

Institute of Education,

Dublin City University

Irish Post-Primary Teachers' Conceptions of Assessment

Dissertation

by

Marie Darmody

B.A., M.A.

Co-Supervisors:

Dr. Zita Lysaght

Professor Michael O'Leary

Submitted to the Institute of Education, DCU,

in Partial Fulfilment of the Requirements for the Award of Doctor of Education.

September 2017

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Education is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

ID No: _____

Date: _____

Acknowledgements

As I near the end of my doctoral experience, I would like to thank all those who have supported me in this very challenging, but incredibly rewarding, journey.

Firstly, to my supervisors Dr. Zita Lysaght and Professor Michael O’Leary, I extend my deepest gratitude. Without your incredible expertise, continuous guidance and invaluable feedback, this endeavour would not have been possible. Thank you both for being true mentors and for showing me great kindness every step of the way. I think we made a brilliant team.

I would like to thank my husband Kevin for his great patience, critical insight and unwavering belief in my ability to succeed. For that, I am truly grateful.

To my parents, Cyril and Anne, and my brother, William, thank you for your endless support and love throughout this journey. Without your continuous understanding, this venture would not have been possible. You will now get to see a lot more of me!

I extend my thanks to my fellow doctoral students on the LTAC strand. The laughter and friendship we shared was a true highlight of the EdD experience.

Finally, to my wonderful friends and colleagues, thank you for your endless messages of support throughout the last four years. Without them, this would have been a far more isolated and difficult experience.

TABLE OF CONTENTS

Abstract.....	5
Chapter 1: A Statement of the Research Context and Problem.....	1
Introduction.....	1
Organisation of the Chapter	3
International Assessment Context	3
Junior Cycle Policy Reform in Ireland	7
Teachers' Beliefs	14
Research Question	18
Terminology.....	18
Summary.....	19
Organisation of the Thesis	20
Chapter 2: Literature Review	22
Introduction.....	22
Organisation of the Review	22
Teachers' Beliefs	23
Nature of Teachers' Beliefs	23
Defining Teachers' Beliefs.....	23
Contents and organisation of teachers' beliefs.....	24
Importance of context in shaping teachers' beliefs.....	25
Role of Teachers' Beliefs	27
Filtering	27
Framing	28
Guiding.....	29
Internal and External Factors Influencing Belief Enactment.....	31
Belief Development and Change	34
Belief change through initial teacher education.....	35
Belief change through teaching experience	36
Belief change through professional development support	38
Summary of Section One	41
Teachers' conceptions of assessment	41
Quantitative studies on Teachers' Conceptions of Assessment.....	42
Qualitative Studies on Teachers' Conceptions of Assessment	57
Mixed-Method Studies.....	61
Summary.....	62
Models of Teachers' Engagement with Assessment	65

Chapter 3: Methodology.....	71
Introduction.....	71
Research Design	73
Introduction	73
Timeline for the Study	74
Target Population	74
Instrumentation	75
Survey Approach.....	81
Sampling Design	83
Piloting	85
Ethics	86
Response Data.....	87
Response Rate	87
Participant Demographic Information.....	88
Current role in education.....	88
Years of teaching experience.....	89
Current school type.....	90
Qualifications.	91
Data Cleaning / Preliminary Analysis	92
Summary.....	95
Chapter 4: Data Analysis.....	97
Introduction.....	97
Factor Analysis	97
Suitability of the Data for Factor Analysis	97
Five-Factor Model.....	99
Factor 1: Assessment is a diagnostic and formative tool.	101
Factor 2: Assessment is irrelevant (bad, ignored & inaccurate).....	106
Factor 3: Assessment is a school accountability tool.	108
Factor 4: Assessment is a measurement and categorisation tool.....	109
Factor 5: Assessment is a valid grading tool.	111
Additional Analyses.....	112
Chapter 5: Conclusions.....	118
Introduction.....	118
Analysis of Findings	119
Factor Structure	119
Factor Endorsement Levels and Factor Correlations	121

Conclusions and Implications	128
Limitations	130
Recommendations for Future Research	131
Epilogue	134
References	135
Appendices	i
Appendix A: Correspondence from the Teaching Council.....	i
Appendix B: Example Page of Online Survey Instrument.....	iv
Appendix C: Feedback Sheet.....	v
Appendix D: Correspondence from Professor Gavin Brown	vi
Appendix E: Plain Language Statement.....	viii
Appendix F: Informed Consent Form	x
Appendix G: Range of Subjects Taught by Participants	xii
Appendix H: Mean and Standard Deviation for each statement on the TCoA-III	
	xiii

LIST OF FIGURES

Figure 1. The dual approach to assessment at Junior Cycle	14
Figure 2. A conceptual framework of teacher assessment literacy in practice.	67
Figure 3. Teacher Assessment Identity (TAI) adapted from Looney et al. (2017)	70
Figure 4. Development of a conceptual framework	72

LIST OF TABLES

1. Brown's (2006) Teachers' Conceptions of Assessment Inventory (Abridged):	
Factor and Corresponding statements	80
2. Organisations who helped with the recruitment of participants	85
3. The point of drop out in partially completed questionnaires.....	88
4. Participants' Current Role in Education.....	88
5. Years of Teaching Experience.....	89
6. Participants' Current School Type	90
7. Qualifications held by participants	91
8. Percentage frequencies for each statement on the TCoA-III A.....	93
9. Comparison of eigenvalues from PCA and criterion values from parallel analysis.	99
10. Total Variance Explained for	
Exploratory Factor Analysis with Oblique Rotation	102
11. Factor Loadings for	
Exploratory Factor Analysis with Oblique Rotation of TCoA-III A	103
12. Communality Values for the TCoA-III A variables following	
Exploratory Factor Analysis with Oblique Rotation	104
13. Mean Factor Endorsement Rankings and Cronbach Alpha Reliabilities	113
14. Factor Correlation Matrix.....	113
15. Factor Mean Differences by Gender	114
16. Factor 1: Statistically significant mean differences by subject	116
17. Factor 2: Statistically significant mean differences by subject	116
18. Factor 3: Statistically significant mean differences by subject	117
19. Factor 4: Statistically significant mean differences by subject	117
20. Factor 5: Statistically significant mean differences by subject	117

Abstract

Title: Irish Post-Primary Teachers' Conceptions of Assessment

Author: Marie Darmody

The purpose of this research was to elicit baseline data about Irish post-primary teachers' conceptions of assessment. Post-primary education in the Republic of Ireland is currently in the midst of significant curriculum and assessment reform at Junior Cycle, the first three years of the secondary school system. Central to this change is the positioning of the teacher at the heart of the assessment process. The successful implementation of the new assessment practices will not only require a high level of teacher assessment literacy, but will also depend upon the extent to which teachers' conceptions of assessment align with the philosophical underpinnings of the reform. Research has indicated that teachers' beliefs serve to filter information entering the cognitive domain, to frame particular educational situations or problems and to guide teachers' intentions and actions (Fives & Buehl, 2012). In light of this evidence, the introduction of new assessment initiatives should take account of how teachers conceive of the nature and purpose of assessment. Adopting a non-experimental cross-sectional design, this study surveyed a large sample ($n=489$) of post-primary teachers using the abridged version of Brown's (2006) Teachers' Conceptions of Assessment Inventory (TCoA-IIIa). This 27-item self-report instrument is designed to elicit teachers' level of agreement with four intercorrelated assessment factors (i.e., school accountability, student accountability, improvement and irrelevance). Quantitative data derived from the survey were analysed using a mixture of descriptive statistics, exploratory factor analysis, independent samples t-tests and one-way analysis of variance. Maximum likelihood exploratory factor analysis resulted in a 5-factor solution for the Irish data which differed somewhat from Brown's (2006) original model. Implications of the results for the conceptualisation of assessment in the Irish post-primary context are considered.

Chapter 1: A Statement of the Research Context and Problem

Introduction

In a recent article on teachers' perceptions of assessment, Gardner and Galanouli (2016) note that situations of change “expose the type of perceptions harboured by teachers who may, up to that point in their careers, have passively followed routine or conventional practices” (p.717). This contention is particularly pertinent to the Irish educational context¹ which is currently facing the challenge of implementing curriculum and assessment reform at Junior Cycle, the first three years of post-primary (lower secondary) education. The reform measures, originally launched in 2012 with the aim of enabling “post-primary schools to provide a quality, inclusive and relevant education with improved learning outcomes for all students” (Department of Education and Skills [DES], 2012, p.1), were greeted with a much greater level of resistance and disquiet than policy-makers might have anticipated. One of the key issues which teaching unions highly contested was summative assessment by teachers for certification purposes. This type of internal assessment in schools was never before part of the assessment system at post-primary level in Ireland. Central to the reform policy is the positioning of the teacher at the heart of the assessment process through an increased emphasis on formative assessment in addition to the introduction of school-based summative assessment by the teacher. Taking the stance that “teachers should be teachers, not judges” (Association of Secondary Teachers of Ireland [ASTI], 2017a), one of the main teaching unions opposing the reform argued that a school-based assessment component “threatens equity, relationships and consistency of standards”

¹ References to the Irish context in this study refer exclusively to the Republic of Ireland

(ASTI, 2014, p.3). In light of the opposition to the proposed changes, an intense period of multi-stakeholder dialogue and mediation ensued, resulting in the publication of a revised policy framework (DES, 2015a) which was only recently accepted by all teaching unions (ASTI, 2017b). School-based assessment remains within this revised framework but the results of this assessment will no longer be combined with students' results on an externally assessed state-certified examination. As pointed out in a recent commentary by Dr. Damian Murchan, "there will now be two sets of results, two assessment currencies" (Murchan, 2017). Teachers will assess their students at two distinct points throughout the Junior Cycle but the reporting of students' achievement in these assessments will be based on descriptors and will be reported on separately from the students' terminal examination grades. The quality descriptors in question are exceptional, above expectations, in line with expectations and yet to meet expectations (www.juniorcycle.ie). While *Framework for Junior Cycle* (DES, 2015a) has been approved in principle, the success of its implementation will depend to a large extent on teachers' interpretation of, and engagement with the policy. As practitioners, teachers enact personal educational philosophies (Yan & Cheng, 2015) which provide a lens for the translation of policy into practice. Research has indicated that teachers' beliefs serve to filter information and content entering the cognitive domain, to frame particular educational situations or problems and to guide teachers' intentions and actions (Fives & Buehl, 2012). Furthermore, in his work on educational change, Fullan (2007) argues that engagement with teachers' beliefs is a crucial step in the pursuit of long-lasting deep reform. This thesis focuses on teachers' conceptions of assessment, a topic that, heretofore, has received little or no attention in the Irish context. It is argued that to ignore how teachers conceive of the nature and purpose of assessment would be tantamount to negligence in the current Irish context. Hence, the purpose of this

research was to elicit baseline data about Irish post-primary teachers' conceptions of assessment at this critical juncture. It has been acknowledged that teachers will require comprehensive professional development in educational assessment in order to be effectively directed and guided through this change process (DES, 2015a). The successful provision of such support requires a base-line starting point from which to help teachers build assessment capacity. For this reason, an understanding of teachers' conceptions of assessment needs to precede professional development support. As rationalised, large-scale engagement with teachers' conceptions of assessment would provide important evidence-based data, currently absent, for use in policy implementation and professional development design.

Organisation of the Chapter

In the interest of directing the reader through the study, the remainder of this chapter is intended to contextualise the research study and to provide a rationale for its focus. The first part of the chapter outlines the international assessment narrative which has informed and framed Junior Cycle reform in Ireland. Specific attention is then paid to the extent of reform in the Irish context, focusing particularly on the challenge of assessment and the consequential necessity for research on Irish post-primary teachers' conceptions of assessment.

International Assessment Context

Wyse, Hayward and Pandya (2016) argue that “conflicts of purpose for assessment, and the differing value systems that lie behind their enactment, present education systems internationally with what may be the major challenge for curriculum

theory in the twenty-first century” (p.3). The conflicts of purpose in question stem from macro-level forces which exert a strong influence on the shape of learning, teaching, curriculum and assessment worldwide. Central to the shaping of the contemporary assessment narrative are economic competitiveness, performativity and contemporary learning theory. These competing agendas are each underpinned by a particular understanding of what constitutes knowledge and learning and how they should consequently be assessed.

Firstly, as discussed in depth by Ball (2013), the increasing competitiveness of the global economic environment has impacted on the perceived role of education in today’s society. Education is often seen as the means to producing skilled workers who will meet the needs of the knowledge economy, an economy which gives priority to such skills as collaboration, communication, creativity and critical thinking (www.p21.org). This instrumentalist view of education is underpinned by a conceptualisation of knowledge in “terms of its uses not its intrinsic worth” (Yates & Young, 2010, p.6). The drive towards a skills-based approach to success and employability, which prioritises ‘knowledge how’ over ‘knowledge what’ (Winter, 2012), has important implications for curriculum and assessment procedures worldwide. If curricula are to successfully promote and embed the metacognitive and creative capital required by the 21st century learner, they need to be supported by a broad approach to assessment which moves beyond the mere recall of discrete knowledge and supports deeper and more contextualised learning. The requirements of such an assessment approach are currently at odds, however, with the priorities of performativity, a second global trend.

The twenty-first century has seen the emergence of a global movement towards performativity and accountability in education (De Luca & Johnson, 2017; Klenowski

& Carter, 2016; Wyse et al., 2016). Central to this movement are the ideas of standards-based curricula, high-stakes assessment, international testing and the increased scrutiny of teachers and schools. The discourse and assessment programmes of powerful policy agents such as the Organisation for Economic Co-operation and Development (OECD) have been a key driving force in the promotion of this performativity culture. The push to meet targets set out by international assessments such as the Programme for International Student Assessment (PISA) and Trends in International Maths and Science Study (TIMSS) has led to a culture of “constant comparison” (Grek et al., 2009, p.5) among nation states. As countries strive for international rank and global competitiveness, international assessment results are sometimes viewed as the prime indicator of the success or failure of students, teachers, schools, and entire education systems. It has been argued that this view in turn can lead to the narrowing of curriculum and to an increase in the use of testing on a national level (Klenowski & Carter, 2016). Klenowski and Carter (2016) note that “if important decisions are presumed to be related to the results of standardized tests, then schools will focus on the test content” (p.793). It is clear therefore, that external assessment for performativity purposes does not align with the type of assessment required for the acquisition of key skills discussed earlier. Neither does it align with the use of assessment as a pedagogical tool (Smith, 2016). The use of assessment for formative purposes will now be discussed in the context of the third force shaping the international assessment narrative, contemporary learning theory.

It has been argued that developments in learning theory have not been paralleled at the same rate by developments in assessment practices (James & Lewis, 2012; Nusche 2016). Developments in learning theory have broadly moved from a behaviourist view of learning, towards a constructivist understanding of learning which

focuses particularly on “how people construct meaning and make sense of the world by developing mental models” (James & Lewis, 2012). Central also to the discourse on learning theory is the socio-cultural perspective where learning is viewed as a social, collaborative and culturally mediated endeavour (Lave & Wenger, 1991; Rogoff, 1990). Within the constructivist domain, while the promise of formative assessment laid out a pathway to enhance student learning and innovate teaching practice (Black & Wiliam, 1998), the successful implementation of this practice proved more complex than had initially been anticipated (Black, 2015; Hayward, 2015; Hopfenbeck, Flórez Petour & Tolo, 2015). As a transformative practice, formative assessment requires deep learning and engagement, requirements which can be very challenging to reconcile with the demands of external accountability regimes as outlined above. James and Lewis (2012) argue that the development of an assessment for learning culture is an essential stepping stone in meeting the challenge of assessment which aligns with socio-cultural learning theory. In a discussion on assessment through a socio-cultural theoretical lens, Elwood and Murphy (2015) argue that “the social and cultural experiences that students and teachers bring to assessment situations cannot be controlled for or ignored; they are part of the tapestry of the assessment tasks and outcomes” (p.187). James and Lewis (2012) contend that in order for assessment practices to align with this type of learning, they would need to be contextualised and conducted by a member of the learning community rather than by an external assessor. They also argue that such assessment should incorporate collaborative as well as individual assessment, place value on process and agency rather than the outcome alone, capture learning outcomes in various ways and be more qualitative and holistic rather than simply quantitative.

Taken together, the varying priorities of the twenty-first century educational landscape in terms of economic competitiveness, performativity and learning theory represent a complex terrain which needs to be carefully navigated. The interplay between multiple assessment purposes can be problematic and undoubtedly presents a great challenge to educational practitioners. As policy-makers attempt to redesign curriculum and assessment policy in order to harmonise conflicting purposes for assessment (Crooks, 2011; Hutchinson & Young, 2011), teachers are faced with the challenge of reconciling these new policy measures with their current assessment beliefs and practices. As Ireland is an open economy, and very responsive to global issues such as competitiveness and performativity, Junior Cycle education here is currently responding to and facing the above challenge.

Junior Cycle Policy Reform in Ireland

The policy for reform, *Framework for Junior Cycle 2015* (DES, 2015a), represents a reconceptualisation of learning, teaching, curriculum and assessment at Junior Cycle – the first three years of the secondary school system. The policy change represents a radical departure from the former Junior Certificate examination which was completely externally assessed. Up to this point, the “cultural script” (Elwood & Murphy) of assessment in Ireland has been dominated at Junior and Senior Cycle by national grading systems which did not involve the teacher adopting the role of assessor for certification purposes. Although this is not the first time Junior Cycle review and reform has taken place, the magnitude of the change in terms of curriculum and assessment combined is unprecedented.

In 1989, the Junior Certificate programme was introduced in place of the former Intermediate Certificate to “provide a single unified programme for students aged broadly between twelve and fifteen years” (DES, 1995, p.46). With the aim of achieving greater breadth, balance and relevance at Junior Cycle, the Junior Certificate subject-based curriculum sought to “extend and deepen the quality of students’ educational experience in terms of knowledge, understanding, skills and competencies and to prepare them for further study at senior cycle” (DES, 1995, p.46). It also sought to “contribute to the moral and spiritual development of students” (DES, 1995, p.46) and to encourage them “to develop qualities of responsible citizenship in a national, European and global context” (p.46). In spite of these developments in curriculum, however, assessment remained, as it had been before, as a state-certified terminal examination at the end of the third year of lower secondary education.

