
- Smyth, E (2009) Buying your way into college? Private tuition and the transition to higher education in Ireland. *Oxford Review of Education*, 35(1) 1–22
- Spratt, M (2005) Washback and the classroom: the implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29
- State Examinations Commission (2012) Draft report of the SEC working group on predictability in the leaving certificate examination
- Strathern, M (1997) “Improving ratings”: audit in the British university system. *European Review*, 5, 305–321
- Sturman, L (2003) Teaching to the test: science or intuition? *Educational Research*, 45(3), 261–273
- Torrance, H (2007) Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: principles, policy & practice*, 14(3), 281–294
- Vernon, P (1956) *The Measurement of Abilities* (2nd edn) London: University of London Press
- Warmington, P & Murphy, R (2004) Could do better? Media depictions of UK educational assessment results. *Journal of Educational Policy*, 19(3), 285–299
- Watanabe, Y (1996) Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318–333
- Watanabe, Y (2000) Washback effects of the English section of Japanese entrance examinations on instruction in pre-college level EFL. *Language Testing Update*, 27, 42–47
- West, R & Crighton, J (1999) Examination reform in central and eastern Europe: issues and trends. *Assessment in Education: principles, policy & practice*, 6(2), 271–289

- Wikström, C (2005) Criterion-referenced measurement for educational evaluation and selection. *Assessment in Education: principles, policy & practice*, 12(2), 125–144
- Wiliam, D (1996) Standards in examinations: a matter of trust? *The Curriculum Journal*, 7, 293–307
- Wolf, A (2002) *Does Education Matter? Myths About Education and Economic Growth*. London: Penguin

Appendix A: Irish media analysis on allegations of predictability in 2012

In Appendix A and Appendix B, we depict the media interest in the Leaving Certificate examinations in Ireland to give background information on how they are treated in this public forum. This helps to contextualise the research programme, as part of the reason for the research arises from the public perception of predictability in the examinations.

Media interest in the Leaving Certificate examination in Ireland differs from other countries in the volume of coverage and in the detail of the analysis of the examination questions (the topics covered, wording and structure). It is unusual to see so many reviews of question papers on a day-by-day basis by stakeholders and media commentators. Also apparent is an unwritten understanding in the education culture, and in society more generally, that the examinations will not contain major surprises, as commentators used this as a point of evaluation of the examinations. In keeping with the findings from research conducted in other countries, there was some indication that some students intended to rely on strategies involving pre-prepared responses and rote learning of examination materials. Media coverage of the Leaving Certificate examinations reduced, was less focused upon predictability and was more positive in 2013 compared with 2012. Policy announcements about tackling predictability could have impacted upon the media narrative in 2013.

Coverage of the Leaving Certificate in the Irish media had distinctive characteristics compared with an analysis of coverage of A-level examinations in England (Warmington and Murphy, 2004). The content of the English coverage was on students' experiences, pass rates, Higher Education entry, standards, hard or easy subjects, examination reforms and examination boards. Generally, the coverage was less detailed and less focused upon particular examination subjects and questions in England. Warmington and Murphy (2004) concluded that much could be done to influence the narratives about examinations presented in the media. As important life events, they have a large human interest value and negative stories can create scandalisation and interest. Whether or not the media narratives have basis in fact, there needs to be public engagement with them if trust in the examinations is to be upheld in the long term. Notably, media use of the term predictability took the meaning to be positive in some coverage and negative in other articles

A search of mainstream online Irish news outlets was conducted using the terms 'leaving certificate' and 'leaving cert'. The search was narrowed to the period 1 April 2012 – 15 September 2012 to focus on the examination period, when there was the most active media interest in the Leaving Certificate. Articles relevant to predictability and related issues were saved on a social bookmarking site, *delicious.com*, and tagged to indicate the main topics and actors in the article. A total of 238 sites were bookmarked, with the most frequent tags being 'good exam' (68 sites), 'predictability' (58 sites), 'bad exam' (43 sites), 'students' (40 sites), and 'maths' (35 sites). The news outlets searched were: *thejournal.ie*, *irishtimes.com*, *independent.ie*, *herald.ie*, *thestar.ie* and *irishexaminer.com*. Also, a small number of blogs and other sites found through links in the newspapers, Google and YouTube were included.