In an effort to prepare young people for a rapidly changing world, further development and refinement of the Junior Certificate curriculum throughout the 1990s led to the identification of revised Junior Cycle outcomes in the Government’s *White Paper on Education: Charting our Education Future* (DES, 1995). Notwithstanding this attempt to better align the curriculum with the educational requirements of young people, a subsequent review of the Junior Certificate programme suggested that curricular goals were not being achieved due to a mismatch between the redesigned curriculum and the terminal examination (National Council for Curriculum and Assessment [NCCA], 1999). The attempts at curricular flexibility and innovation were being strangled by an outdated examination system. It became very apparent, through subsequent review processes, that unless the examination changed, student experience at Junior Cycle would remain the same (NCCA, 2011). Key concerns and criticisms were raised in relation to:

the dominating effect of the Junior Certificate examination on teaching and learning practice and on school organisation and structures, the perception of an inflexible overcrowded curriculum, the disengagement of many students at an early stage in the cycle, the inadequate time for engagement with deeper learning, the narrow range of assessment activity and the limited access to a single qualification. (NCCA, 2010, p.13)

Furthermore, longitudinal research conducted by the Economic and Social Research Institute [ESRI] (Smyth, Dunne, Darmody & McCoy, 2006) pointed to a problem of identity amongst young people at Junior Cycle. In the case of many students, lower secondary education was considered a meaningless period, devoid of any real purpose in its own right. Some students noted that the Junior Certificate examination had become a “dress-rehearsal” for the Leaving Certificate Examination, the terminal examination at the end of secondary education which determines students’ transition to tertiary education. The once-off nature of the Junior Certificate Examination, as well as the narrow range of assessment methods throughout Junior Cycle, were also highlighted as contributing factors to the disengagement of students throughout lower secondary education in Ireland. Overall, the evidence suggested that the dominance of the terminal examination at Junior Cycle resulted in the narrowing of the learning experience for students, an experience ultimately characterised by teaching to the test and rote learning.

In June 2009, ten years after the initial review of the Junior Certificate examination, the then Minister for Education and Science, Batt O’Keeffe T.D., addressed the NCCA conference stating that:

Recent public debate on education (...) has become increasingly critical of a second-level system that’s considered by many to be driven by rote learning and exam pressures rather than the promotion of real understanding and skills. While I realise that this is a gross over-simplification, it’s not unrelated to the overall issue of curriculum overload, the need for greater emphasis on assessment for learning, practical project and portfolio assessment and the time necessary to promote experiential self-directed learning. Our learners need to

be flexible, adaptable, resilient and competent if they are to participate successfully in society and in lifelong learning.

The Minister also acknowledged the ESRI (Smyth et al., 2006) research stating that “the study leads to the inevitable conclusion that the current Junior Cycle curriculum and assessment are in need of reconfiguration to reflect the fundamental principles of education and to take account of current best practice internationally” (O’Keeffe, 2009).

It was from this point onwards that the drive towards Junior Cycle policy reform in Ireland gained real momentum. Progressing through a subsequent six year period of discussion documents (NCCA, 2010), consultation processes (NCCA, 2011), draft policy frameworks (NCCA, 2011) and significant industrial relations unrest, *Framework for Junior Cycle* (DES) was finally published in 2015 and recently endorsed by all teaching unions in Ireland (ASTI, 2017b). In its initial published form in 2012, the framework faced great opposition from teaching unions due mainly to the assessment procedures which were broken down in the form of a school-based assessment component for certification (worth 40% of the marks) and an external examination worth 60% of the overall marks. The resistance to this model, which, arguably, speaks in itself to how teachers might conceive of assessment, led to subsequent review and redesign of the assessment procedures. Union resistance to the original model could suggest that certain teachers believe that school-based assessment and external assessment are separate assessment types with different purposes and should therefore not be combined into one overall percentage or mark. On the other hand, the resistance could also point to an understanding of assessment solely in terms of an external process which does not involve the teacher.

The new Junior Cycle policy framework aims to set out curriculum and assessment procedures that will “provide students with quality learning opportunities that strike a balance between learning knowledge and developing a wide range of skills and thinking abilities” (DES, 2015a, p.29). These opportunities will allow for a much broader range of student learning to be assessed and reported upon. The policy document indicates that learning at Junior Cycle (DES, 2015a, p.7):

- will be guided by the twenty-four statements of learning, eight principles and eight key skills
- will encompass learning in subjects or a combination of subjects and short courses
- will include an area of learning entitled Wellbeing
- will provide a range of other learning experiences (e.g. students participating in a debating competition)
- may include priority learning units (PLUs) that will help to provide a junior cycle programme that is appropriate to the needs of particular students with significant special education needs.

The framework also points to a renewed focus on literacy and numeracy which aligns with Ireland’s national strategy to improve literacy and numeracy (DES, 2011), stemming partially from Ireland’s results in PISA 2009. A decline in the literacy and numeracy performance of Irish post-primary students in PISA (2009) was presented as part of the evidence for the implementation of a national strategy to improve literacy and numeracy standards among children and young people by 2020. The continuation of its prioritisation in this reform policy speaks to the power of the OECD agenda outlined earlier.

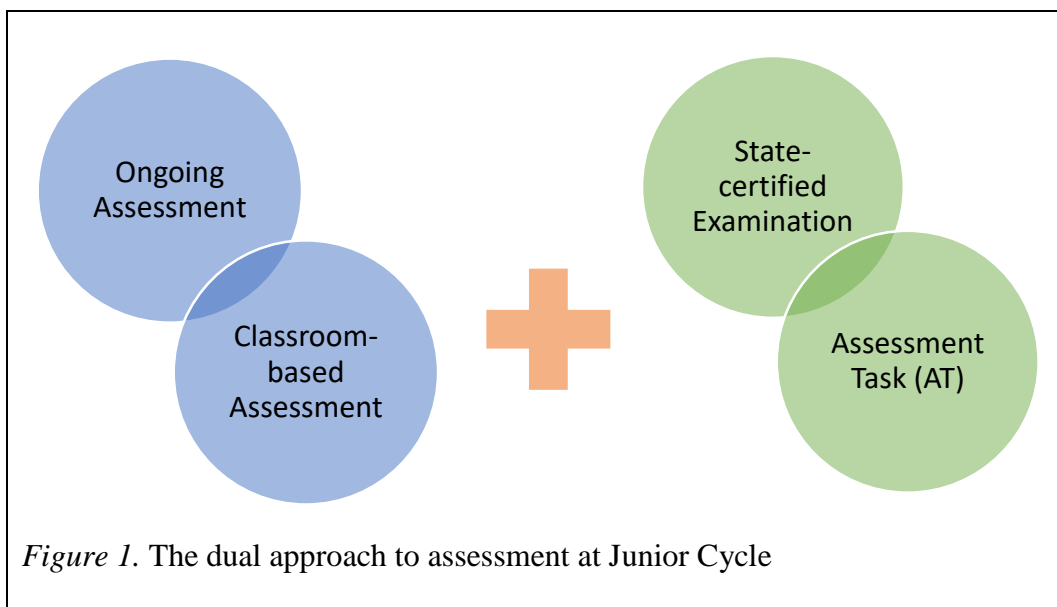
Above all these changes, however, as noted in the policy document, “the most significant change in the new Junior Cycle is in the area of assessment” (DES, 2015a). The reform procedures “present a dual approach to assessment that supports student

learning over three years of junior cycle and also measures achievement at the end of those three years” (DES, 2015a, p.7). Similar to attempts in New Zealand (Crooks, 2011) and Scotland (Hutchinson & Young, 2011) to merge AfL and accountability, it has been argued in the context of the Irish reform that “the dual approach to assessment provides a valuable opportunity to embed classroom based assessment and formative assessment for learning while recognising the role of external assessment” (O’Sullivan, 2015). As illustrated in Figure 1, the broader approach to assessment includes formative assessment, classroom-based assessment, an externally-marked terminal examination in addition to an externally-marked reflective assessment task. In the case of each subject, students undertake two structured classroom-based assessments, one in second year and one in third year. These tasks are set by the National Council for Curriculum and Assessment (NCCA), in conjunction with the State Examinations Commission (SEC), but are assessed by the classroom teacher. Each of the two classroom-based assessments will be drawn from a variety of types of assessment such as project tasks, oral language tasks, investigations, field studies and artistic performance. After the second of the Classroom-Based Assessments, students will complete a written Assessment Task on what they have learned and the skills and competences that they have developed in that assessment. This Assessment Task (AT), along with the final examination at the end of third year, will be marked by the SEC. In order to support the implementation of these new procedures, teachers must engage in what is termed “Subject Learning and Assessment Review” meetings where they have the opportunity to share and discuss samples of their assessments with departmental colleagues (DES, 2015a). Teachers evaluate students’ work against exemplars and “a nationally determined, common set of descriptors” (DES, 2015a, p.38). Students’ overall achievement throughout Junior Cycle will be reported on at the end of the third

year through the Junior Cycle Profile of Achievement (JCPA). Each part of the assessment procedures (state-certified examination (including AT), classroom-based assessments, wellbeing and other learning experiences) will be reported on in a separate section of the JCPA. While this broad approach to reporting aims to value different aspects of students' learning and the assessment approaches that generate evidence of that learning, it has been argued that the success of this approach will depend upon all educational stakeholders valuing the school-based achievement, reported through descriptors, in the same way that grades from state-certified examinations are currently valued (Murchan, 2017). This argument, points, once again, to the significance of understanding at this time how teachers conceive of assessment.

Furthermore, the new assessment procedures are underpinned by, and reliant upon, a strong commitment to formative assessment “as a normal part of learning and teaching in classrooms” (DES, 2015b, p.2) and not as “an additional bolt-on to the work teachers do” (DES, 2015b, p.2). For the change to be successful, teachers will need to believe in the underlying pedagogical principles of formative assessment. The policy also appears to address learning theory more broadly in its aim “to recognise and value the different types of learning that take place in schools”, thereby allowing “for a more rounded assessment of the educational achievements of each young person” (DES, 2015a, p.8). Without question, the new model of school-based assessment represents a significant departure from the former Junior Certificate examination programme in which the teacher was positioned apart from the assessment process. Now, however, teachers are positioned at the heart of the assessment process and will play a vital role in the translation of assessment policy into practice. The policy reform measures will require active engagement, critical reflection and deep understanding on the part of the teacher. Teachers will need to effectively use and interpret assessment information in a

new way, to reconcile formative and summative assessment procedures and to create a classroom environment where all parties engage with learning and assessment as an iterative reflective process. A challenging time lies ahead and it has been acknowledged by policy makers that “successful implementation will depend to a very significant degree on the professional skills and abilities of teachers” (DES, 2015a, p.35). While acknowledgement of skills and abilities is essential, it would appear that these aspects constitute but one part of the complex jigsaw of teacher assessment capacity, a jigsaw which appears to foreground beliefs in times of change.



Teachers’ Beliefs

Drawing on reconceptualised models of the role of teacher as assessor, which are detailed in the literature review (Looney, Cumming, van Der Kleij & Harris, 2017; Xu & Brown, 2016), a broader approach to assessment capacity is proposed which expands on the traditional concept of teacher assessment literacy (Stiggins, 1991). This broader understanding of teachers’ engagement with assessment highlights the importance of knowledge, beliefs, feelings, skills, learning and context. While each of

these dimensions is worthy of consideration in its own right, research has shown that beliefs play a significant role in times of educational change (Fullan, 2007). In their concluding remarks to an analysis of teachers' perceptions of assessment, Gardner and Galanouli (2016) contend those attempting to implement change programmes in relation to assessment "should be mindful of the teachers' perceptions of the changes they propose" (p.721). Furthermore, in a seminal review of teachers' beliefs, Fives and Buehl (2012) conclude that teachers' beliefs filter information and content entering the cognitive domain, frame particular educational situations or problems and guide teachers' intentions and actions. A teacher may, therefore, accept, alter or potentially reject policy change depending on the degree to which the new knowledge is consistent with their current beliefs. The filtering role of beliefs is of particular importance in the current Irish situation given the changing policy context and the role of teachers as street level bureaucrats (Lipsky, 2010) to filter and implement the policy message. In addition to the significance of the filtering role of beliefs in the Irish reform context, a number of other key research findings in relation to beliefs further justify the need to explore how post-primary teachers currently conceive of assessment.

Firstly, much research has shown that a reciprocal relationship appears to exist between teachers' beliefs, context and experience (Fives & Buehl, 2012). The notion of context can include the school as well as the broader social and political realm (Levin, 2015). Specifically, in relation to the school context, school culture appears to play a role in shaping teachers' beliefs (Tschannen-Moran, Salloum & Goddard, 2015). Within school cultures, teachers' sense of collective efficacy to execute a course of action has been shown to effect teachers' sense of self-efficacy (Goddard & Goddard, 2001), which in turn has been linked to teachers' beliefs (Looney et al., 2017). These relationships are important in the context of Junior Cycle reform. Given the reluctance

on the part of teacher unions to engage with the new assessment initiatives, there is a chance the sense of collective efficacy towards the new procedures might be low in many schools. If this were the case, then teachers may be feeling low in confidence which in turn could negatively influence their conceptions of assessment and their consequential adoption of policy mandates.

Secondly, in relation to the broader political and social context, international studies specifically investigating teachers' conceptions of assessment suggest that teachers' conceptions of assessment are often ecologically rational in that they tend to reflect the policy priorities and educational traditions experienced by teachers within a particular society. Some of these studies have also shown that teachers may simultaneously hold multiple and sometimes conflicting conceptions of assessment (Brown, Lake & Matters, 2011; Harris & Brown, 2009), which speak, perhaps, to the competing educational agendas outlined earlier in this chapter. In light of this evidence, it was deemed of interest to consider how Irish post-primary teachers' conceptions of assessment might reflect the current policy mandate as well as the high-stakes assessment tradition in which these teachers have been working for so long.

Thirdly, aligned with the importance of context and experience, is Lortie's (1975) argument on the power of teachers' apprenticeship of observation. By the time teachers begin their careers, they already hold an established set of educational beliefs. Research has shown that these beliefs exist as part of an interconnected system in which beliefs can be held with varying degrees of conviction (Green, 1971; Rokeach, 1968). Given that post-primary teachers in Ireland have experienced a high-stakes examination system throughout their own schooling years, it was considered important, in the context of this study, to examine the degree of conviction with which teachers

hold their assessment-related beliefs and to explore the possible consequences that this might have for curriculum and assessment reform.

Finally, research has shown that if teachers' beliefs are to develop or change, then teachers need to firstly be given the opportunity to become aware of their beliefs and to make them explicit (Cabaroğlu & Roberts, 2000). As is detailed in the literature review, the breadth of international, and to a lesser extent national, research engaging with teachers' beliefs spans many disciplines, levels of education, theoretical paradigms and methodologies. Many of the small-scale studies have adopted qualitative approaches to investigation whereas those studies attempting to engage with beliefs on a larger scale have predominantly used quantitative methods of inquiry. This study aimed to contribute to the quantitative body of literature by providing post-primary teachers in Ireland with the opportunity to make their conceptions of assessment explicit. A clear understanding of how teachers conceptualise assessment is an important springboard for the introduction of policy change and the planning of professional development.

Research Question

In the absence of evidence-based data on how Irish post-primary teachers conceive of the nature and purpose of assessment, this study focused on teachers' conceptions of assessment as a particularly important part of teachers' engagement with assessment reform. The research attempted to answer the following question: What are Irish post-primary teachers' conceptions of assessment?

Terminology

In relation to conceptions, the study adopted Thompson's (1992) definition of a conception as a "general mental structure encompassing beliefs, meanings, concepts, propositions, rules, mental images, preferences and the like" (p.130).

This study adopted the widely documented understanding of assessment purposes as operating on a continuum or spectrum (Barnes, Fives & Dacey, 2015; Brown & Gao, 2015; Gardner & Galanouli, 2016; Remesal, 2011) ranging broadly from a formative or diagnostic perspective at one end to a summative perspective on the opposite end. Formative assessment, used interchangeably in this study with the term Assessment for Learning (AfL), refers to "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (Assessment Reform Group, 2002). Formative assessment, therefore, supports learning as an ongoing part of the teaching and learning process. Learners and teachers mutually engage in the learning process through such means as sharing learning intentions and success criteria, providing and receiving feedback, questioning, partaking in class discussion and

engaging in peer and self-assessment (Black & Wiliam, 2009). Summative assessment, on the other hand, takes place after a particular period of learning as a means of gauging “the extent of the knowledge, understanding or skill that the student has acquired” (Gardner & Galanouli, 2016, p.711). Summative assessment, conducted by the teacher or by an external body, generally serves the purpose of summarising learning, reporting achievement or monitoring progress. In its extreme form, summative assessment is also sometimes used for accountability purposes. While formative and summative assessment have distinct purposes, they do not necessarily have to be mutually exclusive. Summative assessment may be used in a formative way to support learning if reviewed through a feedback discussion, for example. As argued by Gardner and Galanouli (2016), the interpretation of assessment as operating on a continuum helps to reflect the complexity of the assessment process and the potential for assessment types to serve more than one purpose.

Summary

As noted at the beginning of the chapter, this introduction was intended to contextualise the study for the reader and to provide a rationale for its focus. The key issues raised throughout this chapter can be summarised as follows: (i) the 21st century assessment narrative is being shaped and directed by the competing priorities of economic competitiveness, key skills, performativity and contemporary learning theory; (ii) policy makers worldwide are attempting to harmonise these conflicting educational agendas through the redesign of curriculum and assessment policy; (iii) Junior Cycle education in the Republic of Ireland is currently in the midst of such

reform; (iv) a key element of the reform relates to the positioning of the teacher at the heart of the assessment process.

The significance of beliefs in relation to the role of teacher as assessor is explored in depth in the literature review. The discussion is guided by the following key points: teachers' competence in assessment, which was traditionally conceptualised as teacher assessment literacy, has recently been reconceptualised in a much broader sense; this broader conceptualisation encapsulates such aspects as knowledge, conceptions, feelings, skills and learning; while each dimension of teachers' assessment work is worthy of consideration, research has shown that conceptions play a vital role in times of change; foregrounded in the idea of conceptions are beliefs which filter, frame and guide; these beliefs frame and guide practice but this relationship is not necessarily linear in nature and is open to contextual influence; international research focusing on teachers' conceptions of assessment has reported that teachers' conceptions may be multi-faceted and sometimes conflicting and that they are often ecologically rational.

Organisation of the Thesis

This introductory chapter is the first of five chapters.

Chapter 2 focuses on literature and previous research in relation to teachers' beliefs generally as well as teachers' conceptions of assessment specifically.

Chapter 3 details the methodology employed in the study. Specifically, this chapter outlines the conceptual framework and research design as well as discussing the participant profile and the methodological instrument employed in this study. Furthermore, this chapter addresses the ethical considerations raised by the study.

Chapter 4 outlines the analytical approach adopted in the study and presents an account of the findings from the data collected.

Finally, Chapter 5 provides a summary and critical review of the findings in conjunction with previous research and literature in this field. The implications of this study in relation to policy implementation and curriculum reform in the Irish post-primary context are addressed. This chapter also acknowledges the limitations of the study and the potential for future research.