Public examinations appear to be of great media interest in Ireland. During the exam period, *The Irish Times*, the *Irish Independent*, the *Irish Examiner* and *thejournal.ie* all publish articles reviewing each of the major Leaving and Junior Certificate exams. The articles reviewing each of the major

Leaving and Junior Certificate exams offered question-by-question breakdowns, and were based around quotations from students, teachers, teachers at 'grinds'⁴ companies and representatives of subject teacher associations. We give counts relating to particular codes, but it is obvious that media representation cannot be taken to reflect the views of stakeholders generally and there are sometimes different views about the same examination paper expressed in the media.

Positivity about predictability

Referring to such reviews, and in the context of the NCCA (National Council for Curriculum and Assessment) and HEA (Higher Education Authority) report⁵ released in September 2011, Ruairí Quinn, Education Minister, is quoted in the *Irish Independent* (Donnelly, 2011), as follows:

When the exams come out for both the Junior Cert and the senior Leaving Cert the commentary from education commentators frequently is, 'It was a good exam, there were no surprises, it was as predicted'...

An initial content analysis of the exam reviews does largely support the equating of predictable with 'good'. In 68 articles from summer 2012, where exams were reviewed as largely 'good', five mentioned that there were 'no surprises', two referred to 'few surprises' and one to 'no nasty surprises'. Three articles presented the exams in a positive light as they were consistent with past papers. One review praised the conventional format of the examination. Another article highlighted the predictable range of questions, while other reviewers praised the inclusion of 'key' or 'popular' topics. The range of topics tested was also mentioned as a positive feature, with six articles mentioning 'a broad range' and nine mentioning either 'a lot' or 'plenty' of choice. Three articles referred to 'a variety of topics'.

The adjective 'nice' was used frequently (9 times) to describe question papers; 'manageable' was also used frequently (7 times) as were 'do-able' (6 times), 'straightforward' (6 times), 'accessible' (5 times), 'appropriate' (4 times) and 'student-friendly' (3 times). 'Fair' was used 23 times to describe 'good' exams. The level of challenge was also associated with whether an exam was presented as 'good': one paper was described as 'challenging but fair', one as 'nothing too taxing', one as 'nothing too tricky', two as 'easy' and one as 'simple enough'.

Fairness was sometimes equated with the clarity of questions, with five comments relating to this, and three praising questions for being 'well structured' or 'well scaffolded'. 'Good questions' were 'interesting' (6) and 'relevant' (4). Four reviews of 'good' exams also mentioned that papers gave a 'good balance of theory and practice' or 'theory and application'.

Examinations, but not as we know them

Examinations were criticised in 43 newspaper review articles of the 2012 summer session. Where reviews were negative, comparisons were often made with previous years, both implicitly and explicitly. Exams were described as 'tricky' (2), 'difficult' (4) and 'tough' (2). When exams did not meet expectations, more emotive language was used. The absence of the expected poets on English Paper 2 was considered a 'betrayal' by a student writing for a national newspaper, and the Maths

⁴ This is the term used for crammer colleges and/or private tutoring.

⁵ http://www.transition.ie/files/HEA_NCCA_Transitions.pdf

papers, introduced as part of Project Maths,⁶ were considered ‘disastrous’, ‘traumatic’, ‘one of the worst in the history of the State’, and prompted student ‘outrage’. Two articles commented on the new Maths papers being ‘quite different to’ or ‘not reflective of’ the sample paper.

In some cases, articles implied that there was a change in the style of questions in 2012 in response to the political announcements about predictability; this is presented largely positively in subjects where the shift was considered to be slight (*The Irish Times*, 2012; Holden, 2012a, 2012b; Faller, 2012; *Irish Independent*, 2012b). For example, one article (Donnelly, 2012) claimed that

the trend towards requiring Leaving Certificate candidates to apply their knowledge rather than simply regurgitate what they could learn off was evident in today’s Business Higher Level paper.

Nonetheless, later in the article, a ‘grinds’ teacher is quoted as judging it as a ‘good paper with lots of choice, no nasty surprises’, noting that while one question asked students to make an unexpected comparison in their calculations, another had come up for the fifth consecutive year.