Chapter 2: Literature Review

Introduction

In a recent historical overview of teacher belief research, Ashton (2015) tracked the evolution of the discipline over 60 years. She aligned the rise of teacher belief research with the emergence of what Shulman (1986) referred to in the third *Handbook of Research on Teaching* as a new paradigm of teacher cognition and decision making. Representative of this new perspective was a chapter in that handbook by Clark and Peterson (1986) on teachers' thought processes. Shortly after this, growing enthusiasm for the potential of teachers' beliefs to inform educational practice saw the publication of some key papers in the field (Kagan, 1992; Nespor, 1987; Pajares, 1992; Pintrich, 1990; Thompson, 1992). Over twenty years on from the publication of these papers, Fives and Buehl (2012) completed a seminal review of what had become a complex and far-reaching body of literature. Examining 627 peer-reviewed articles spanning multiple disciplinary perspectives, the authors attempted to consolidate the belief topics which had been addressed, demonstrate how teachers' beliefs had been defined, and synthesise trends in empirical findings across different research paradigms.

Organisation of the Review

Drawing closely on the categorisation of ideas in the work of Fives and Buehl (2012), the first section in this chapter explores some of the key findings from teacher belief literature up to and beyond the seminal review by Fives and Buehl (2012). Four key categories which Fives and Buehl (2012) used to organise their findings and which are used to guide the discussion in the first part of this chapter are the nature of

teachers' beliefs; the role of teachers' beliefs; the relationship between beliefs and practice; and belief development and change. While the main aim of this section is to introduce the reader in a general sense to the broad field of teacher belief research, efforts are made where possible to link the discussion to assessment-related studies. With this broad context in place, the second section of the review focuses solely on literature related to teachers' conceptions of assessment. Finally, two models of teachers' engagement with assessment are explored by way of positioning the contents of the review within a broader framework of teacher assessment capacity more generally.

Teachers' Beliefs

Nature of Teachers' Beliefs

The discussion in this section considers how teachers' beliefs are characterised. The construct of teachers' beliefs will be examined by focusing on the manner in which these beliefs are defined, composed, organised and related to context.

Defining Teachers' Beliefs.

Fives and Buehl (2012) note a considerable amount of diversity in the literature in how teachers' beliefs are conceptualised. Teachers' beliefs are a "messy construct" (Pajares, 1992, p.307) which can be defined and understood in a multitude of ways. This "lack of cohesion" (Fives & Buehl, 2012, p.2) has led to notable variances in the terminology used throughout this field of research. Pajares (1992) conceptualised beliefs as an "individual's judgement of the truth or falsity of a proposition" (p.316) but he also contended that many other terms such as attitudes, values, axioms, opinions, ideology, perceptions, preconceptions, dispositions, and implicit theories are actually beliefs in disguise. Kagan (1992) referred to teachers' beliefs as "a particularly

provocative form of personal knowledge” (p.65) and Richardson’s (1996) understanding, which derived from Green (1971), is that a belief “describes a proposition that is accepted as true by the individual holding the belief” (p.104). In a key paper on how teachers conceive of mathematics, Thompson (1992) suggested it would be more helpful for researchers to think more broadly in terms of conceptions rather than beliefs which she defined as “a more general mental structure encompassing beliefs, meanings, concepts, propositions, rules, mental images and the like” (p.141). These few examples from some key contributors to the field illustrate the complex and elusive nature of the construct in question. However, despite the nuanced nature of the literature, a clear commitment to the importance and centrality of teacher cognition in educational research is present throughout.

Contents and organisation of teachers’ beliefs.

Fives and Buehl (2012) provide a categorisation framework for the general contents of teachers’ beliefs. They report that the content of teachers’ beliefs may be broadly framed to include beliefs about self, context or environment, content or knowledge, specific teaching practices, teaching approach, and students. It was noted in the review that the various beliefs that a teacher may have are “overwhelmingly recognized” (Fives & Buehl, 2012, p.477) in the literature as existing within a belief system. Teachers’ educational beliefs could be described in terms of an interconnected subsystem or substructure within one’s broader overarching belief system. The idea of a belief system which Fives and Buehl (2012) refer to, stems from the work of Rokeach (1968) and Green (1971) who identified some key aspects of these systems. Firstly, belief systems have a quasi-logical structure in which some beliefs are primary and some are derivative of those primary beliefs. Secondly, beliefs may be psychologically central or peripheral based on the degree of conviction with which they are endorsed.

The more central a belief is in the system, the more likely it will be to resist change and to have implications and consequences for other beliefs in the system. In addition to these dimensions, Green (1971) also argued that beliefs may exist in isolated clusters thereby allowing for the co-existence of conflicting beliefs. Rokeach (1968) and Green's (1971) theorisation of a belief system highlights the multidimensional and complex nature of this construct and the consequential importance for researchers to be mindful of this complexity when looking at any aspect of the belief system. As noted in the introductory chapter, the degree of conviction with which teachers hold their assessment-related beliefs is of interest in the Irish context. Also of note are the possible interconnections which might exist between potentially conflicting beliefs and the possible influence these might have in shaping teachers' practice.

Importance of context in shaping teachers' beliefs.

Another key aspect of teachers' beliefs is their contextual and situated nature. A growing body of research acknowledges, in varying ways, that teachers' beliefs cannot be isolated from the environments in which they occur (Fives & Buehl, 2012; Levin, 2015; Mansour, 2009; Pajares, 1992; Skott, 2009, 2015). The notion of context, as pointed out by Levin (2015), includes "the larger social, political, and economic climate as well as the immediate school context" (p.51). A systematic body of cross-cultural research suggests that teachers' conceptions of assessment are reflective of the policy priorities within the social and cultural contexts in which they teach (Brown & Harris, 2009). This is a noteworthy finding in the current Irish context and is explored in detail in the later section on teachers' conceptions of assessment. In line with Bandura's (1997) theory of triadic reciprocal determinism which posits that a person's behaviour both influences and is influenced by personal factors (including beliefs) and the social environment, Fives and Buehl (2012) contend from their review that "beliefs

are individually held conceptions that are in constant relation to the context and teachers' experiences" (p.476). With reference to the school context they suggest that "a teacher working in a school holds personal beliefs about the school community that in turn influence and are influenced by the school environment and his or her own behaviors" (p.476). A key aspect of the school environment which appears to shape teachers' beliefs is school culture. Tschannen-Moran, Salloum and Goddard (2015) define school culture as:

a set of tacit assumptions and beliefs that have arisen as a group of educators has wrestled with the problems of practice over time, and that has worked well enough to be considered valid and that is consequently passed along to new organizational members as the proper way to think, perceive and behave. (p.302)

This understanding would appear to resonate with the research by Rokeach (1968) and Green (1971) on belief systems as the power of a shared norm could manifest in a central belief within a teacher's belief system which in turn could influence one's other beliefs or practice. Research has shown that the culture within a school can exert a particular influence upon teachers' self-efficacy beliefs (Fives & Alexander, 2004; Goddard & Goddard, 2001). A specific aspect of school culture which has been linked to teachers' self-efficacy beliefs is a construct known as collective efficacy. As defined by Goddard, Hoy and Woolfolk Hoy (2004), collective efficacy beliefs are "the perceptions of teachers in a school that the faculty as a whole can organize and execute the courses of action required to have a positive effect on students" (p.7). Goddard and Goddard (2001) show that when compared with other contextual variables such as school size, perceived collective efficacy is the strongest predictor of teachers' individual self-efficacy. In addition to shaping teachers' beliefs, context and culture

may also hinder or support the enactment of teachers' beliefs, an important point discussed later in the chapter.

Role of Teachers' Beliefs

One of the key conclusions from Fives and Buehl's (2012) review was that teachers' beliefs may function in different ways depending on the situation in question. They argue that teachers' beliefs serve to filter information and content entering the cognitive domain, to frame particular educational situations or problems and to guide teachers' intentions and actions. As outlined in the introduction, in light of the changing policy context in Ireland, the power of beliefs to filter, frame and guide is of particular significance at the current time.

Filtering

Firstly, as filters, beliefs can influence what new information teachers allow enter the cognitive domain and how that new information is interpreted. The idea of beliefs as filters is particularly significant in the context of educational reform as they act as the initial lens through which teachers may accept, modify or reject change (Fullan, 2007). The filtering role of beliefs is exemplified in a recent study by Tam (2015a) which investigated Chinese teachers' beliefs about learning in Professional Learning Communities (PLCs) and the way in which these beliefs governed the teachers' inclination to engage in collegial practice. A qualitative case study, involving semi-structured interviews and observation, was carried out with the teachers in the Chinese and English Departments of a Hong Kong Secondary school. The findings demonstrated that the Chinese teachers believed that "professional knowledge can be co-constructed through critical collaboration with colleagues" (Tam, 2015a, p.437) and

were therefore receptive to the idea of taking part in a PLC. These teachers took part in many collegial activities which facilitated their professional growth. Conversely, their English Department colleagues “held deeply rooted beliefs that privatisation is less risky than opening the doors of the classroom to colleagues” (Tam, 2015a, p.438). These teachers did not believe that collaboration with colleagues would foster learning and the attempt at implementing a PLC with these teachers was not successful. The success and failure of the PLC implementation in these cases was consistent with the teachers’ underlying beliefs about learning in a PLC.

Framing

The second function which Fives and Buehl (2012) identify for teachers’ beliefs is that of a framing device. Teachers may draw on certain beliefs to define or shape a particular task or problem. In the case of reform initiatives such as assessment for learning for example, the level of congruence between teachers’ beliefs and the underlying pedagogical foundations of formative assessment would appear to shape their degree of engagement with it (Marshall & Drummond, 2006; Wallace & Priestley, 2011; Warwick, Shaw & Johnson, 2015). In a study of teachers’ engagement with assessment for learning, Marshall and Drummond (2006) found that some teachers embraced the “spirit” (p.137) of AfL, whereas others merely conformed to the “letter” (p.137). The authors used the term “spirit” (p.137) to refer to full engagement with the promotion of pupil autonomy and “letter” (p.137) to describe surface-level engagement with AfL procedures. The authors noted that the teachers who fully embraced AfL in the study also saw the classroom as a site of their own learning. This led the authors to conclude that the teachers’ beliefs about pupil autonomy as well as their beliefs about learning framed their application of AfL in their classrooms. These findings resonate with the challenge of large-scale implementation of AfL in Norway which experienced

difficulties in relation to “teachers who did not really understand the principles behind AfL” (Hopfenbeck, Flórez Petour & Tolo, 2015, p.55). The teachers’ narrow beliefs as to what constituted AfL appeared to frame their application of it in the classroom. The authors noted that “many teachers believed that the view of The Directorate of Education and Training (DET) was the only correct way of doing AfL” (p.55) and that this belief prevented the teachers from engaging with the broader potential of AfL. The manner in which Irish post-primary teachers frame their application of AfL is likely to play a significant role in the success of the policy reform implementation. This is another aspect of the reform which could pose challenges given that recent national reviews suggest that formative assessment has not, as yet, been fully embedded at in the post-primary sector in Ireland (DES, 2013, 2014).

Guiding

The third role reported for teachers’ beliefs is that of guiding teachers’ intentions and actions (Fives & Buehl, 2012). Beliefs are often identified in the literature as “precursors to behaviour” (Buehl & Beck, 2015, p.68) and “predictors of practice” (p.68). A body of literature exists exploring the influence of beliefs upon teacher practice in many domains. Some recent contributions to the field include studies related to teaching (Devine, Fahie & McGillicuddy, 2013), writing instruction (Gaitas & Martins, 2015), outdoor scientific learning (Glackin, 2016), student inquiry-based learning (Song & Looi, 2012), educational technology (Ertmer, Ottenbreit-Leftwich & Tondeur, 2015) and special needs and inclusion (Kiely, Brownell, Lauterbach & Benedict, 2015). Given that this thesis focuses on assessment, some recent studies in that domain are now discussed. In a New Zealand study of 518 practicing teachers, Brown, Harris and Harnett (2012) used the *Teacher Conceptions of Feedback Inventory* (TCoF) to evaluate the teachers’ conceptions of feedback. This

self-report survey instrument was devised by the authors drawing primarily on the work of Hattie and Timperley (2007) and Irving, Harris and Peterson (2011). The authors then analysed the relationship between these conceptions and feedback practice definitions using structural equation modelling. The results indicated that “there were conceptually meaningful relations between teachers’ conceptions of feedback and the practices they associate with it” (Brown, Harris & Harnett, 2012, p.974). Panadero, Brown and Courtney (2014) conducted a study with 944 Spanish teachers (at primary, secondary and university levels) to explore their beliefs about student self-assessment (SSA) and how these beliefs influence their self-reported uses of SSA in the classroom. Upon evaluation of the self-report instrument, using confirmatory factor analysis and structure equation modelling, five statistically significant predictors for the use of SSA were found. They were positive experience with SSA, belief in student participation in assessment, willingness to include self-assessment as a percentage of final grade, endorsement of self-assessment advantages and previous attendance at assessment courses. The authors concluded that “self-reported use of SSA in classroom settings across educational settings in Spain is strongly influenced by teachers’ values, attitudes and prior experiences with this type of assessment” (Panadero et al., 2014, p.377). Two further studies adopted Ajzen’s (2002) theory of planned behaviour as a framework to predict teachers’ intentions to implement school-based assessment (Yan, 2014) and formative assessment (Yan & Cheng, 2015). According to the theory of planned behaviour (Ajzen, 2002), three key determinants – attitude toward the behaviour, subjective norm, and perceived behavioural control – combine to form a behavioural intention. It is contended that the level of this intention together with the degree of perceived behaviour control can predict behaviour. Attitude toward the behaviour refers to one’s general outlook on performing the behaviour in question. Subjective

norm refers to “perceived social pressure” (Ajzen, 2002, p.665) to perform the behaviour. Perceived behaviour control is concerned with “perceived ability to perform a behaviour” (Ajzen, 2002, p.668) and is often measured as one’s sense of self-efficacy. The results from both of these studies indicate that teachers’ self-efficacy beliefs are strong predictors of their intention to implement formative assessment (Yan & Cheng, 2015) and school-based assessment (Yan, 2014). This finding connects with previously discussed findings in relation to school culture (collective efficacy) and self-efficacy (Goddard & Goddard, 2001) and points to potentially interesting links between collective efficacy, teacher’ self-efficacy and teachers’ assessment beliefs and practices. While these studies illustrate how beliefs can act as an explanatory principle for practice, other empirical findings report varying degrees of consistency between beliefs and practice (Fives & Buehl, 2012). Attempts to explain the lack of congruence which can sometimes exist between beliefs and practice draw attention to the complexity of the belief-practice relationship and the factors which can influence it. These factors are now discussed.

Internal and External Factors Influencing Belief Enactment

Thompson (1992) argued that the belief-practice relationship is not a “simple cause-and-effect relationship” (p.140). This is best understood when the relationship is positioned “within a broader multi-leveled context of various internal and external factors” (Buehl & Beck, 2015, p.74). Drawing on work by Bronfenbrenner (1989) and Fives and Buehl (2012), Buehl and Beck (2015) provide a useful ecological framework to demonstrate how the enactment of teachers’ beliefs can be supported or hindered by internal factors related to the teacher and by external factors at classroom, school and

state level. This overarching framework supports much of the research discussed thus far (Bandura, 1997; Green 1971; Rokeach, 1968).

Internal factors include elements such as teacher knowledge, experience, self-reflection and other beliefs within the teacher belief system. Focusing on teacher knowledge for example, Rushton, Lotter and Singer (2011) illustrate the mediating effect of pedagogical knowledge in a study of chemistry teachers who embarked upon a professional development course in inquiry-based teaching. The teachers were interviewed prior to the workshop in relation to their learning beliefs and instructional practices. A lack of alignment was apparent from the interview data between the teachers' espoused constructivist beliefs and their teacher-centred instructional practices. The authors noted that "the teachers' misunderstanding of inquiry-based teaching and their inability to see how inquiry could be used to cover their content standards initially inhibited them from using the student-centered practices that they believed would result in student learning" (Rushton et al., 2011, p.42). Interactions among different aspects of a teacher's own belief system may also account for possible inconsistencies between belief and practice. As noted in the previous section on the role of beliefs, Marshall and Drummond (2006) cited beliefs about pupil autonomy and learning as influential in the implementation of assessment for learning. Buehl and Beck (2015) also refer to a range of studies in which factors such as self-efficacy beliefs, teachers' sense of responsibility for student learning and teachers' beliefs about students are mentioned as potential mediators in the belief-practice relationship.

Turning now to external factors, there are many elements at various levels which may facilitate or impede belief enactment. Examples include national educational policy, testing and accountability, school culture, parental expectations, instructional resources, class size, student ability, classroom management and time

(Buehl & Beck, 2015; Fives & Buehl, 2012). Phipps and Borg (2009) noted how classroom management concerns and student expectations presented a challenge to a group of English language teachers in the enactment of their grammar teaching beliefs. Despite stating that they did not believe in the value of controlled grammar practices, the teachers used mechanical gap-fill exercises in class as they felt that the students expected these and that they were a good control mechanism to keep the students calm. In a study across 12 primary and secondary schools in Ireland, Devine et al., (2013) found a contradiction between teachers' beliefs and practices in relation to good teaching. The authors concluded that the teachers' practices were mediated by "the sociocultural context of the school (gender, social class and migrant children), teacher expectations for different types of students and leadership practices within the school" (Devine et al., 2014, p.83). Crawford (2007) demonstrated how the established culture within a school can sometimes inhibit teachers with fresh ideas from acting on their beliefs. The study in question followed 5 prospective high school science teachers as they engaged in a year-long work placement. Each prospective teacher had a mentor who was already teaching in the school. The study found that the "mentor teachers' beliefs and preferred pedagogical approaches appeared to deter at least some of the prospective teachers from deviating from the mentor's established classroom culture" (Crawford, 2007, p.623). These teachers were afraid to try out inquiry-based approaches as they knew that such approaches did not align with their mentors' pedagogical beliefs. The impact of broader policy issues on belief enactment was highlighted in a series of studies in England (James & Pedder, 2006; Winterbottom et al., 2008) and elsewhere (Warwick et al., 2015) which examined the gap between teachers' assessment values and practices. James and Pedder (2006) developed a 30-item likert-format questionnaire which measures teachers' frequency of use of certain

assessment practices as well as the extent to which they claim to value each of these practices. Analysis of 558 survey responses grouped the practices into three main dimensions – making learning explicit, promoting learner autonomy and performance orientation. Results indicated that while teachers most valued practices associated with making learning explicit and promoting learner autonomy, their implementation of practices associated with performance orientation was far greater than the other two dimensions. James and Pedder (2006) attributed these results to the curriculum and testing regime which prevailed in England at the time of this study. This finding begs a question in relation to the extent of the influence of the high-stakes examination culture on the belief-practice relationship amongst post-primary teachers in Ireland. The overall influence of internal and external mediating factors in the belief-practice relationship points to the importance of exerting caution when making any inferences from the findings of belief studies.

Belief Development and Change

Belief development and change is a complex process which appears to be underpinned and influenced by aspects related to the very nature of beliefs themselves. As argued by Rokeach (1968) and Green (1971), if beliefs are held as part of a system in which certain beliefs may be more deeply ingrained than others, then it is likely that peripheral beliefs held with less conviction will be more susceptible to change than core beliefs. Secondly, given that research points to a reciprocal relationship between teachers' beliefs, environment and experiences (Bandura, 1997; Fives & Buehl, 2012), it is likely that this contextual relationship will exert an influence on belief development. Fives and Buehl (2012) provide a useful categorisation framework for

discussing belief change. They distinguish between changes in trainee teachers' beliefs through initial teacher education, developmental changes over time through teaching experience and changes in practicing teachers' beliefs resulting from professional development in-service support.

Belief change through initial teacher education

Teachers beginning teacher education programmes bring with them an established set of beliefs which have been developed during their own school years, or what Lortie (1975) referred to as their "apprenticeship of observation". The power of this enculturation process can sometimes leave one with beliefs which are resistant to change (Kagan, 1992; Pajares, 1992). Helping teachers to become more aware of their beliefs would appear to be a key first step in belief change (Fives & Buehl, 2012; Kagan, 1992). Kagan (1992) contends that:

if a program is to promote growth among novices, it must require them to make their personal beliefs explicit; it must challenge the adequacy of those beliefs; and it must give novices extended opportunities to examine, elaborate and integrate new information into their existing belief system. (p.77)

Cabaroglu and Roberts (2000) reported on a study inquiring into development in student secondary teachers' beliefs on language teaching and learning. Twenty participants participated in a sequence of three in-depth interviews over a 36 week period. Analysis of the interview data indicated that all but one of the participants demonstrated changes in their beliefs. The authors noted variations "in the content areas which changed, in the degree of change, and in the time at which change seemed to occur" (Cabaroglu & Roberts, 2000, p.398). This finding is arguably linked to the notion of central and peripheral beliefs with certain beliefs being more deeply engrained than others and therefore difficult to change (Green, 1971; Rokeach, 1968).

In the overall analysis of the data, a set of change process categories were observed. These were awareness/realization, consolidation/confirmation, elaboration/polishing, addition, re-ordering, re-labelling, linking up, disagreement, reversal, pseudo change and no change.

Drawing on Cabaroglu and Roberts' (2000) framework, Yuan and Lee (2014) investigated the process of belief change among three trainee language teachers during 10 weeks of the teaching practicum in a Chinese university. Adopting a socio-cultural perspective, the authors also examined the socio-cultural factors affecting belief change in their participants. Upon analysis of multiple forms of qualitative data, a range of belief change processes similar to Cabaroglu and Roberts (2000) emerged with the addition of two new change processes, namely integration and modification. Socio-cultural factors which influenced belief change included social learning activities, mentor support and guidance and contextual obstacles and challenges which activated a sense of agency in the trainee teachers.

Belief change through teaching experience

In a study that moved beyond the training stage to the initial years of one's teaching career, Brownlee (2003) interviewed 11 primary school teachers at the beginning and end of their teacher education programme and again at the end of their third year of teaching. The interviews focused on the teachers' beliefs about knowing and the role of experts in knowledge construction. Shifts in teachers' beliefs were noted at each stage of the study. In terms of their beliefs about knowing, teachers predominantly reported mixed beliefs at the first interview, mainly constructivist beliefs at the second interview and most reported constructivist beliefs after three years of teaching. It should be noted, however, that the number of teachers holding

constructivist epistemological beliefs at the end point of the study was less than that at the second interview. Brownlee (2003) points out that one participant whose beliefs became less constructivist “described a context of needing to comply with school requirements that were quite transmissive” (p.93). This finding exemplifies the idea of school culture or context inhibiting belief enactment. In relation to the role of experts, most of the teachers initially believed that “experts facilitate the reception of absolute truths” (Brownlee, 2003, p.92). This belief may have been influenced by the participants’ experience during their apprenticeship of observation (Lortie, 1975). At the end of the teacher education programme, however, this belief changed dramatically to a constructivist view which held that experts “facilitate the construction of reasoned truths” (Brownlee, 2003, p.92). After three years of teaching, the constructivist viewpoint was still held by the vast majority of the participants. Despite the small scale nature of the study, the overall results suggest that the participants’ beliefs were more susceptible to change during their preparation programme than in their initial teaching years. However, the drop in the number of teachers holding constructivist beliefs about knowing between interviews two and three also suggests that the beliefs of some novice teachers were not fully embedded at that point and may have been sensitive, as mentioned above, to context or other issues.

Levin, He and Allen (2010) examined content changes in the pedagogical beliefs of 18 practicing primary teachers, with 1-6 years of experience, who had previously been part of a larger (n=84) belief study (Levin & He, 2008) prior to starting their teaching careers. The authors elicited the teachers’ pedagogical beliefs in the form of their personal practical theories (PPTs) using the personal theorising processes described by Cornett (1990). As part of this process, teachers were given the opportunity to “articulate their tacit beliefs and make them explicit, either orally or in

writing” (Levin, 2015, p.54). The process also “allowed teachers to choose what they wanted to reveal about their beliefs in their own words” (Levin, 2015, p.54). Analysis of the data revealed six content categories for the teachers’ stated beliefs. These were classroom community (21.88%), differentiation (18.75%), student-centred instruction (17.71%), organisation and classroom management (15.63%), professionalism (15.63%) and expectation (10.42%). Some notable differences appeared in the frequency of certain beliefs cited by the teachers in this study in comparison to the same beliefs expressed by the larger cohort prior to beginning their teaching career. In the pre-service study (Levin & He, 2008), 6.14% of the beliefs expressed were about differentiation, 8.05% of the beliefs referred to student-centred instruction, 6.99% of the beliefs came under the umbrella of professionalism and 5.51% were about expectation. The authors noted that the significant increase in percentages for these beliefs in the Levin et al. (2010) study indicated a shift towards more student-centred beliefs as the teachers gained experience. Another noteworthy finding in terms of belief development related to the absence of organisational or classroom management beliefs from teachers with 5 or 6 years teaching experience. One participant referred to such procedures as being “ingrained” (Levin et al., 2010, p.17) in her teaching and were therefore no longer a salient feature of her explicit pedagogical beliefs. Perhaps with experience over time, certain beliefs are held tacitly rather than consciously.

Belief change through professional development support

While teaching experience over time contributes to belief development of practising teachers, so too does engagement in professional development opportunities. Bandura (1997) argued that changing or deepening one’s belief systems can happen by experiencing success, observing success, emotional arousal, or through verbal persuasion. A positive professional development experience has the potential to

support such developments. Fives and Buehl (2012) note in their review that “two salient features of programs that report adaptive belief change in teachers are (a) a task or strategy focus and (b) the development of a community of practice among participants” (p.486). Beswick (2007) reported on an in-service programme conducted to influence teachers’ beliefs about teaching mathematics to students with mathematics learning difficulties (MLD). A mixture of 22 primary and secondary teachers participated voluntarily in the study which involved attending three workshops. During this time the participants discussed and analysed material presented to them, shared ideas and experiences, trialled new methods and undertook readings. Quantitative survey data obtained from the participants prior to and following the series of workshops revealed changes in their beliefs. In terms of beliefs about mathematics generally, there was a marked reduction in the number of teachers who believed that mathematics requires a good memory, that it requires logic and not intuition and that mathematicians solve problems quickly in their head. Some statistically significant differences were noted in relation to certain beliefs about students with and without MLD at the beginning and the end of the programme. For instance, participants were significantly less inclined at the beginning of the programme than at the end to see conceptual understanding as an appropriate goal for students with MLD compared to students generally. The authors remarked that this change in belief was consistent with the participants’ conversation which took place in the first workshop based on a collaborative task as well as a discussion of brief readings related to inclusion. Another finding revealed that participants were more inclined after the programme “to reject the notion that students with MLD should use concrete materials as a substitute for thinking to get answers” (Beswick, 2007, p.15). Notwithstanding the small scale nature of this study and the consequential caution that is necessary in drawing conclusions from it,

the results would suggest that teachers' mathematical beliefs are susceptible to change with the support of a professional learning community.

Tam (2015b) also examined the role of a professional learning community in changing teachers' beliefs. Working with 12 teachers in the Chinese Department of a Hong Kong Secondary school, Tam (2015b) conducted a longitudinal qualitative study which aimed to identify the features of a professional learning community that facilitate change in teachers' beliefs and practices and to identify what those changes were. Of particular relevance in the context of this study are the findings in relation to belief change. From 2007 to 2011, multiple sources of data were collected and analysed. Throughout this period, teachers engaged in many practices such as discussions, observations, reflection and action research. Findings from the data analysed indicated that the features of a PLC facilitating teacher change were development of a coherent structure, a collaborative culture, and effective learning activities. Belief change was reported in relation to curriculum, teaching, learning, role of teacher and learning to teach. Examples of changes included shifts in beliefs about teaching from direct knowledge transmission to a hybrid approach and shifts in beliefs about the role of teacher from director, disciplinarian and manager to curriculum planner, implementer, action researcher and reflective practitioner. Similar to Beswick (2007), although this was a small scale case study, the findings point once again to the possibility of belief change with the support of a professional learning community. While these are but two studies out of a broad range of research in the area of teacher/professional learning communities (Liebermann & Miller, 2008), they highlight the potential of this practice to guide teachers through the processes of belief development and change which Caboroglu and Roberts (2000) and Yuan and Lee (2014) described.

Summary of Section One

Belief research is an extensive and complex domain marked by a significant amount of conceptual diversity. It is widely accepted in the literature that teachers' beliefs exist as part of an interconnected system (Fives & Buehl, 2012) which allows for the co-existence of potentially conflicting beliefs. It would appear that context (school and social/political) can play an important role in shaping teachers' beliefs as well as sometimes acting to support or hinder the enactment of teachers' espoused beliefs. In addition to context, teachers' beliefs can also be influenced by other factors internal to the teacher such as teacher knowledge, self-efficacy, and experience. Research suggests that teachers' beliefs function to mediate information and predict practice in different ways. Fives and Buehl (2012) contend that beliefs filter information and content entering the cognitive domain, frame particular educational situations or problems and guide teachers' intentions and actions. Teacher belief development and change is a slow and complex process especially given the power of the enculturation process during teachers' own years of schooling (Lortie, 1975). Fives and Buehl (2012) contend that development in teachers' beliefs generally occurs through initial teacher education, teaching experience and professional development support.

Teachers' conceptions of assessment

Assessment can mean many things to many people (Newton, 2007). This section of the chapter focuses on the corpus of international literature exploring teachers' beliefs about the nature and purpose of assessment. It would appear from previous research in this field that the main construct employed in the examination of

teachers' assessment-related beliefs is a "conception" (Barnes et al., 2015). Barnes et al. (2015) suggest that the use of this term "provides a framework for describing teachers' overall perception and awareness of assessment" (p.285). Upon a review of 28 peer-reviewed empirical articles related to this topic, Barnes et al. (2015) identify teachers' conceptions of assessment as existing mainly along a continuum of purposes ranging from an extreme pedagogical perspective on one end, to an extreme perspective for high-stakes accountability purposes on the other. Studies examining teachers' conceptions of assessment vary in size, context, participant profile and methodology. While the review by Barnes et al. (2015) captures a wide range of the studies focusing on teachers' conceptions of assessment, a number of other key studies have been conducted since its publication and are included in this review to further build on the knowledge base for the topic. As a means of organising the discussion in this section, the studies reviewed are grouped according to the methodological approach underpinning them beginning with those that adopted a quantitative approach.

Quantitative studies on Teachers' Conceptions of Assessment

Upon review of the quantitative literature on teachers' conceptions of assessment, it became apparent that a considerable amount of cross-cultural research had been conducted using Brown's (2006) Teachers' Conceptions of Assessment Inventory (TCoA-III). The TCoA-III (Full and Abridged versions) is a self-report multidimensional survey instrument designed to elicit teachers' level of agreement with four contrasting purposes for assessment (i.e., improvement, school accountability, student accountability, and irrelevance). Brown (2008) has claimed that three main purposes for assessment are well established in the literature. These are improvement of teaching and learning, student accountability and school accountability. Drawing on the work of Crooks (1988), Sadler (1989), Scriven (1991), Black and Wiliam (1998),

Popham (2000) and others, the improvement purpose focuses on assessment as informing teaching instruction and guiding student learning. This view aligns with the formative use of assessment as an ongoing part of the teaching and learning process. Rooted in literature by Kahn (2000), McMillan (2001), Guthrie (2002), Broadfoot (2002), Gipps (2002) and others, Brown (2008) argued that the notion of assessment as holding students accountable can be seen in the judgement of student performance against criteria or standards, the assignment of grades and the certification of students based on scores. While it must be acknowledged that the judgement of student performance through grades, criteria or standards can have a formative function when used to provide feedback to students about their current level of attainment, this conception of assessment relates to the judgement of student performance for summative purposes only. Brown's third purpose for assessment, grounded in the work of Gipps, Brown, McCallum & McAlister (1995), Linn (2000), Darling-Hammond (2003) and others, refers to the use of assessment for accountability purposes to determine the quality of schools' and teachers' performance. This conception of assessment aligns with the performativity and accountability culture outlined in the introduction. Finally, in addition to these three purposes, Brown (2008) argued for the inclusion of a fourth conception or *antipurpose* which maintains that assessment is essentially bad or irrelevant to the teaching and learning process. Drawing on the work of Torrance and Pryor (1998) and Hall (2000) amongst others, Brown (2008) maintained that there are number of reasons for teachers to conceive of assessment as irrelevant to their work. He noted that teachers may believe that assessment impacts unfairly on certain students, that they are being forced to implement assessment that they do not believe in and therefore choose to ignore it, or that assessment may be seen as inaccurate and therefore unreliable.

Based on his understanding of assessment in terms of these four conceptions, Brown (2004) developed the TCoA-III (full version) with a large sample of New Zealand primary teachers and the TCoA-IIIA (abridged version) was subsequently validated with large samples of Queensland primary and secondary school teachers (Brown, 2006). The instrument allows for the interpretation of mean scores for the four scales as well as more complex interpretations of the intercorrelations between the scales. It has been posited by Brown and colleagues (Brown & Harris, 2009; Brown et al., 2011) that teachers' conceptions of assessment are ecologically rational (Rieskamp & Reimer, 2007) in that they reflect the social and cultural priorities of a particular policy context. The notion of assessment beliefs as ecologically rational constructs is echoed in an article by Elwood and Murphy (2015) in which they discuss assessment and its associated cultural scripts. The authors maintain that as an educational practice, assessment is subsumed under a broader cultural and social structure which projects certain cultural beliefs onto society. They argue that:

These cultural beliefs are embedded in our thinking about learners, learning and assessment and inscribed in our routines and behaviours as we engage with, and in, assessment activity. As learners and their assessors participate in assessment practices, they produce, reproduce and transform an instance of 'collective life', of society. (Elwood & Murphy, 2015, p.184)

The studies which have employed Brown's (2006) TCoA-III (full and abridged) appear, for the most part, to support the proposition of ecological rationality. Given that Brown's (2006) survey instrument is so prolific within the quantitative body of literature on teachers' conceptions of assessment, a broad range of these studies are now discussed.

The line of research using Brown's (2006) TCoA-III (full and abridged) spans many countries, education systems and cultures. The full version of the TCoA-III was

originally validated in New Zealand when Brown (2004) used it to examine primary teachers' conceptions of assessment. When this study was conducted, assessment practices in New Zealand primary schools were largely orientated around low-stakes classroom-based assessment (Crooks, 2011). Brown (2004) noted that emphasis was placed at this level on "voluntary, school-based assessment for the purpose of raising achievement and improving the quality of teaching programmes" (Brown, 2004, p.306). Brown (2004) reported that primary teachers ($n=525$) endorsed the improvement conception as their dominant purpose for assessment. They also agreed with the conception of assessment as school accountability but they rejected the notion that assessment holds students accountable and that assessment is irrelevant. The teachers' endorsement of assessment as a tool to improve teaching and learning was unsurprising given the assessment for learning culture embedded at primary level in New Zealand at that time. It would be interesting to see whether the teachers' rejection of the *student accountability* conception would still hold today in light of the New Zealand Government's introduction of standards-based reporting of achievement at primary level in 2010 (Crooks, 2011). A somewhat unexpected finding from this study was a correlation ($r=.58$) between the improvement and school accountability conceptions, that is teachers who tended to agree with assessment for the purpose of improving teaching and learning were also likely to agree with the idea of assessment for school accountability purposes. Brown (2004) proposed that this may have been due to the self-management of New Zealand schools where teachers were "accountable to their colleagues and to a school-based board of trustees made up of parents of pupils for the effectiveness of their work in changing student learning outcomes" (p.313). Another finding from this study was that the author reported no statistically significant differences in mean factor scores regardless of teacher gender, years of training, years

of experience, or role. This could have been due to data in this study being from a “relatively homogenous sample of New Zealand primary school teachers” (Brown, 2004, p.311).

In a subsequent study comparing New Zealand primary ($n=573$) and secondary school ($n= 404$) teachers’ conceptions of assessment (Brown, 2011), confirmatory factor analysis validated Brown’s (2006) original four-factor model. The teachers from the two samples demonstrated similar levels of agreement with regard to three of the conceptions (improvement, school accountability and irrelevance) but a statistically significant difference in mean scores was found in relation to the conception of assessment as holding students accountable. As might have been expected in light of assessment for certification at second level in New Zealand, this conception was endorsed more strongly by the secondary teachers. As noted by Brown (2011), this finding is “consistent with real-world differences in how assessment is used at the primary and secondary levels of schooling in New Zealand” (p.14). New Zealand second level teachers are regularly involved in evaluating and grading students given the requirements of the state examination system at this level. However, despite the difference in teacher responses on this sub-scale, their agreement across three of the conceptions is noteworthy and may suggest that the educational culture across the two sectors in general has led to a “a common set of professional attitudes” (Brown, 2011, p.14) among the two groups of teachers. Interestingly, however, while the broad sample of secondary teachers in the study shared a generally common understanding of assessment with their primary colleagues, the findings of recent research (Yates & Johnston, 2017) focusing solely on senior-level secondary teachers in New Zealand diverge from Brown’s (2011) findings.