English Paper 2 in 2012

The Irish Times reported that teachers saw English Paper 2 as a ‘game changer’, with one anonymous teacher quoted as saying,

the days of the pre-cooked, single transferable response to questions on the English paper are over. This was a much more searching exam. (Flynn, 2012)

Much coverage was given to student reaction to this exam, as in the final poetry section, where four questions are posed from eight possible poets, Seamus Heaney and Sylvia Plath did not appear. A student writing for the *Irish Independent* described the choice of poets as ‘betrayal’, and the Leaving Certificate examinations as a ‘battle’. She (McGirr, 2012) described it as fortunate that one of her prepared poets did appear on the paper, writing that

otherwise the examiner would have gotten a Heaney essay whether they wanted one or not.

Other student writers also saw the selection of poets as unfair (*The Irish Times*, 2012; Freeman, 2012). An unpredictable exam here is seen as a bad exam as it confounds the preparation tactic by some students of pre-prepared essay plans.

Maths and Predictability

Similarly, the Maths exams received negative reviews in the press, and the lack of predictability was an issue. Before the exams, the lack of a corpus of past papers was seen as a problem for exam preparation. In May, one student blogger (Kelleher, 2012) wrote that

we haven’t really got a clue what to expect. Anything on the entire course could come up, which is certainly making things more difficult for us.

⁶ <http://www.projectmaths.ie>

With regard to the examinations themselves, one student (MacSweeney, 2012) referred to Maths Paper 1 as leaving students 'distraught', writing that

Most of the classmates I spoke to had 'never seen a paper like it'. I found there were a few unusual questions I really had to think about.

However, reviews from teachers were more balanced. While the coverage highlighted concerns from teachers about the wording of the questions, responses from teachers on the content of questions was more positive (Holden, 2012a; Murray, 2012).

Question difficulty

Where individual questions were highlighted for criticism, even as part of a 'good' paper, they too were often considered difficult in some way: 13 mentions of difficulty include 'unfairly difficult', 'extremely difficult', 'too taxing' and 'too demanding'. This was sometimes linked to the question being 'unusual' (2). For instance, the German exam included an applied grammar task that was considered unfamiliar, which was reported to have caused alarm among students.

A key criticism of questions was also the wording and the level of vocabulary used (18 mentions). This was seen to affect question clarity and also difficulty. For example, 'some of the language used on the ordinary paper was "slightly unfair"'.

Curriculum and revision load

Some articles voiced concern about curriculum load and whether the volume of material tested in the terminal exams was manageable (O'Brien, 2012; *Irish Independent*, 2012a). A student blogger expressed this concern about Project Maths (Kelleher, 2012). In an opinion piece, Derry Cotter (2012), a university lecturer, argued that

the time teachers can devote to higher-order learning is limited by the pressures of syllabus coverage.

Similarly, in another column, Denise Deegan (2012) expressed concern about her daughter's workload and advocated cutting down the curriculum to leave more time for enjoyment and critical thinking.

In the sources analysed here, predicting topics and questions was seen by some as a tactic that allows students to reduce the revision load. A student who achieved six A1s advised others to use past papers to narrow the revision focus as 'you can't learn everything' (Faulkner, 2011). Similarly, a teacher writing in *The Irish Times* (O'Brien, 2012) claimed that

however unwisely, students begin to gamble on certain questions because they feel overwhelmed by the sheer volume of work.

This leads to the 'consternation' when a widely predicted topic or poet does not appear on the paper (ibid).

Rote learning of examination materials

However, some articles featured students reporting their use of past papers not just for prediction, but as rote learning material. Writing an article on how to prepare for the exams, one student (Crean, 2012) advised compiling ‘collections of past papers’ not just in order to familiarise oneself with the format and task styles but also

the repetition of the more common topics [which] will mean that if they do crop up again, you’ll know the subject inside out.

The student goes on to assert that

the best way to learn anything is through repetition, and making your way through the past papers is pure repetition. (Crean, 2012)

Similar tactics are recommended elsewhere (Faulkner, 2011) and this is in keeping with the finding by Daly et al (2012) in England.

The Grinds Industry

A key source of supplementary notes, exam guidance and practice materials is the private tuition sector. A survey conducted with nearly 1,500 students in Ireland concluded that 45% had received private tutoring in their last year of school and this had risen from 32% in 1994 (Smyth, 2009). These figures were higher than those in countries such as France, Germany and England (European Commission, 2011, pp22–23). The grinds market is reported overall as being worth between €20 and €50 million per year (Murphy, 2012b; Newenham, 2013), although a report in the mid-market tabloid the *Evening Herald* suggested that some parents were cutting back on tuition as a result of the recession (Murphy, 2012a). Online grinds services are reportedly becoming more popular, with *The Irish Times* recently running a review of grinds sites (Holden, 2013). Smyth (2009) noted that this commodification of education did not appear to advantage students in terms of their Leaving Certificate examination grades, once she controlled for the effects of social characteristics, prior performance and attitudes to school.

Public perceptions of the predictability of examinations might allow grinds providers to flourish, by conveying the message that they have the key to unlocking more advantageous pathways to examination success. The owners of *mocks.ie* do not have a background in education or assessment (Newenham, 2013). They launched the website in 2009, offering study notes devised with the help of two students who had achieved A grades in their Leaving Certificates. The website now offers the services of teachers who also work as examiners for the State Examinations Commission to offer marking and feedback services and supplies schools with mock papers (ibid), as do other private companies. An advantage of this process, according to an article in the *Irish Independent* (Bielenberg, 2013), is that the mock tests are not biased by the class teacher, as

the teacher cannot pack the test with questions that have been well covered in class.

Media sources cited

- Bielenberg, K (2013, 23 January) Don't knock the mocks, and don't keel over with shock. Irish Independent. Retrieved from <http://www.independent.ie/lifestyle/education/dont-knock-the-mocks-and-dont-keel-over-with-shock-29022429.html>
- Cotter, D (2012, 16 May) In my opinion: first-year students entering third level have been taught to learn – not to reason. Irish Independent. Retrieved from <http://www.independent.ie/lifestyle/education/in-my-opinion-firstyear-students-entering-third-level-have-been-taught-to-learn-not-to-reason-26854569.html>
- Crean, R (2012, 10 April) Your leaving cert: 57 days to go. The Irish Times. Retrieved from <http://www.irishtimes.com/news/education/your-leaving-cert-57-days-to-go-1.498807?page=1>
- Deegan, D (2012, 27 November) We need to teach our teenagers how to chill. Irish Independent. Retrieved from <http://www.independent.ie/opinion/analysis/denise-deegan-we-need-to-teach-our-teenagers-how-to-chill-28940476.html>
- Donnelly, K (2011, 22 December) Leaving Cert set to become tougher in system overhaul. Irish Independent. Retrieved from <http://www.independent.ie/lifestyle/education/leaving-cert-set-to-become-tougher-in-system-overhaul-26804720.html>
- Donnelly, K (2012, 14 June) Leaving Cert 2012: new trend evident in Business Higher Level paper – teachers. Irish Independent. Retrieved from <http://www.independent.ie/lifestyle/education/leaving-cert-2012-new-trend-evident-in-business-higher-level-paper-teachers-26864750.html>
- Faller, G (2012, 18 June) Reality stars in Physics papers. The Irish Times. Retrieved from <http://www.irishtimes.com/news/reality-stars-in-physics-papers-1.720544>
- Faulkner, W (radio presenter) [Midlands103Official] (2011, 1 June) How to get six A1s in your Leaving Cert. [Audio file] Retrieved from http://www.YouTube.com/watch?v=LkOBxqw_XaY
- Flynn, S (2012, 8 June) Black day for poetry students as Plath question fails to come up. The Irish Times. Retrieved from <http://www.irishtimes.com/news/black-day-for-poetry-students-as-plath-question-fails-to-come-up-1.1064621>
- Freeman, M (2012, 8 June) 'Worst day ever': Leaving Cert students react to English Paper Two [Online news article]. Retrieved from <http://www.thejournal.ie/worst-day-ever-leaving-cert-students-react-to-english-paper-two-478829-Jun2012/#slide-slideshow1>
- Holden, L (2012a, 11 June) Higher Maths gets mixed response. The Irish Times. Retrieved from <http://www.irishtimes.com/news/higher-maths-gets-mixed-response-1.719329>
- Holden, L (2012b, 15 June) 'Good and honest' paper with some 'little twists'. The Irish Times. Retrieved from <http://www.irishtimes.com/news/good-and-honest-paper-with-some-little-twists-1.1066289>
- Holden, L (2013, 5 March) The new daily grinds: online and on your terms. The Irish Times. Retrieved from <http://www.irishtimes.com/news/education/the-new-daily-grinds-online-and-on-your-terms-1.1317532>
- Irish Independent (2012a, 21 June) Broad view is 'par for the course'. Retrieved from <http://www.independent.ie/irish-news/broad-view-is-par-for-the-course-26867512.html>
- Irish Independent (2012b, 22 June) Leaving Cert 2012: religious education paper a shift in style, on reflection. Retrieved from <http://www.independent.ie/lifestyle/education/leaving-cert-2012-religious-education-paper-a-shift-in-style-on-reflection-26867781.html>