Yates and Johnston (2017) used Brown's (2006) TCoA-IIIa to investigate the impact of school-based assessment for qualifications on New Zealand secondary teachers' ($n = 135$) conceptions of assessment. The authors note that the "participating sample comprised teachers who work almost exclusively at the senior level of high school and are responsible for school-based assessment programmes that contribute at least 50% of students' final grades towards qualifications" (Yates & Johnston, 2017, p.1). When confirmatory factor analysis indicated an inadequate fit of the data to the original model, maximum likelihood exploratory factor analysis was conducted resulting in the extraction of a six-factor solution. Two of these factors refer to the use of assessment for summative purposes, one points to its formative use, another describes assessment as a valid and reliable practice, one factor describes assessment as detrimental to learning and one factor aligns with Brown's (2006) school accountability factor. A key finding from this study was the emergence of a new factor which the authors labelled *assessment is for qualifications*. Yates and Johnston (2017) interpreted this factor as being illustrative of the ecological influence of "summative assessment for qualifications on high school teachers' conceptions of assessment" (p.15). The authors highlighted some positive correlations ($r = .354$ / $r = .210$) between the formative and summative assessment purposes that emerged. They contend that these correlations could "indicate a tendency for teachers to see a dual purpose for National Certificate of Educational Achievement (NCEA) school-based assessment" but could also be indicative of "tensions between using assessment for both formative and summative purposes" (Yates & Johnston, 2017, p.14). The authors explain the second argument further stating that the weak correlations might point to an uncertainty amongst teachers as to what actually constitutes formative and summative assessment. They argue that the teachers may conceive of formative assessment "in line with

Torrance's (2007) view that compliance with assessment criteria can be a proxy for actual improvement of learning" (p.14). A relatively weak positive correlation emerged also between assessment for school accountability and assessment for the purpose of qualifications ($r = .165$). It was suggested that these teachers may not totally agree with the "notion that school quality can be measured through assessment results, in particular when those assessments are also used to award qualifications" (p.11). Yates and Johnston (2017) concluded overall that the teachers participating in their study displayed conceptions of assessment that align more closely with the results from high-stakes examination contexts than with Brown's previous (2011) New Zealand sample of primary and secondary teachers. The pressure of internal assessment for certification purposes appears to be impacting on teachers' conceptions of assessment in New Zealand. Crooks (2011) noted in a commentary on second level in this context that "classroom assessment is a complex mixture of assessment for learning and high-stakes summative assessment" (p.76). It would appear from the results of the factor analysis in this study that the goal of formative assessment can be a difficult one to achieve in the face of accountability pressures from internal assessment for accreditation purposes. Given that the Irish education system is currently implementing internal assessment at second level, the findings from this study further justify and highlight the need to investigate Irish teachers' conceptions of assessment at this point.

Brown, Lake and Matters (2011) examined primary ($n=784$) and secondary ($n=614$) teachers' conceptions of assessment in Queensland using the TCoA-IIIa. Similar to New Zealand, assessment practices in Queensland at the primary and lower secondary school levels are low-stakes in nature and during the last two years of secondary school a "rigorous system of externally moderated school-based assessment" (Brown et al., 2011, p.213) is applied. Using confirmatory factor analysis,

the original TCoA-IIIA was deemed admissible for the Queensland teachers with the addition of two extra paths. In terms of the mean conception scores across the two groups, primary teachers agreed slightly more than secondary teachers that improvement of learning and teaching is the dominant purpose of assessment whereas secondary teachers placed slightly more emphasis on student accountability. This finding is consistent with Brown's (2011) New Zealand study. The tensions which can sometimes exist in relation to multiple assessment purposes were also noted in this study. In the case of the two groups of teachers, school accountability was moderately correlated with student accountability ($r = .38$ [primary], $r = .44$ [secondary]) and improvement ($r = .43$ [primary], $r = .45$ [secondary]). The authors remarked in their analysis that this correlation suggested that "accountability at the school level, assessing students and improvement were intertwined rather than juxtaposed" (Brown et al., 2011, p.217) and that Queensland teachers "did not exhibit the simplistic notion of formative assessment good, summative assessment bad" (p.217). These type of correlations reflect the inter-connected nature of beliefs as presented by Green (1971) and Rokeach (1968). The findings also highlight the complex inter-relationship between potentially conflicting conceptions of assessment such as assessment for the purposes of student learning and assessment for accountability purposes.

Another comparative study of primary and secondary teachers' conceptions of assessment was conducted in Cyprus (Brown & Michaelides, 2011). In the Cypriot context, "assessment is explicitly discussed in the formal curricula as an indispensable part of the teaching-learning process" (Brown & Michaelides, 2011, p.322). Teacher-designed classroom-based assessments are implemented throughout primary and secondary school with the exception of a high-stakes national exam for certification purposes in the final year of secondary school (Brown & Michaelides, 2011).

Confirmatory factor analysis revealed that the original TCoA-III model did not fit the Cyprus data and so exploratory factor analysis was conducted to find the most likely model for the data. A model was found which reduced the conceptual structure of the original questionnaire to two major conceptions – a positive and negative orientation towards assessment. The positive conception consisted of three subordinate factors (*assessment improves teaching*, *assessment improves student learning*, and *assessment holds schools accountable*) and the negative conception consisted of two subordinate factors (*assessment is bad* and *assessment is ignored*). The teachers endorsed the positive conception of assessment to a greater degree than the negative one, with student learning being the factor most strongly supported. This finding is consistent with the underpinning philosophy of the Cypriot educational system which views assessment as a core part of the teaching and learning process in which teacher judgement plays a crucial role. The authors noted that the teachers' endorsement of school accountability within the positive conception of assessment is most likely related to the nature of the educational system itself. They suggested "that Cypriot teachers conceive that evaluating schools with assessments is legitimate, since the assessment system and policy are consistent with high respect for and trust in teachers' professionalism in evaluating, monitoring, and responding to student learning progressions" (Brown & Michaelides, 2011, p.331). The idea of systematic respect and trust in teacher judgements is sometimes referred to as "intelligent accountability" (Smith, 2016). Drawing on the work of Crooks (2003), Smith (2016) notes intelligent accountability "adheres to specific criteria such as trust among the key participants – students and teachers, teachers and parents, teachers and the school leadership and, above all, the politicians' trust in teachers and their professionalism" (p.750). This idea of trust among stakeholders is an important issue to consider in the context of reform in

Ireland. Perhaps teacher resistance to assessment change in Ireland is underpinned to some degree by a lack of trust in teacher professionalism nationally? As mentioned previously in the introductory chapter, the dominance of high-stakes assessment in Ireland has created a cultural understanding of external examinations as the most trustworthy type of assessment. Might it be that the power of such a narrative undermines the professional judgement of teachers, leaving them in a vulnerable position as assessors? Perhaps a lack of trust in teacher judgement can lead to teachers themselves questioning their role as assessors and sometimes, as a result of this, resisting engagement with such a role. The results from the Cyprus study would suggest that a cultural mind-shift regarding teacher professionalism is a necessary ingredient for successful implementation of school-based assessment. This finding resonates with Murchan's (2017) argument, highlighted in the introduction, on the needs for educational stakeholders and the public at large to value teachers' judgements if school-based assessment in Ireland is to work.

Similar to the study in Cyprus (Brown & Michaelides, 2011), exploratory factor analysis using the TCoA-IIIA was also conducted in the Netherlands with a group of 351 Dutch secondary teachers (Segers & Tillema, 2011). The secondary school system in the Netherlands consists of ongoing formative assessment as well as a national examination at the end of upper secondary which provides access scores for entry to tertiary institutions. Half of the exam is developed externally and half of it is developed internally by schools. Results from the analysis indicated a four-factor model which differed somewhat from Brown's (2006) original structure. At the time of inquiry, Dutch secondary teachers expressed a belief in assessment as a construct which (1) informs performance and learning, (2) holds schools accountable, (3) is inaccurate and contains measurement errors and (4) guides instructional decisions and measures

higher order thinking skills. One of the key findings from this study relates to the first factor. Dutch teachers did not make a distinction between formative and summative purposes for assessment. Segers and Tillema (2011) suggest that this finding may be a reflection of the Dutch secondary system where classroom assessment serves both a formative and summative purpose. The extraction of the fourth factor, which was new, was attributed to national debate at the time emphasising a broader use of assessment moving beyond knowledge recall and the use of assessment to inform and adapt teacher practice. This finding once again highlights the influence of ecological priorities (Brown & Harris, 2009) on teachers' conceptions of assessment.

In a similar vein to the contexts discussed thus far, Spain also has a low-stakes accountability system with the intention being to ensure that assessment is “continuous, formative, and holistic” (Brown & Remesal, 2012, p.77). A Spanish translation of the TCoA-IIIA was used to investigate 672 preservice teachers' conceptions of assessment in Spain (Brown & Remesal, 2012). Brown and Remesal (2012) note, however, that at the time of this study a pilot initiative was in place in the autonomous community of Catalonia to introduce external evaluation of student achievement at the end of primary education. It is worth noting that this was the area of Spain from which the study sample was drawn. After conducting a combination of exploratory and confirmatory factor analysis, a model incorporating five factors was found. These factors were labelled assessment improves student learning and teaching, is ignored and inaccurate, is bad, measures school quality validly, and grades students. Apart from the school quality factor, similar mean scores were reported for the remainder of the factors. Brown and Remesal (2012) suggested that the narrow range in the mean scores may have been due to a reluctance on the part of these inexperienced participants to “express extreme opinions” (p.83). Furthermore, the authors highlighted that at the time of this

study the participants were enrolled in an introductory psychology course which did not include assessment as a learning topic, and so may not have been “mentally attuned” (Brown & Remesal, 2012, p.83) to assessment issues. Finally, it is noteworthy that the participants in this study primarily conceived of assessment as a negative construct. Despite the overall low-stakes formative approach to assessment in the Spanish context, Brown and Remesal (2012) noted that a common practice in Spanish schools “to establish within-school achievement levels for differential grouping of students” (p.83) may have been interpreted by the participants in a discriminatory light, thereby negatively affecting their conceptions. This suggestion aligns with Lortie’s (1975) argument regarding teachers’ apprenticeship of observation. The trainee teachers in the Spanish study were in their first year of teacher education and so their experience of assessment throughout their schooling years previous to then may still have been quite live in their minds, and may have exerted an influence on how they conceived of assessment at the time of the study.

Trainee teachers’ conceptions of assessment were also the focus of a Canadian study using the TCoA-IIIA (Daniels, Poth, Papile & Hutchison, 2014). The Canadian context is one in which standardized tests form an established part of the assessment culture. In addition to standardised testing, “external agencies such as the Fraser Institute publish report cards that rank individual schools according to their performance on provincially administered tests” (Volante & Earl, 2013). Despite these summative accountability measures, Daniels et al. (2014) noted that all provinces in Canada involve teachers in the construction of standardised tests and that teachers are being provided with professional development opportunities in formative assessment. They also highlighted the fact that preservice teachers had previously reported (Lejeune, Poth, & Daniels, 2010) experiencing formative assessment methods

throughout their schooling. The sample for the study consisted of 436 prospective teachers who were enrolled on their first assessment course as part of their teacher education programme. Upon completion of confirmatory factor analysis, a model of best fit was found which comprised of nine conceptions of assessment which corresponded exactly to the nine first order factors on the original TCoA-IIIa (Brown, 2006). In general the improvement-orientated factors were rated most highly by participants and the negative factors were rated lowest. A significant exception to this pattern, however, lay in the prospective teachers relatively strong endorsement of the *assessment is inaccurate* factor. This finding is similar to the Spanish prospective teachers' endorsement of assessment as a negative construct. Daniels et al. (2014) suggest that as the teachers were learning about assessment error and inaccuracy at that time, they may have been "hypersensitive to these issues" (p.153). Apart from Brown (2004), this Canadian study is the only one to date (known to the researcher) to test for gender differences across the conceptions. Respondents did not differ systematically based on gender.

Turning now to more high-stakes contexts, a Chinese translation of the TCoA-IIIa was used to examine Hong Kong teachers' conceptions of assessment (Brown, Kennedy, Fok, Chan & Yu, 2009). The education system within Chinese contexts places a very strong emphasis on high-stakes examinations which are used to determine placement into different levels of education and entry into third-level institutions. Rote learning, direct transmission of information and strictly controlled teaching to the test are all common features of the examination-orientated culture within Chinese contexts (Brown & Gao, 2015). In addition to the accountability purposes for assessment as mentioned above, the importance of summative assessment is also deeply engrained in Chinese culture to the extent that one's self-worth appears to be determined by it.

Brown, Hui, Yu and Kennedy (2011) note that “in the Chinese society of Hong Kong, high academic performance demonstrates perfection and superiority of character” (p.309). As part of a reform agenda towards more formative assessment in Hong Kong, the Assessment for Productive Learning Project was developed which involved sixteen schools. The sample of teachers for the study by Brown et al. (2009) was taken from these schools. A modified but conceptually equivalent model to that developed by Brown (2006) was extracted from the data. While the Hong Kong teachers in this study endorsed the improvement conception in a similar way to the studies previously discussed, the key finding from this study was the extremely high correlation ($r = .91$) between improvement and student accountability. The teachers saw the improvement of learning as being almost synonymous with holding students accountable. Notwithstanding these teachers engagement in a reform project, the strength of the high-stakes culture in Hong Kong remained to the fore in their beliefs, reflecting once again the arguments by Brown and Harris (2009) and Elwood and Murphy (2015) on the influence of cultural priorities on beliefs.

To examine the Chinese context further, Brown et al. (2011) examined the beliefs of 1,912 primary and secondary teachers from Hong Kong ($n=1014$) and Guangzhou ($n=898$). Grounded in data from small scale qualitative studies (Hui, 2012; Wang, 2010), a context-specific version (C-TCOA) of Brown’s (2006) questionnaire was developed which included the two new constructs of development and control. Repeated data analyses resulted in the reduction of the original factors into a hierarchical three factor model – *assessment for improvement*, *assessment for accountability* and *assessment is irrelevant*. Teachers from the two regions responded very similarly to the questionnaire. Consistent with Brown et al. (2009), improvement was strongly correlated with accountability ($r = .80$). This finding once again reinforces

the influence of “the Chinese tradition and policy of using examinations to drive teaching quality and student learning and as a force for merit based decisions” (Brown et al., 2011, p.307).

A three-factor model was also found in a study of pre-service and practicing teachers in Egypt (Gebril & Brown, 2014). Similar to China, Education in Egypt is also examination-driven. Summative examinations are used at all levels of schooling to select students for access to further educational opportunities. An Arabic translation of the TCoA-IIIa was administered to a group of pre-service teachers in the final year of teacher education ($n = 305$) and a group of full-time practising teachers who were enrolled on a diploma in education by night ($n = 202$). Using confirmatory factor analysis, the teachers’ beliefs about assessment were conceptualised as a three-factor model – *improvement*, *school accountability* and *irrelevance*. Both groups of teachers most strongly endorsed the improvement conception of assessment but the model also indicated a strong relationship ($r = .89$) between improvement and school accountability. This finding would appear ecologically rational given the high-stakes environment in which Egyptian teachers work. This strong correlation aligns with the findings from the previously discussed Chinese studies (Brown et al., 2009; Brown et al., 2011). The authors argued that “greater changes to the examination system are required if teacher beliefs are expected to be more positive about the priority of formative, improvement-oriented uses of assessment” (Gebril & Brown, 2014, p.16). In light of the dominance of high-stakes examinations in Ireland to date, it was deemed of interest in this study to investigate whether the Irish data would reveal a correlation between accountability and improvement as extracted from the Chinese and Egyptian contexts.

Qualitative Studies on Teachers' Conceptions of Assessment

In addition to the wealth of quantitative research in this field, a range of qualitative studies are also available. Remesal (2011) analysed interview data and assessment material from 30 primary and 20 secondary mathematics teachers in Spain. The study was carried out in the context of systematic school reform focusing on formative assessment. The research aimed to elicit data regarding teachers' beliefs about assessment with respect to teaching, learning, students' accreditation of learning and teacher accountability. Interview data were categorised and a four-dimensional bi-polar model of teachers' conceptions of assessment was formulated. Within this model, teachers' beliefs about assessment in relation to teaching, learning, accreditation of learning and teacher accountability were positioned along a continuum between two poles. One pole represented the "pedagogical function" (Remesal, 2011, p.473) of assessment which presents assessment as "a device capable of promoting reflection and change in education by monitoring both teaching and learning" (Remesal, 2011, p.473) and the other pole reflects the "societal function" (Remesal, 2011, p.473) of assessment which holds that assessment is a "tool for certification to different audiences in society, such as families and school administrators" (Remesal, 2011, p.473). The participants' overall conceptions of assessment were identified in the following ways. If one's beliefs for the four dimensions could all be positioned on one of the two poles, then that participant was labelled as having an extreme pedagogical or extreme societal conception of assessment. Those participants whose beliefs for the four dimensions were located at different poles in the ratio of 3:1 were characterised as having a mixed pedagogical or mixed societal conception of assessment depending on the dominant pole. The data revealed that more than twice as many teachers held mixed conceptions (i.e. 3:1 societal or pedagogical) as extreme ones. It was noted by the author that this

finding reflects the complex nature of assessment within schools. The most frequent conception within the full sample of participants was a mixed societal conception. However, when the distribution of conceptions across levels was considered, the primary teachers' conceptions were mainly pedagogical in nature whereas the secondary teachers had predominantly mixed or pure societal conceptions. This finding, which mirrors some of the findings from aforementioned quantitative studies (Brown, 2011; Brown et al., 2011), points once again to the prioritisation of summative assessment by teachers at secondary level, even within a reform context. Remesal (2011) notes, however, that at the time of data collection for the study the formative assessment mandate in Spanish schools may not have completely taken root and that were the study to be replicated at some point in the future, the results may be different.

Davis and Neitzel (2011) conducted semi-structured interviews with 15 middle school teachers in the United States to elicit their beliefs about the forms and functions of classroom assessment. In order “to avoid creating a situation in which teachers would feel compelled to give socially desirable answers” (Davis & Neitzel, 2011, p.204) the interviews were based around five tasks. The tasks involved sketching what assessment looks like, discussing a portfolio of assessment activities that the teachers had compiled, choosing between a list of assessment metaphors, completing a card sorting activity related to assessment tools and responding to a hypothetical scenario. In relation to assessment forms, the data revealed that the teachers conceived of assessment opportunities as being artifactual (written tests, written assignments projects, student-produced writing) and transient (interactions, verbal explanations/presentations, in situ class activities). These two labels, which the authors borrowed from the work of Sadler (1989), are used to distinguish between assessment opportunities that produce “tangible products and those that happen through

interactions” (Davis & Neitzel, 2011, p.207). In terms of assessment functions, the teachers conceived of assessment as having ten distinct purposes in relation to four different audiences. The audiences in question were teachers, students, parents and state and district level “higher ups” (Davis & Neitzel, 2011, p.208). Four of the purposes described how assessment informs teachers: assessment is used to evaluate and inform instruction, identify students for remediation, evaluate student learning and gauge student investment in learning. Three of the functions related to assessment as it impacts upon students: assessment is used to hold students accountable, to guide and expand student knowledge and to provide feedback to students. Participants also believed that assessment is used to inform parents and in the case of state and district level audiences that it is used to prepare for high-stakes testing and that it holds teachers accountable. The findings presented in this study point once again to a sense of plurality in teachers’ conceptions of assessment and also reflect the concept of a continuum of assessment purposes with some being pedagogically focused and others aligning more with the accountability end of the continuum.