- Kelleher, P (2012, 20 May) Project Maths – what do you think? [Blog entry] Retrieved from <http://leavingcertdiaries.wordpress.com/2012/05/20/project-maths>
- MacSweeney, J (2012, 9 June) Game of Russian roulette? I prefer Ultimate Frisbee. Irish Examiner. Retrieved from <http://www.irishexaminer.com/archives/2012/0609/world/game-of-russian-roulette-i-prefer-ultimate-frisbee-196771.html>
- McGirr, I (2012, 8 June) Blood will be spilled; betrayal like that comes at a high price. Irish Independent. Retrieved from <http://www.independent.ie/lifestyle/education/blood-will-be-spilled-betrayal-like-that-comes-at-a-high-price-26862548.html>
- Murphy, C. (2012a, 16 August) How a lot of parents have lost faith in grind schools. Evening Herald. Retrieved from <http://www.herald.ie/news/how-a-lot-of-parents-have-lost-faith-in-grind-schools-28015346.html>
- Murphy, C (2012b, 4 September) Quinn to target €20m spent on grind schools. Evening Herald. Retrieved from <http://www.herald.ie/news/quinn-to-target-20m-spent-on-grind-schools-28016805.html>
- Murray, N (2012, 12 June) Fair Maths papers should leave most happy. Irish Examiner. Retrieved from <http://www.irishexaminer.com/archives/2012/0612/world/fair-maths-papers-should-leave-most-happy-197048.html>
- Newenham, P (2013, 17 January) Figures add up for exam preparation website. The Irish Times. Retrieved from <http://www.irishtimes.com/news/education/figures-add-up-for-exam-preparation-website-1.962418>
- O'Brien, B (2012, 18 August) Leaving Cert's core issues still not confronted. The Irish Times. Retrieved from <http://www.irishtimes.com/debate/leaving-cert-s-core-issues-still-not-confronted-1.539119>
- The Irish Times (2012, 7 June) Two exams down. Timing abominable, but sure we'll see. Retrieved from <http://www.irishtimes.com/news/two-exams-down-timing-abominable-but-sure-we-ll-see-1.1064403>

Appendix B: Preliminary analysis of media coverage 2013

Search methods

A search was undertaken for the terms 'leaving certificate' and 'leaving cert' on the mainstream media websites of the *Irish Examiner*, *The Irish Times*, the *Irish Independent* and *thejournal.ie*. The date range was 1 May 2013 to 1 September 2013. In total, 86 articles were imported into NVivo for further analysis and then coded inductively.

Discussion of initial results

The majority of the articles analysed presented reviews of particular examination papers. The examination papers sat in 2013 seem to be presented in a less controversial light than in 2012, when English and Maths in particular drew criticism from commentators. These two subjects will be reviewed in further detail below. The majority of examinations were presented in a positive light by commentators; only five articles focused predominantly on the negative reception of papers or parts of papers.⁷

Fairness

In 34 of the articles, examination papers were described as 'fair', as seen in the example below, in which Art History is commented upon:

Students were very happy with the paper, described as fair but challenging. The Iron Age came up as predicted, and there were no trick questions. (Donnelly, 2013c)

In order to examine the relationship between fairness and other themes, a coding matrix was run. Table A1 below presents the most frequent results of co-occurrence of 'fair' with other codes within the same reference.

⁷ Chemistry ordinary level, Irish paper 2 ordinary level, Applied Maths, Religious Education, Spanish, Home Economics.

Table A1: Coding matrix: nodes most commonly appearing in the same reference as ‘fair’

	Frequency of co- incidence with ‘fair’ in same reference	Notes
Accessible or approachable	5	
Choice of questions	4	
Level of challenge – negative	3	Used where commentators describe one item as overly challenging although the rest or most of the paper was ‘fair’.
Level of challenge – welcome	11	This is somewhat skewed by the newspapers recycling the same commentator quotations – if these are removed, the frequency is 9.
Predicted or expected	5	
Range of question or paper	3	
Relevant	5	

The most frequent co-occurrence was ‘fair’ with comments that described the level of challenge presented by the paper in a positive light. For example, the French examination was presented as follows:

‘The ordinary level paper was fair and well pitched,’ said Ms Ní Chiarba, ‘with topics including festivals, French culture and celebrities.’ (McGuire, 2013a)

The same paper was described elsewhere as ‘approachable’, a term that, similarly to ‘accessible’, seems to imply an appropriate level of challenge:

An approachable and fair paper that offered students a good choice of questions on a range of interesting, relevant and topical subjects, was how one teacher described Leaving Certificate French Higher Level. (Donnelly, 2013b)

Combining both codes, 16 reviews that refer to examinations as fair also consider the level of challenge as appropriate. This is almost half of the reviews that refer to examinations as ‘fair’ (34).

Level of challenge and predictability

In some articles, whether the level of challenge presented by a paper was appropriate was linked to whether it met the commentator’s expectations based on past papers. A coding matrix query was run to investigate the correlation of codes pertaining to level of challenge with those pertaining to predictability (Table A2).

Table A2: Coding matrix of difficulty with predictability

	1: Accessible or approachable	2: Level of challenge – negative	3: Level of challenge – neutral	4: Level of challenge – welcome
A: Catching students out	0	1	0	0
B: Deviation from the usual	0	5	0	2
C: No surprises	0	1	0	0
D: Positive unpredictability	1	0	1	1
E: Predicted or expected	1	1	0	4
F: Straightforward	2	2	1	2
G: Surprise	1	0	1	0

Five sources contained references which were coded for some kind of deviation from what the commentator saw as usual, and also an inappropriate level of challenge. For instance, ordinary level Irish Paper 2 was directly compared to the 2012 examination:

However, Ms Grealy described the ordinary level paper as ‘tough and quite challenging for weaker students’ (...) ‘This was a lot more challenging than last year’s paper and some of the questions, especially in the Prose section, were unpredictable,’ she said. (Donnelly, 2013a)

Here, unpredictability appears to be associated with an unreasonable level of demand for ‘weaker’ students.

Four sources contained references which were coded for both predictability or expectedness and a welcome or appropriate level of challenge. This was the case for Maths Paper 1, which examined the Project Maths curriculum. In 2012, the first year of Project Maths, the examinations were poorly reviewed in the media and perceived to be difficult. One student referred to Maths Paper 1 as leaving students ‘distracted’ and stated that the questions were ‘unusual’ (MacSweeney, 2012). One student blog (Kelleher, 2012) stated that the lack of a corpus of past papers was seen as a problem for exam preparation:

We haven’t really got a clue what to expect. Anything on the entire course could come up, which is certainly making things more difficult for us.

In contrast, Paper 1 in 2013 was reviewed positively. One detailed review noted the consistency of parts of the paper with sample papers, implying that predictability ensures an appropriate level of challenge:

Aidan Roantree, senior maths teacher at the Institute of Education, said that higher-level students will have entered the exam with apprehension and dread but will have been very relieved by what they saw. **‘All the questions in section A were fully**

anticipated and wouldn't have posed any difficulty for students who were prepared. The questions in section B were similar to those posed in the sample paper provided during the year; this would have taken the sting out of them for students. Section C, meanwhile, was **much easier than previously, with no challenging integration whatsoever.**' (McGuire, 2013b)

Here, similarity with the sample paper ensures that the level of challenge is appropriate as students' preparations towards teacher 'anticipated' questions is rewarded. Predictability is therefore seen to be a positive feature of the examination.