In a similar vein to Davis and Neitzel (2011), Karp and Woods (2008) also identified conceptions of assessment with respect to different audiences. They investigated preservice physical education teachers’ beliefs about assessment ($n=17$) while enrolled on a teacher education course. Influenced by their own apprenticeship of observation (Lortie, 1975) throughout their school years, the participants provided interview and survey data outlining their conceptions of assessment with respect to teachers and students. Ranked in the order of most important to least important, the participants believed the purpose of assessment for teachers was to facilitate increased student learning, to show achievement of standards, to evaluate teacher effectiveness and to determine students’ level of knowledge and skills. In relation to students, the

two purposes identified for assessment were to show them where they are in relation to goals or standards and to motivate them. The majority of these conceptions fall along the pedagogy-accountability continuum but the notion of assessment as a motivator is new. Referring to the idea of assessment as a motivating tool, Barnes et al. (2015) note that alignment of such a conception with one or other end of the assessment purposes continuum depends on how the assessment in question is employed in context. In other words, one would need to know whether the motivation was triggered by formative or summative assessment practices. In the case of Karp and Woods' (2008) investigation, reference to the ideas of competition and comparison in the context of motivation would suggest that the participants in this study viewed summative assessment as motivating.

The tensions between conflicting purposes of assessment came to the fore in a study of graduate trainee secondary teachers' ($n=17$) beliefs about the nature and purposes of assessment in England (Taber et al., 2011). The participants' conceptions of assessment aligned mainly with the accountability end of the assessment purposes continuum. Assessment was primarily viewed as having a summative function. The ideas of measurement and quantification of learning appeared regularly throughout the data. Participants also referred to assessment as an important accountability measure for external stakeholders. Assessment was considered to have a motivating function when delivered in the form of results and grades. Some of the more pedagogically-focused conceptions that appeared were the ideas that assessment informs decision-making and that assessment is an effective means of enquiring into one's own teaching practice. Notably very few participants conceived of assessment as a support for learning. Despite efforts to champion the "orthodoxy of AfL" (Taber et al., 2011, p.181), the primacy of summative conceptions of assessment among the participants

would suggest, in a similar way to the findings from James and Pedder (2006), that the constraints of high-stakes testing can supersede the goals of formative assessment.

Mixed-Method Studies

Similar to Karp and Woods (2008), the idea of assessment as an extrinsic motivator for students was also identified as one of seven purposes ascribed to assessment by a group of 26 New Zealand teachers participating in a phenomenographic study (Harris & Brown, 2009). Harris and Brown conducted the study in order to assess the adequacy of Brown's (2006) model of conceptions of assessment. From a group of 161 teachers who had completed Brown's (2006) TCoA-III A, 26 teachers "with noticeably different conception profiles" (Harris & Brown, 2009, p.368) were selected for interview. Systematic analysis of the interview data led to the extraction of seven purposes of assessment. They were compliance, external reporting, reporting to parents, extrinsically motivating students, facilitating group instruction, teacher use for individualising learning, and joint teacher and student use for individualising learning. Harris and Brown (2009) contended that the seven purposes extracted could be subsumed under three major purposes for assessment – accountability, student improvement and irrelevance. The authors argued, therefore, that this finding provided independent validation for the categories in Brown's TCoA-III A (2006). External reporting, reporting to parents and extrinsically motivating students were all seen as coming under the umbrella of accountability. More specifically, the reporting purposes were linked to school accountability and extrinsic motivation by grades and qualifications were associated with student accountability. The categories of facilitating group instruction, teacher use for individualising learning, and joint teacher and student use for individualising learning were all deemed to be aligned with an improvement orientation towards assessment. The compliance purpose

was subsumed under the irrelevance conception. Participants spoke of having to comply with national mandates and school-wide directives regarding standardised testing and other practices which they rejected as “irrelevant, inaccurate, or negative for teachers, pupils, and learning” (Harris & Brown, 2009, p.370). The notion of compliance as being irrelevant to teachers’ thinking about assessment ties in with the previously discussed idea of national and school policy acting as factors which can hinder belief enactment (Buehl & Beck, 2015). Overall this study demonstrates that teachers view assessment as having a multi-faceted range of purposes which are sometimes conflicting in nature. Attention is drawn to the difficulty for teachers with “balancing school and student needs” (Harris & Brown, 2009, p.377). Some of the key tensions highlighted in the study were those between school and student, compliance and improvement and positive and negative evaluation.

Summary

Belief research is a complex and nuanced field of inquiry spanning multiple disciplines, theoretical paradigms and methodological approaches. The elusive nature of beliefs as a psychological construct has led to a domain of research characterised by a significant amount of terminological and conceptual variance. Notwithstanding the overall lack of consensus within the literature generally, a trend would appear to be emerging towards acceptance of the term ‘conception’ as the primary construct for examining teachers’ assessment-related beliefs.

Research would suggest that teachers’ beliefs develop throughout their lifetimes from an early age (Beswick, 2007; Brownlee, 2003; Buehl & Fives, 2009; Cabaroglu & Roberts, 2000; Fives & Buehl, 2012; Shulman, 1987; Tam, 2015b; Yuan & Lee, 2014).

By the time teachers begin their teaching careers, they already possess an established set of beliefs (Kagan, 1997; Levin & He, 2008; Lortie, 1975; Pajares, 1992). These beliefs exist as part of an interconnected multidimensional system in which beliefs may be central or peripheral and which allows for the existence of conflicting beliefs (Green, 1971; Rokeach, 1968). Much research has acknowledged the fact that teachers' beliefs cannot be disassociated from the contexts in which they occur (Fives & Buehl, 2012; Levin, 2015; Mansour, 2009; Pajares, 1992). This notion of context includes both the school environment as well as the larger social and political sphere. In line with Bandura's (1997) theory of triadic reciprocal determinism, there appears to be a reciprocal relationship between teachers' beliefs, experiences and context (Fives & Buehl, 2012).

Fives and Buehl (2012) provided a useful categorisation framework for the role of teachers' beliefs. They argued that teachers' beliefs may filter information entering the cognitive domain, frame particular situations, tasks or problems and guide teachers' intentions and actions. Teachers' beliefs are essentially a cognitive screen through which new information is interpreted. When considering teachers' beliefs as a guiding force for action or practice, an element of caution must be exerted. While one's beliefs can and often do predict one's practice, the relationship between the two may not always be linear in nature. The enactment of one's espoused beliefs can sometimes be hindered by internal and external mediating factors (Buehl & Beck, 2015).

In terms of the literature on conceptions of assessment, a strong line of quantitative cross-cultural research has been established using Brown's (2006) TCoA-III. These studies appear to support Brown and Harris's (2009) proposition that teachers' conceptions of assessment are ecologically rational (Rieskamp & Reimer, 2007) in that they tend to reflect the policy priorities and educational traditions

experienced by teachers within a particular society. Studies have shown that in high-stakes examination-driven societies such as Egypt (Gebril & Brown, 2014) and China (Brown, Hui, Yu & Kennedy, 2011), teachers' conceptions of assessment are predominantly consistent with an understanding of assessment for accountability and evaluation purposes, whereas in more low-stakes environments such as New Zealand (Brown, 2011), Queensland (Brown et al., 2011) and Cyprus (Brown & Michaelides, 2011), teachers are more committed to an improvement-oriented notion of assessment. In studies comparing primary and secondary teachers (Brown, 2011; Brown et al., 2011), secondary teachers tended to agree more with the notion of assessment for student accountability purposes. This finding was seen as being in line with the reality of how assessment is used at the two levels. The intercorrelations between different assessment purposes identified in these studies demonstrates the complex multidimensional nature of assessment and the tensions which can arise as a result of this.

It would appear, from both the quantitative and qualitative studies examined, that teachers' conceptions of assessment can be positioned along a continuum of assessment purposes ranging from pedagogical conceptions to summative conceptions of assessment for accountability purposes. In addition to falling along this continuum, findings from some of the qualitative studies point to teachers' conceptualisations of assessment as corresponding to different audiences (Davis & Neitzel, 2011; Harris & Brown, 2009; Karp & Woods, 2008). In line with the previously discussed quantitative literature, these studies suggest that teachers hold numerous and sometimes conflicting conceptions of assessment which can sometimes be challenging to balance and reconcile. This study attempts to contribute to the above body of literature by unearthing Irish post-primary teachers' conceptions of assessment and investigating the

nature, strength, interconnectedness and potential implications of these conceptions in the context of Junior Cycle policy reform. While conceptions are the main focus of this work, it is important to remain mindful of their interrelationship with other dimensions of teachers' engagement with assessment. Recent research offerings from Xu and Brown (2016) and Looney et al. (2017) provide succinct conceptual models which draw together the contents of this review and position its focus within the broader frames of teacher assessment literacy in practice (Xu and Brown, 2016) and teacher assessment identity respectively (Looney et al., 2017). These models are now considered.

Models of Teachers' Engagement with Assessment

Teachers' competence in educational assessment is often presented in the literature as teacher assessment literacy (DeLuca, LaPointe-McEwan & Luhanga, 2016; Stiggins, 1991). In its original form, as defined by Stiggins (1991), assessment literacy involved an understanding of "high and low quality assessment" (p.535) and the ability to "apply that knowledge to various measures of student achievement" (p.535).

Stiggins (1991) also argued that assessment literacy involves recognising and taking the appropriate steps to modify inappropriate assessment procedures. In a recent review of assessment literacy standards in a variety of regions since the 1990s, DeLuca et al.

(2016) report "a gradual shift in conceptions of assessment literacy over time" (p.267).

They note that standards initially emphasised "teachers' abilities to construct, administer, and use primarily summative forms of assessment" (DeLuca et al., 2016, p.267) but that the integration of "assessment for learning and assessment education" (DeLuca et al., 2016, p.267) has become more apparent since the year 2000 in light of the surge of literature on AfL since Black and Wiliam's (1998) seminal review.

Notwithstanding this gradual change, recent conceptual analysis of teacher assessment

competence suggests that the construct of teacher assessment literacy may be too narrow and instrumentalist to properly capture the complexity of teachers' engagement with assessment. Xu and Brown (2016) and Looney et al. (2017) present expanded models of teachers' assessment competence in their respective Teacher Assessment Literacy in Practice (TALiP) and Teacher Assessment Identity (TAI). Xu and Brown (2016) reconceptualise teacher assessment literacy as part of an "iterative and dynamic system" (p.149). The system is represented by a pyramid consisting of six interrelated components (See Figure 2). From bottom to top, the components in question are knowledge base, teachers' conceptions of assessment, institutional and socio-cultural contexts, TALiP, teacher learning and teacher identity (re)construction as assessors. The knowledge base represents a "general body of key theoretical principles of assessment" (Xu & Brown, 2016, p.155) that a teacher must acquire for effective assessment practice to take place. Teachers' conceptions of assessment "denote the belief systems that teachers have about the nature and purpose of assessment" (Xu & Brown, 2016, p.156). The authors argue that teachers' conceptions of assessment have both a cognitive and affective dimension. Drawing on the work of Barnes et al. (2015) and Fives & Buehl (2012), addressed earlier in this review, the authors outline the cognitive aspect as teachers' belief systems which act as "an interpretive and guiding framework by which they mediate their uptake of theoretical knowledge and its implementation" (Xu & Brown, 2016, p.156). They ground the affective dimension, on the other hand, in work of Green (1971) on central and peripheral beliefs as previously discussed. The third component in the framework is representative of the micro- and macro-level variables which can influence teacher assessment practice "individually or in concert" (p.157). This component closely aligns with the discussion throughout this review regarding the influence of context in shaping as well as mediating teachers'

conceptions. The authors contend that the reconciliation of these three components within one's assessment work characterises TALiP. It has been illustrated in many studies in this review that this reconciliation process can prove extremely complex especially since “accountability has become the prevailing watchword” (DeLuca & Johnson, 2017) in education. Finally, toward the top of the pyramid, Xu and Brown (2016) conceptualise the advancement of TALiP through teacher learning which in turn can lead to teachers re-evaluating their identity as assessors throughout their professional lives. The inclusion of teacher learning in the framework resonates with the discussion in this chapter on belief development and change through professional learning communities. The overall framework provided by Xu and Brown (2016) clearly encapsulates the complex, inter-related and multi-layered nature of teacher assessment literacy in practice. The framework positions teachers' conceptions of assessment as a key guiding force within this process.

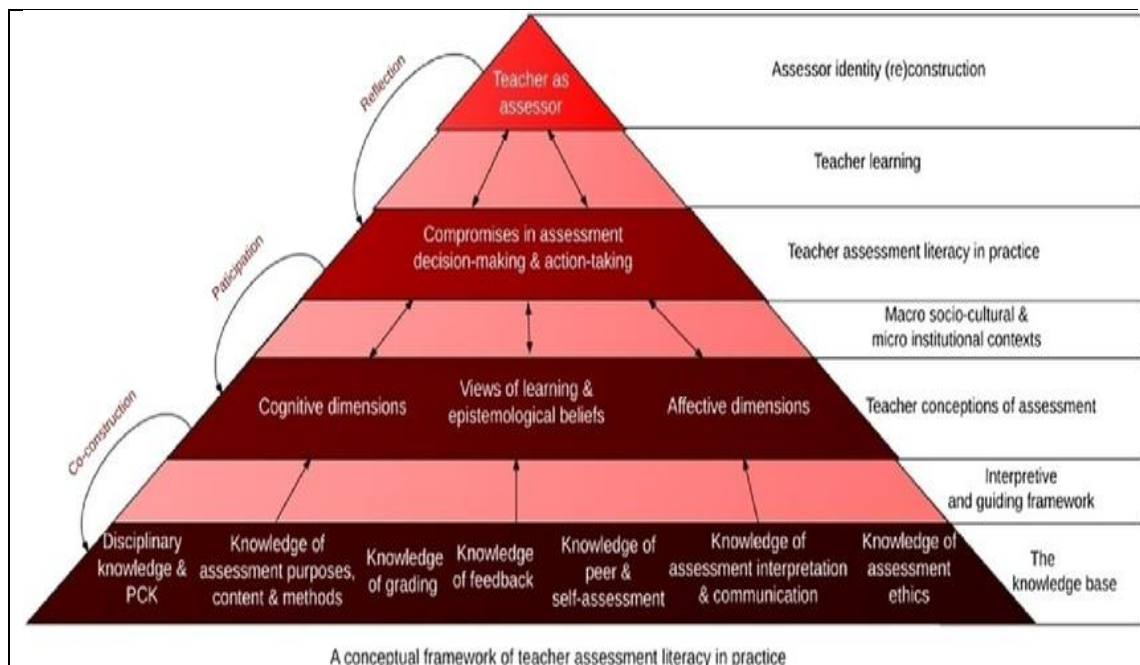


Figure 2. A conceptual framework of teacher assessment literacy in practice (Xu and Brown, 2016, p.155)

Drawing on research on teacher assessment literacy, teacher identity, teachers' beliefs about assessment and teacher self-efficacy, Looney et al. (2017) propose "a dynamic and interactive teacher assessment identity" (p.14) (see Figure 3). While the dimensions in the TAI model closely align with the work of Xu and Brown (2016), Looney et al. (2017) appear to have placed a particular emphasis on research relating to teacher identity in the creation of their framework. The TAI model consists of five dimensions: I know, I feel, I believe, I am confident and My role. The 'I know' dimension reflects assessment literacy as described by Stiggins (1991) and Popham (2009). Drawing on the work of Beijaard, Meijer & Verloop (2004), the emotional aspect of assessment identity is included in the 'I feel' category. The 'I believe' dimension reflects the role that teachers' beliefs play in shaping assessment identity (Broadfoot, 1996; Brown, 2011; Shulman, 1986) and once again highlights the significance of beliefs in teachers' engagement with assessment. Drawing on the work of Bandura (1986), 'I am confident' represents teachers' sense of self-efficacy in relation to assessment. The inclusion of this dimension aligns with the earlier discussion in this chapter of self-efficacy as it relates to collective efficacy and self-efficacy as a predictor of assessment practice (Yan, 2014; Yan & Cheng, 2015). Finally, the 'My Role' dimension, which draws on the work of Pryor and Croussouard (2010) and Ecclestone and Pryor (2003), speaks to the tensions of the dual role teachers have in trying "to strike a difficult balance between being supportive of learners and being critical of them" (Looney et al., 2017, p.7). This idea echoes Xu and Brown's (2016) notion of reconciliation between different components of the assessment jigsaw. Similar to Xu and Brown (2016), Looney et al. (2017) highlight the interconnected nature of the different dimensions by interlinking the disparate factors in their model. While teachers may have the necessary assessment knowledge, for example, they may

not feel confident enough to implement it in practice. On the other hand, teachers may be knowledgeable and confident but may not believe that a certain practice is valuable. Looney et al. (2017) contend that their model of teacher assessment identity allows for broader, deeper and more complete engagement with teacher assessment capability by focusing “not simply on what teachers do, but on who they are” (p.16).

It would appear, therefore, that the two frameworks resonate with each other while demonstrating a nuanced difference in emphasis. Xu and Brown (2016) seem to emphasise the complexity of teacher assessment literacy as it plays out in the context of many interconnected internal and external variables. Looney et al. (2017), on the other hand, appear to focus more on how teachers identify and feel as assessors and how that shapes their engagement with assessment. While the two frameworks are new and therefore largely untested, the findings from the literature review suggest that both models highlight elements which have been shown to be significant to understanding teachers’ conceptions of assessment in the broader context of teachers’ engagement with assessment. For this reason, it is suggested that side by side these models provide a useful conceptual frame for this study.