English 2013

Similarly to Maths, English Paper 2 was poorly reviewed in 2012 but reviewed more positively in 2013. Sixteen sources in the search period referred to English. Before the examination, one student writer explained the relationship between the 'notorious' English Paper 2 and predictability:

Every year it promises to break some hearts through the poets it does, or doesn't examine. It has become de rigueur for Leaving Certs to try predicting which poets will appear on Paper 2. (Gaynor, 2013)

Most reviews made a comparison to the 2012 paper, in which the popular poet Sylvia Plath did not come up. One news source which otherwise did not carry reviews of exam papers in 2013 ran a short 'just so you know' article informing readers that 'Sylvia Plath *did* appear' on Paper 2 Higher (*thejournal.ie*, 2013; original emphasis). A more detailed review of the paper opened as follows:

A 'delightful, dream paper' and the 'most settled English Paper 2 for many years' were among the comments of teachers to what Leaving Certificate higher-level candidates were treated to yesterday afternoon. Even Sylvia Plath obliged. (Donnelly, 2013d)

Similarly, another review welcomed the predicted and expected nature of both the choice of poets and other parts of the examination:

'There were no surprises, in either the selection of poets or the questions set: a huge relief for everyone,' said teacher Jim Lusby of the Institute of Education (...) The majority of candidates studied the Shakespeare play *Macbeth*, and the expected focus on character and language appeared on the paper. (Holden, 2013)

The reviews of the examination were highly positive, with few negative comments like the one below:

On the comparative study question, Mr Lusby said examiners should give more thought to the phrasing of the questions. (Donnelly, 2013d)

However, the tone of the reviews was overwhelmingly positive as most of the paper complied with commentators' predictions and expectations.

References

- Donnelly, K (2013a, 11 June) Leaving Cert Irish Paper II: easy task for honours students but ordinary level a challenge. *Irish Independent*. Retrieved from: <http://www.independent.ie/lifestyle/education/exams/leaving-cert-irish-paper-ii-easy-task-for-honours-students-but-ordinary-level-a-challenge-29336560.html>
- Donnelly, K (2013b, 12 June) Leaving Cert Higher French deemed 'approachable and fair'. *Irish Independent*. Retrieved from: <http://www.independent.ie/lifestyle/education/exams/leaving-cert-higher-french-deemed-approachable-and-fair-29340060.html>
- Donnelly, K (2013c, 13 June) LC Art History: Iron Age put smiles on their faces. *Irish Independent*. Retrieved from: <http://www.independent.ie/lifestyle/education/exams/lc-art-history-iron-age-put-smiles-on-their-faces-29343080.html>
- Donnelly, K (2013d, 7 June) Leaving Cert English Higher Paper II: 'a dream paper – the most settled for many years.' *Irish Independent*. Retrieved from: <http://www.independent.ie/lifestyle/education/exams/leaving-cert-english-higher-paper-ii-a-dream-paper-the-most-settled-for-many-years-29326927.html>
- European Commission (2011) *The Challenge of Shadow Education: Private Tutoring and its implications for policy makers in the European Union*, An independent report prepared for the European Commission by the NESSE network of experts.
- Gaynor, L (2013, 6 June) Exam diary: I left the exam hall fairly exhausted – but happy. *Irish Independent*. Retrieved from: <http://www.independent.ie/lifestyle/education/exams/exam-diary-i-left-the-exam-hall-fairly-exhausted-but-happy-29324331.html>
- Holden, L (2013, 7 June) Sigh of relief as Sylvia Plath returns in triumph for higher level English paper 2. *The Irish Times*. Retrieved from: <http://www.irishtimes.com/news/education/sigh-of-relief-as-sylvia-plath-returns-in-triumph-for-higher-level-english-paper-2-1.1419765>
- Kelleher, P (2012, 20 May) Project Maths – what do you think? [Blog entry] Retrieved from <http://leavingcertdiaries.wordpress.com/2012/05/20/project-maths/>
- MacSweeney, J (2012, 9 June) Game of Russian roulette? I prefer Ultimate Frisbee. *Irish Examiner*. Retrieved from <http://www.irissexaminer.com/archives/2012/0609/world/game-of-russian-roulette-i-prefer-ultimate-frisbee-196771.html>
- McGuire, P (2013a, 12 June) Rote learning hits a cul-de-sac for Leaving Cert French. *The Irish Times*. Retrieved from: <http://www.irishtimes.com/news/education/rote-learning-hits-a-cul-de-sac-for-leaving-cert-french-1.1426183>
- McGuire, P (2013b, 8 June) Students overjoyed with maths paper but teachers fear dumbing down. *The Irish Times*. Retrieved from: <http://www.irishtimes.com/news/education/students-overjoyed-with-maths-paper-but-teachers-fear-dumbing-down-1.1421237>
- www.thejournal.ie (2013, 6 June) Just so you know. [Online news article] Retrieved from: <http://thescore.thejournal.ie/just-so-you-know/news/>