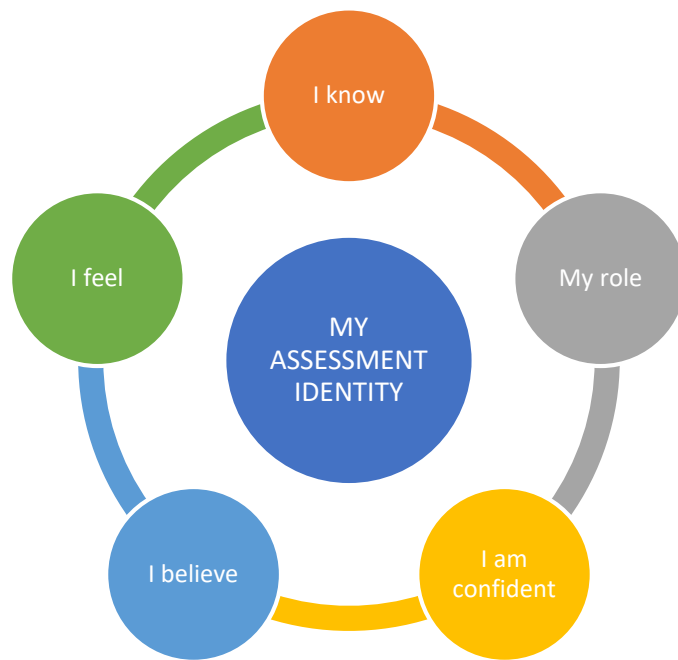
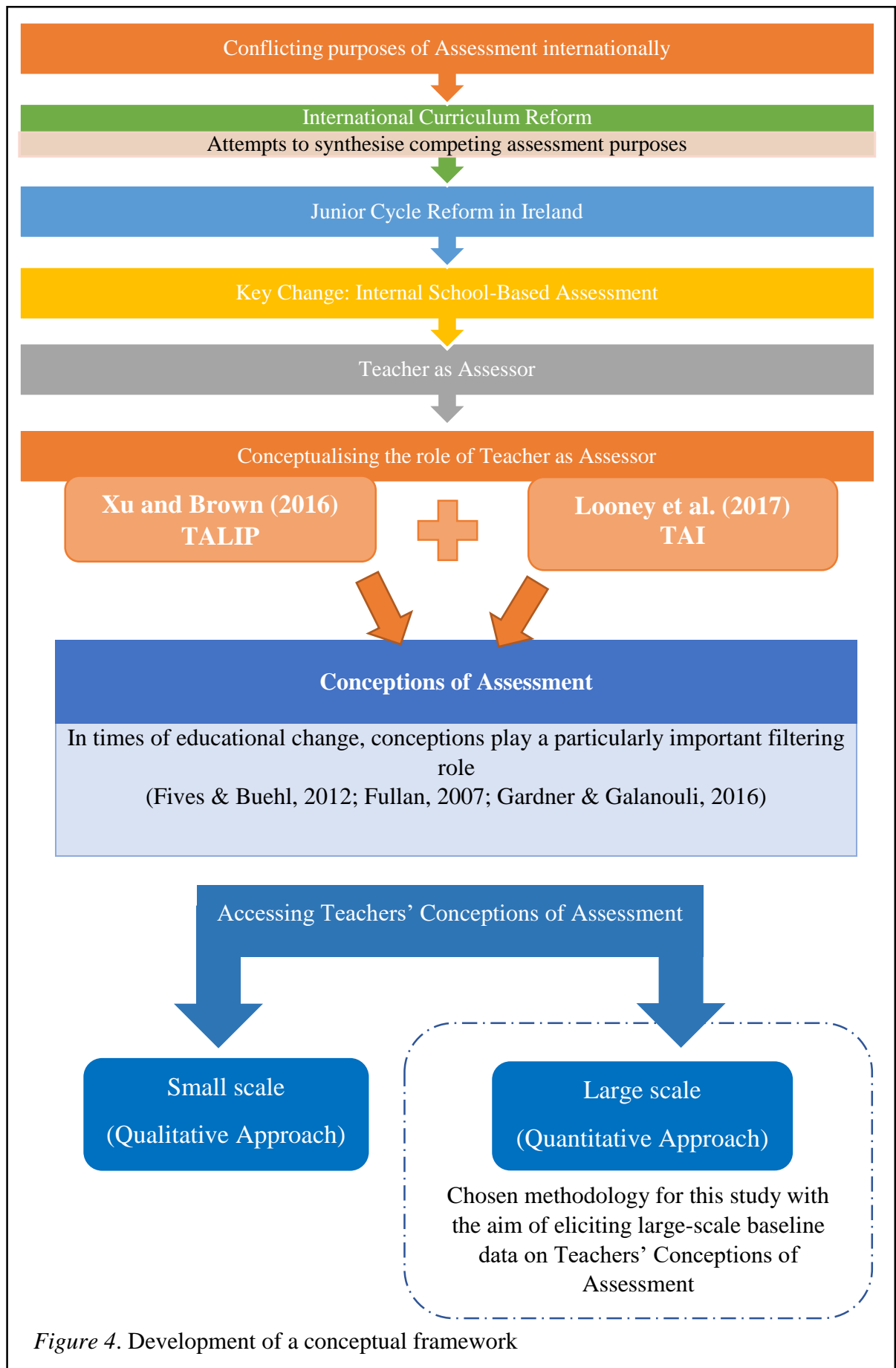


Figure 3. Teacher Assessment Identity (TAI) adapted from Looney et al. (2017)

Chapter 3: Methodology

Introduction

Drawing on the conceptual framework provided in the first section of this chapter (see Figure 4), careful consideration was given to the creation of a research design which would best enable the researcher to elicit baseline data about Irish post-primary teachers' conceptions of assessment. This chapter accordingly outlines and justifies the methodological approach underpinning the research. Four discrete sections are presented; research design, ethics, response data, and data cleaning/preliminary analysis.



Research Design

Introduction

In deciding how best to plan this research project, Creswell's (1998) framework for design, which points to the interconnection of worldviews, strategies of inquiry and research methods, provided a useful starting point. Assuming a pragmatic lens, the research was underpinned by a multi-layered understanding of reality which rejects the "polarization of qualitative and quantitative research" (Hartas, 2010, p.26) in favour of a context-driven approach to inquiry where both hard data and interpretive approaches have a role to play. When considering how best to investigate Irish post-primary teachers' conceptions of assessment within this worldview, it was acknowledged that both quantitative and qualitative approaches could be employed to address various elements of the topic as outlined in the conceptual framework. However, in order to capture the conceptions of a large number of participants in a manageable time period, a purely quantitative approach was adopted for the purposes of this doctoral thesis. A quantitative method of inquiry would allow for the collection of large-scale baseline data on teachers' conceptions of assessment, thus providing the researcher with an initial snap-shot of teachers' current thinking regarding assessment. With an evidence-based starting point in place, these data could then be explored qualitatively in a future study in order to consider broader issues such as contextual influences, the relationship between conceptions and practice, belief development and more.

A non-experimental cross-sectional model was adopted for this study. This approach aims to capture and investigate the views of a representative sample of a particular population at "a frozen moment in time" (Cohen, Manion & Morrison, 2011, p.267). It enables large-scale sampling, comparison between sub-groups within a

sample and macro-level analysis. It does not, however, permit analysis of causal relationships. This contrasts with an experimental design which is undertaken over time and which “seeks to determine if a specific treatment influences an outcome” (Creswell, 2014, p.13). A non-experimental cross-sectional design was deemed appropriate for this research as it enabled collection of data from a large sample of post-primary teachers in order to acquire a snapshot of their current conceptions of assessment. As detailed in Chapter 1, the study specifically aimed to ascertain teachers’ conceptions of assessment in order to inform Junior Cycle policy implementation in Ireland.

Timeline for the Study

This section briefly outlines the steps taken in operationalising the research design discussed above. Firstly, the target population for the study was established. Following this, a key data collection instrument in the field was identified. This instrument (TCoA-IIIa), which is discussed in an upcoming section, was carefully considered in the context of Irish education. Permission was granted from the author to use the instrument in this study. A sampling strategy was then outlined. In advance of disseminating the questionnaire nationally, a pilot study was conducted (in three iterations) with 15 teachers. The questionnaire was adapted in accordance with feedback from the pilot study. Upon completion of the official data collection process, the data were analysed using various methods of statistical analysis.

Target Population

In order to conduct the research, the population of post-primary teachers in Ireland needed to be officially identified. Two key sources were consulted for this information; the Department of Education and Skills (DES) and the Teaching Council.

The DES (2016a) indicated a population size of 26,804 post-primary teachers. This figure refers to the number of post-primary teachers paid by the DES in the year 2015-16. A second population figure of 42,589 post-primary teachers was provided to the researcher by the Teaching Council, the regulatory body for all teachers in Ireland (see Appendix A for email correspondence from the Teaching Council). This higher figure refers to the number of post-primary teachers officially registered with the Council on the 21 February 2017 and is inclusive of teachers who may be working full-time, part-time, as substitutes, on career break or retired but still officially registered. Both of these population figures were used in the study when considering the representativeness of the sample in relation to different categorical variables.² Given the large target population, a questionnaire was chosen as the most practical and appropriate method of data collection for this piece of research. This decision is now discussed in more detail.

Instrumentation

A range of data collection techniques are present within the literature on teachers' beliefs. These include questionnaires, verbal reports, observations, self-reflective essays, portfolios, tests and exams, vignettes, scales, classroom artifacts and metaphor analysis (Bullough, 2015; Schraw & Olafson, 2015). Within the quantitative domain, self-report mechanisms are very common (Hoffman & Seidel, 2015). These attempt to "describe or quantify beliefs and the cognitive underpinnings that guide learning and instruction" (Hoffman & Seidel, 2015, p.108). In a meta-analysis of belief measures in 345 studies related to teachers' beliefs, self-report questionnaires (often incorporating a scale) emerged as the dominant method of belief measurement

² The sample size for this study was 489 participants. Had this been a random sample, the margin of error at a 95% confidence level would be 4.39 for a population size of 26,804 and 4.41 for a population size of 42,542.

(Hoffman & Seidel, 2015). This finding is not surprising given the many uses questionnaires have as a data collection strategy. Schraw and Olafson (2015) note that questionnaires are “easy to administer and score, measure multiple constructs within a single set of questions, are amenable to sophisticated statistical analyses, and provide a comparative baseline across different studies” (p.92). The advantages of questionnaires, notwithstanding, evaluation of self-report data must be “approached with interpretative caution” (Hoffman & Seidel, 2015, p.118) given that these data can be subject to response bias or inaccuracies in reporting.

A key instrument within the extant body of research on teachers’ conceptions of assessment is the abridged version of the Teachers’ Conceptions of Assessment (TCoA-III) Inventory (Brown, 2006). As previously discussed in the literature review, this inventory is well recognised and has been used in many international studies (Brown, 2011; Brown et al., 2011; Brown & Michaelides, 2011; Brown & Remesal, 2012; Daniels et al., 2014; Gebril & Brown, 2014; Segers & Tillema, 2011). Validation of the instrument through confirmatory factor analysis is reported by Brown (2006). Through conducting two confirmatory studies, Brown (2006) established that the instrument had good fit characteristics for teachers in both New Zealand and Queensland. The instrument is a multidimensional 27-item, self-report questionnaire in which teachers are asked to indicate their level of agreement with statements related to four over-arching purposes for educational assessment (Brown, 2004). As detailed in Chapter 2, Brown (2004) claimed that three main purposes for assessment are well established in the literature. These are improvement of teaching and learning, student accountability and school accountability. In addition to these three purposes, Brown (2004) argued for the inclusion of a fourth conception or *antipurpose* which maintains that assessment “has no legitimate place within teaching and learning” (p.305). The

instrument's structure and items, which are discussed below, were carefully reviewed in the context of this research in order to ensure consistency with existing literature on assessment beliefs as well as their relevance within the Irish post-primary assessment context. The instrument's clear alignment with the continuum of assessment purposes, as outlined in the literature, was apparent. The intercorrelated, multi-factorial structure of the instrument was considered relevant to the Irish context given the multiple uses of assessment within the Irish education system at present (e.g. high-stakes examinations for third level entry, formative assessment, school accountability). Potential for comparison across studies as well as the possibility to position the Irish data internationally was also considered important when determining the appropriateness of using Brown's instrument in this research. Consideration is now given to the factorial composition of the instrument. This discussion is accompanied by a visual representation of the TCoA-IIIa factor structure in Table 1.

The first main factor, school accountability, posits that assessment can be used to evaluate the effectiveness of teachers and schools and consequently improve the quality of instruction. Within this conception of assessment, results are used to “publicly demonstrate that teachers or schools are doing a good job” and to impose consequences “for schools or teachers for reaching or not reaching a required standard” (Brown, 2008, p.18). In the TCoA-IIIa, this factor consists of three items (i.e. *assessment provides information on how well schools are doing, assessment is an accurate indicator of a school's quality and assessment is a good way to evaluate a school*).

The premise of the second main factor, student accountability, is that the goal of assessment is to hold students “individually accountable for their learning through giving of grades or scores, checking off performance against criteria, and reporting

grades to parents, future employers, and other educators” (Brown et al., 2011). Three items (i.e., *assessment places students into categories*, *assessment is assigning a grade to student work* and *assessment determines if students meet qualification standards*) combine to form this factor.

The third factor, improvement, refers to assessment as a means of improving student learning and the quality of teaching. Brown (2008) draws attention to two caveats in relation to this conception. Firstly, “assessment must describe or diagnose the nature of student performance” (Brown, 2008, p.15) and secondly, “the information must be a valid, reliable, and accurate description of student performance” (Brown, 2008, p.15). The improvement conception is operationalised by 4 first-order factors (i.e. *assessment describes abilities*, *assessment improves learning*, *assessment improves teaching* and *assessment is valid*) each measured by three items (e.g. assessment provides feedback to student about their performance). The first-order factors load onto the overarching improvement factor.

The fourth and final factor in Brown’s model, the irrelevance factor, reflects the view that evaluation processes are inadequate, unfair, inaccurate, and/or irrelevant to the teachers’ ability to improve student learning. Brown (2008) argues that “assessment may unfairly impact on certain students, teachers may be forced to implement assessment but choose to ignore it, or assessment may be so inaccurate that it is unreliable” (p.29). As conceived, the irrelevance factor comprises of three first-order factors (i.e. *assessment is bad*, *assessment is ignored* and *assessment is inaccurate*) each measured by three items. The first-order factors load onto the overarching irrelevance factor.

The TCoA-IIIa allows for interpretation of mean scores for each of the four main factors as well as the intercorrelations between them. The capacity to examine the interconnectedness of conceptions is important given that one's conceptions exist as part of a system, are not mutually exclusive, and may be contradictory (Green, 1971; Rokeach, 1968). The TCoA-IIIa uses a positively-packed agreement rating scale which means that the scale has more positive options than negative ones. Teachers choose from two negative options (mostly and strongly disagree) and four positive options (slightly, moderately, mostly and strongly agree) (Brown, 2006). This type of scale was chosen by Brown (2008) in an attempt to maximise the amount of variation in participants' responses, thereby allowing for a more precise description of their conceptions. Brown (2008) noted that balanced response anchors often elicit inadequate information and restrict variance when respondents are inclined to be positive toward the psychological object being rated. Given that research on conceptions indicates that teachers tend to agree with multiple conceptions, Brown (2008) maintained that a balanced response format would be too restrictive and he therefore opted for "four shades of positive orientation" (p.64).

In addition to completing the TCoA-IIIa inventory, the teachers in this study were asked to provide some demographic information in order to aid data interpretation and thus gain a more nuanced insight into the sample of teachers in question. Details were requested from participants in relation to gender, years teaching experience, current role in education, school type, subjects taught and teaching qualifications. Participants were also given the option to provide an email address if they were interested in being contacted to take part in a possible follow-up focus group on this topic in a future study.

Table 1

Brown's (2006) Teachers' Conceptions of Assessment Inventory (Abridged): Factor and Corresponding statements

Factors and Statements
<p>Assessment Makes Schools Accountable</p> <ul style="list-style-type: none"> • Assessment provides information on how well schools are doing • Assessment is an accurate indicator of a school's quality • Assessment is a good way to evaluate a school <p>Assessment Makes Students Accountable</p> <ul style="list-style-type: none"> • Assessment places students into categories • Assessment is assigning a grade or level to student work • Assessment determines if students meet qualifications standards <p>Assessment Improves Education</p> <p><i>Assessment Describes Abilities</i></p> <ul style="list-style-type: none"> • Assessment is a way to determine how much students have learned from teaching • Assessment establishes what students have learned • Assessment measures students' higher order thinking skills <p><i>Assessment Improves Learning</i></p> <ul style="list-style-type: none"> • Assessment provides feedback to students about their performance • Assessment feeds back to students their learning needs • Assessment helps students improve their learning <p><i>Assessment Improves Teaching</i></p> <ul style="list-style-type: none"> • Assessment is integrated with teaching practice • Assessment information modifies ongoing teaching of students • Assessment allows different students to get different instruction <p><i>Assessment is Valid</i></p> <ul style="list-style-type: none"> • Assessment results are trustworthy • Assessment results are consistent • Assessment results can be depended upon <p>Assessment is Irrelevant</p> <p><i>Assessment is Bad</i></p> <ul style="list-style-type: none"> • Assessment forces teachers to teach in a way against their beliefs • Assessment is unfair to students • Assessment interferes with teaching <p><i>Assessment is ignored</i></p> <ul style="list-style-type: none"> • Teachers conduct assessments but make little use of the results • Assessment results are filed and ignored • Assessment has little impact on teaching <p><i>Assessment is Inaccurate</i></p> <ul style="list-style-type: none"> • Assessment results should be treated cautiously because of measurement error • Teachers should take into account the error and imprecision in all assessment • Assessment is an imprecise process

Note. The factors in bold are intercorrelated. The first-order factors in italics load onto the factor in bold above them.

Survey Approach

One of the key considerations of survey research is how best to collect meaningful data inexpensively. With the ongoing technological developments of the 21st century, researchers now have a range of options to choose from for data collection purposes. However, despite the element of choice which now exists in relation to data collection, each method has its own strengths and difficulties and must therefore be analysed carefully in advance. Dillman (2007) speaks of a tailored design approach to survey research which he defines as:

the development of survey procedures that create respondent trust and perceptions of increased rewards and reduced costs for being a respondent, which take into account features of the survey situation and have as their goal the overall reduction of survey error (p.4).

This definition is underpinned by social exchange theory which maintains that one's actions are motivated by the return these actions are likely to bring and the extent to which the reward gained outweighs the effort required.

In applying a tailored design approach to this study, a decision was made to opt for an internet-based survey instead of a postal one. Given that the target population for this study was large in nature, the cost and practicalities of postage proved prohibitive in this case. In addition, it was noted in light of the researcher's professional experience in the post-primary field that post-primary teachers are already overloaded with paperwork in their daily jobs and would therefore be unlikely to respond in a timely fashion to a postal questionnaire. An internet-based survey allowed the researcher to access participants through a number of channels, to gain an immediate response, and to have results automatically collated. It also benefited the respondent in that s/he could respond from any location at a time that suited them and

their response was immediately registered without any further action (such as postage) required. Opting for an internet-based survey also allowed for reminder communication to be sent to the target population in a quick and easy manner.

Despite the many advantages attached to internet-based surveys, one must still be mindful of the challenges they can pose. Cohen et al. (2011) note that the “visual aspect of questionnaires is heightened in internet based surveys” (p.284). While the range of colours, graphics, fonts, lists and drop-down menus are attractive, they can also be distracting. Further, a range of computer competency levels are likely to exist within any given sample of respondents and so the questionnaire must be coherent and easy to follow. Dillman (2007) states that two aspects of a questionnaire which must be “developed and placed in concert with one another” (p.24) are information organisation and navigational guides. Moreover, when completing online questionnaires, respondents may not know how long the questionnaire is and may therefore lose interest and abandon it. Inclusion of some type of progress bar provides the respondent with important progression information which may encourage them to complete the questionnaire in full (Cohen et al., 2011). Another motivational factor to take note of is the order in which items appear in the questionnaire. It is suggested that the questionnaire should begin with simple questions which maintain the respondent’s attention (Cohen et al., 2011; Dillman, 2007).

The web-based survey template www.esurveycreator.com was used for this study. This particular survey template was chosen for its presentational clarity, ease of use, ability to export data to Microsoft excel and Statistical Package for the Social Sciences (SPSS) and its compatibility with all modern browsers. Considerable attention and time was given to ensuring that the layout of the questionnaire was clear, user-friendly and attractive. Questions were presented in a simple fashion against a

white background so as to ease reading (see Appendix B for an example page from the online survey instrument). Clear instructions were provided to participants at the beginning of the questionnaire and again at the top of each new screen page. A progress bar was included at the top of the screen so that participants would know how much of the questionnaire they had completed. Easy to answer demographic questions were positioned towards the beginning of the questionnaire in order to maintain respondent enthusiasm and reduce potential drop-out. A plain language statement as well as an informed consent form were included for completion. These are discussed further in the section on ethical considerations.

Sampling Design

In order to obtain a sample of the target population, two forms of volunteer sampling were conducted. In this type of non-probability sampling, participants opt to take part in the research study. Given the element of choice associated with this sampling strategy, one has to be careful in making any claims of generalisability as participants may have a variety of motives for volunteering. Notwithstanding this caveat, volunteer sampling was considered to be a more feasible approach than random sampling in this case. As outlined in Chapter 1, the study was conducted during a time of industrial relations unrest among teaching unions in relation to assessment reform, and so the possibility of obtaining a sampling frame from such sources was not considered to be a viable option. A list of contact details for all post-primary schools in the Republic of Ireland was obtained through the Department of Education and Skills (DES) website. Every school on this list was contacted by email. In the correspondence, the school authorities were provided with a link to the online questionnaire being used in the study and a request was made for the questionnaire link to be forwarded to all staff members. Although every school was contacted directly, it

is important to acknowledge that it is unlikely that every teacher in each school had equal access to the questionnaire. Gatekeeping issues were a key concern in this project. Access to teachers through school contact was reliant on the co-operation of management and administrative staff. Cohen et al. (2011) refer to this type of difficulty in their discussion on sampling in internet-based surveys. They maintain that the sampling frame can be unclear as it is “difficult to know how many or what kind of people saw a particular survey on a website” (Cohen et al., 2011, p.286). In an effort to maximise equality of access to the questionnaire and reach as high a proportion of the post-primary teaching population as possible, a second volunteer sampling technique was employed.

Specifically, contact was made via email and Twitter with a range of national educational bodies such as the Teaching Council, teaching unions, subject associations, education centres and managerial bodies. These particular organisations were selected due to their direct association with post-primary teachers. Reips (2002) contends that this type of multiple site entry technique is a useful means of increasing response rates on an internet-based survey as it widens the number of channels through which participants can gain access to a questionnaire. When contacted, the organisations were asked to disseminate the questionnaire to their members by whatever means possible and to encourage their members to support the research. A list of the bodies that co-operated with the research, and by what means, is outlined in Table 2.

Table 2

Organisations who helped with the recruitment of participants

Organisation Name	Method of Communication
Association of Secondary Teachers of Ireland (ASTI)	Advertisement published in the ASTIR magazine
Teachers' Union of Ireland (TUI)	Advertisement published in the TUI News magazine
Irish Maths Teachers' Association	Email with survey link sent to all 1310 members
German Teachers' Association of Ireland	Link to survey posted via Twitter and Facebook
Spanish Teachers' Association of Ireland	Link to survey posted via Facebook
Blackrock Education Centre	Link to survey posted on their website
Joint Managerial Body (JMB)	Link to survey posted on their website
Association of Teachers'/Education Centres in Ireland (ATECI)	Survey link sent to member centres
Physical Education Association of Ireland	Link to survey posted via Twitter
Religion Teachers' Association of Ireland	Link to survey posted via Twitter
An Foras Pátrúnachta	Link to survey posted via Twitter
Gael Linn	Link to survey posted via Twitter
Educate Together	Link to survey posted via Twitter
French Teachers' Association of Ireland	Link to survey posted via Twitter
Development and Intercultural Education (DICE) Froebel	Link to survey posted via Twitter
National Association of Principals and Deputy Principals (NAPD)	Link to survey posted via Twitter
National Induction Programme for Teachers (NIPT)	Link to survey posted via Twitter
Junior Cycle for Teachers (JCT)	Link to survey posted via Twitter

Piloting

Piloting is a useful and important way of checking how valid, reliable and user-friendly a questionnaire is. The questionnaire for this study was piloted with 15 teachers. The pilot study was conducted in three iterations with 5 participants taking part in each phase. Convenience samples from the researcher's own second-level school (south county Dublin) were used for the first and third iterations and a convenience sample from another second-level school (north county Dublin) was used

for the second iteration. In the pilot study, the researcher was particularly interested in receiving feedback on the clarity of questionnaire items, instructions and layout and on the duration of time it took to complete the questionnaire. A feedback sheet was included for completion with the questionnaire (See Appendix C). The questionnaire was revised based on the feedback received during each iteration. While no major difficulties arose and the questionnaire was deemed user-friendly by most teachers, some confusion between beliefs and practices was identified. Two of the teachers were uncertain as to whether they should be responding to the questionnaire based on their assessment beliefs or their assessment practices. The following instructions were provided at the top of each screen on foot of this feedback:

Please respond to each statement by choosing the response that best reflects your beliefs about assessment, whatever that term means to you. Please note: you are asked to focus on your assessment beliefs not your assessment practices. Further, you need to provide a response to each statement in order to progress through the survey.

Ethics

Permission was sought from Professor Gavin Brown for the use of the TCoA-III A in this study. Professor Brown granted permission to proceed with the TCoA-III A subject to him being notified of any adaptations made and being sent a copy of the research results once the thesis is approved for release (see personal correspondence, Appendix D). A plain language statement and informed consent form were included at the beginning of the questionnaire (see Appendices E & F). The plain language statement outlined the rationale for the research and the benefits to the respondent of taking part in the research. The academic background and institutional affiliation of the researcher were clearly outlined. Contact details were provided for both the researcher and the university administration in the event of questions or concerns from

participants. The issues of anonymity and confidentiality were addressed in the plain language statement and again before the declaration of consent. It was made clear to participants that the data gathered would only be used for the purposes of the research and possible future conferences and publications.

Response Data

Response Rate

A total of 586 post-primary teachers participated in the survey. The sample was subsequently reduced to 489 in order to meet the criteria of full data on all survey items. This meant that only questionnaires with all 27 statements completed were analysed. Analysis of questionnaires not fully completed indicated no clear pattern in the point of dropout with the majority of dropouts occurring at the declaration of consent and demographic information stages. A more detailed breakdown of this information is presented in Table 3. The sample size obtained was considered satisfactory as a probability sample of the same size would result in a margin of error (at a 95% confidence level) of 4.39 for a population size of 26,804 and 4.41 for a population size of 42,542.

Table 3

The point of drop out in partially completed questionnaires

Stage of drop out	Participants (n=97)
Confirmation of Understanding / Declaration of Consent	25
Demographic Information	32
Statements 1-9	17
Statements 9-18	17
Statements 18-27	6

Participant Demographic Information

Basic descriptive statistics (frequencies) were obtained for all the categorical variables on the questionnaire. The gender breakdown in the final working sample of 489 participants was 68.5% female and 31.5% male. These figures closely reflect the most recently available statistics for gender breakdown in the post-primary teaching population where females accounted for 68.7% of teachers and males accounted for 31.3% of the population (See Appendix A for correspondence from the Teaching Council).

Current role in education.

Table 4

Participants' Current Role in Education

Current role in education	Participants (n=489)	% of total sample
Principal	48	9.8
Deputy Principal	21	4.3
Classroom Teacher	347	71.0
Learning Support Teacher	19	3.9
Trainee Teacher	14	2.9
Guidance Counsellor	8	1.6
Retired/Career Break	14	2.9
Other	18	3.7

In terms of the figures for educational roles, three findings are noteworthy.

Principal and Deputy Principals are over-represented in this sample whereas classroom

teachers appear to be under-represented. Based on the DES (2016a) population size of 26,804 teachers and the DES (2016b) post-primary teacher allocation information, principals account for 2.7% of the post-primary teaching population and deputy principals account for 2.8% of this population. The over-representation of these two groups could be related to the co-operation of the National Association of Principals and Deputy Principals with the dissemination of the questionnaire. While an official number of classroom teachers could not be obtained, it would seem reasonable to suggest that if management roles constitute 5.5% of the population, then classroom teachers would most likely represent around 85%-90% of the population allowing for other roles such as learning support teachers.

Years of teaching experience.

Table 5

Years of Teaching Experience

Years Teaching Experience	Participants (n=489)	% of total sample
0-5	95	19.4
6-10	98	20.0
11-20	132	27.0
20+	164	33.5

In terms of the age profile of participants in the sample, while one cannot assume that those with less experience will necessarily be younger, the overall profile appears to be quite young as 66.4% of the sample reported having less than 20 years of teaching experience. As such the age profile of the respondents in this study appears to be younger than average when compared with a recent statistic from *Education at a Glance 2016* (OECD) which reports that 44% of secondary teachers in Ireland are under the age of 40.

Current school type.

Table 6

Participants' Current School Type

Current School Type	Participants (n=489)	% of total sample
Voluntary	190	38.9
Community / Comprehensive	116	23.7
Vocational (ETB)	119	24.3
Private (Fee-paying)	48	9.8
Other	16	3.3

The participants in this study represented a broad range of school types. For the purpose of comparison with national statistics, the voluntary and private schools will be amalgamated under one category as is done by the DES in relation to teacher statistics. This is due to the similar management structure in these schools as well as the fact that a lot of teachers in private schools are fully or partially paid by the DES. The most recently available national statistics for the number of teachers per school type are from 2012. These statistics are still a useful base for comparison, however, as the number of teachers on the DES payroll that year was marginally less than in the population for this study. Overall, the representation from voluntary and private schools (48.7%) aligns well with the national breakdown (50%). Teachers from vocational schools, however, appear to be under-represented in this study as they accounted for 33.6% of teachers in the 2012 breakdown. Furthermore, teachers from Community/Comprehensive schools appear to be over-represented in the sample as they accounted for 16.4% of the 2012 national breakdown.

Qualifications.

Table 7

Qualifications held by participants

Qualifications	Participants (n=489)	% of total sample
Degree	305	62.4
Higher Diploma/Postgraduate Diploma in Education	381	77.9
Master's Degree	190	38.9
Doctoral Degree	11	2.2
Other Postgraduate Qualification	52	10.6

In terms of qualifications, the teachers in this sample seem to be particularly well qualified. Recent statistics obtained from the Teaching Council (See Appendix A for correspondence from the Teaching Council) state that approximately 2562 teachers are registered as holding a master's degree or higher. This figure represents 6% of post-primary teachers registered with the Council at the time of correspondence. While the Council pointed out that some teachers choose not to record additional qualifications on their registration profile, the percentage of teachers in this study holding a master's degree or higher (41.1%) would appear to be particularly high.

A broad range of teaching subject domains are represented in this study (See Appendix G). Most highly represented are Mathematics (24.5%), English (21.7%), Irish (19.6%), CSPE (19.6%), History (15.5%), Business (13.3%), Geography (13.1%), Science (13.1 %) and French (11.7%). The over-representation of Mathematics, Irish, French and Religion teachers could be related to the co-operation of associated subject associations in the data collection process. The over-representation of teachers teaching Civil Social and Political Education (CSPE) and Social, Personal and Health Education (SPHE), on the other hand, is likely related to the teaching of these subjects

as supplementary subjects without Teaching Council registration (based on the researcher's experience of the post-primary sector).

Data Cleaning / Preliminary Analysis

In preparation for data analysis, descriptive statistics were conducted for all the continuous variables in order to check for errors and out of range values. A breakdown of the mean and standard deviation for each of the 27 items on the questionnaire is provided in Appendix H. Total scale scores were obtained for each of the four subscales in the TCoA-III A. The internal consistency (Cronbach alpha) of each subscale within the TCoA-III A was checked. The improvement subscale showed high reliability (.83). The irrelevance (.76) and school accountability subscales (.75) showed good reliability but the student accountability subscale had very poor internal consistency (.36) with this particular sample.

Percentage frequencies for each of the 27 statements are presented in Table 8. The statements are sorted by the scores for "strongly agree" and are colour coded according to Brown's original factor structure – improvement (green), school accountability (purple), student accountability (blue) and irrelevance (orange). The presentation of percentage frequencies in Table 8 is followed by a brief commentary on the general picture of agreement in the data in relation to Brown's (2006) original factor structure.

Table 8

Percentage frequencies for each statement on the TCoA-III A
(Sorted by Mostly/Strongly Agree)

	Strongly/ Mostly Disagree	Slightly/ Moderately Agree	Mostly/ Strongly Agree	Total Agree
Assessment provides feedback to students about their performance	10	26	64	90
Assessment is integrated with teaching practice	11	27	62	89
Assessment is a way to determine how much students have learned from teaching	11	29	60	89
Assessment places students into categories	13	36	50	86
Assessment information modifies ongoing teaching of students	6	45	49	94
Assessment establishes what students have learned	7	49	44	93
Assessment helps students improve their learning	9	48	44	92
Assessment feeds back to students their learning needs	14	43	43	86
Teachers should take into account the error and imprecision in all assessment	13	49	37	86
Assessment determines if students meet qualifications standards	17	46	37	83
Assessment results are trustworthy	15	53	31	84
Assessment results can be depended on	16	56	28	84
Assessment provides information on how well schools are doing	20	55	25	80
Assessment allows different students to get different instruction	32	42	25	67
Assessment results should be treated cautiously because of measurement error	19	56	24	80
Assessment measures students' higher order thinking skills	27	51	22	73
Assessment is assigning a grade or level to student work	40	39	22	61
Assessment results are consistent	28	53	19	72
Assessment forces teachers to teach in a way against their beliefs	42	40	19	59
Assessment is an imprecise process	39	47	13	60
Teachers conduct assessments but make little use of the results	48	39	13	52
Assessment interferes with teaching	60	31	10	41
Assessment results are filed & ignored	57	34	9	43
Assessment is a good way to evaluate a school	58	34	8	42
Assessment is an accurate indicator of a school's quality	59	34	7	41
Assessment has little impact on teaching	77	18	5	23
Assessment is unfair to students	71	27	1	28

Looking firstly at the general picture of agreement in the data, the percentage frequencies for total agreement indicate high amounts of agreement overall with the statements that relate to Brown's (2006) factor of assessment for improvement (green). It is noteworthy that teachers agreed strongly with the ideas of assessment as integrated with teaching practice (62%), as a means of determining how much students have learned from teaching (60%) and as a means of providing feedback to students about their performance (64%). There are less noteworthy but nonetheless high percentages of agreement with the idea of assessment for student accountability (blue) and lower overall percentages of agreement with Brown's notion of assessment as irrelevant (orange) and assessment for school accountability (purple). Although these broad trends provide an overall sense of the data, some clear exceptions to the trends are apparent and must be noted.

While the majority of statements which participants agreed with (mostly/strongly) lay within Brown's assessment for improvement factor, it is worth noting that 50% of the sample also agreed (mostly/strongly) with the statement that assessment places students into categories. It would appear, therefore, that participants at once agreed with the idea of assessment as a formative construct closely aligned with teaching and learning whilst also agreeing (86% in total) that assessment separates students in terms of their ability. It could be argued that this categorisation might be by some means other than grades (e.g. criteria/descriptors) as participants did not endorse to the same degree (61% in total) the idea of assessment as the assignment of a grade or level to student work.

In relation to Brown's (2006) irrelevance factor (orange), there are noticeably higher levels of agreement (86% and 80% respectively) with the two statements urging caution in relation to measurement error than with any of the other statements in this

overarching factor. Teachers appear to believe in large numbers that the assessment process is one which can be subject to error or imprecision. Notwithstanding this potential for measurement imprecision, however, the majority of participants did not agree that assessment interferes with teaching (60%), is ignored (57%), is unfair to students (71%) or has little impact on teaching (77%),

In terms of Brown's (2006) school accountability factor, the participants in this study appear to have clearly distinguished between the ideas of assessment as providing information about a school and assessment as being an effective manner of evaluating a school. A total of 80% of the participants agreed that assessment provides information on how well schools are doing but over half of sample strongly/mostly disagreed that assessment is a good way to evaluate a school (58%) or that assessment is an accurate indicator of a school's quality (59%).

Summary

This study adopted a non-experimental cross-sectional design. Participants completed an online version of the TCoA-IIIa (Brown, 2006). A sample size of 489 participants was obtained with a gender breakdown that closely reflects that of the population. Overall, the participants in this study appear to be younger and more qualified than average. Those occupying management positions are over-represented in the sample whereas classroom teachers appear to be under-represented. The respondents came from a broad range of subject disciplines, with English, Irish and Mathematics most highly represented. A range of school types are reflected in the sample. The representation from the voluntary/private sector aligns with population norms but the vocational sector is underrepresented and the community/comprehensive

sector is over-represented. Some basic descriptive statistics on the questionnaire items suggest that participants tended to agree for the most part with the conceptions of assessment for improvement and student accountability but agreed less with the school accountability and irrelevance conceptions. Further data analysis incorporating factor analysis, correlations, t-tests and one-way analysis of variance is reported on in full in Chapter 4.

Chapter 4: Data Analysis

Introduction

This chapter presents an analysis of the findings from the quantitative data collected. The main analytical technique employed in this study was factor analysis. The suitability of the data for factor analysis is addressed early on in the chapter and this is then followed by an in-depth discussion of the factor structure extracted from the data. Additional analyses involving correlations, t-tests and one-way analysis of variance are addressed in the second half of the chapter.

Factor Analysis

Suitability of the Data for Factor Analysis

The 27 items of the TCoA-III A were subjected to exploratory factor analysis using SPSS version 23.0. Prior to performing the analysis, the suitability of the data for factor analysis was assessed. Notwithstanding the absence of any real consensus in relation to sample size for factor analysis, it is acknowledged that the sample must be “large enough that correlations are reliably estimated” (Tabachnick & Fidell, 2007, p.613). Field (2009) refers to a minimum amount of “10-15 participants per variable” (p.647) while Tabachnick and Fidell (2007) suggest that “it is comforting to have at least 300 cases for factor analysis” (p.613). In line with these general recommendations in the literature, the sample size of 489 in this study was deemed sufficient. Secondly, inspection of the correlation matrix revealed the presence of many coefficients of 0.3 and above, thereby fulfilling another key recommendation for factor analysis suitability. Furthermore, the Kaiser-Meyer-Olkin measure of sampling adequacy was

0.825, exceeding the recommended value of 0.6 (Kaiser, 1974) and Bartlett's test of Sphericity (Bartlett, 1954) reached statistical significance, supporting the factorability of the data.

Using maximum likelihood extraction and direct oblimin rotation as recommended by Costello and Osborne (2005) and Fabrigar and Wegener (2012), exploratory factor analysis (EFA) was conducted to assess the underlying structure of the 27 items on the TCoA-III. The initial analysis extracted 8 factors with eigenvalues greater than 1. Inspection of the scree plot, however, showed inflexions that would justify retaining either 3, 4 or 5 factors. Three further factor analyses were then conducted forcing 3, 4 and 5 factors respectively. The pattern matrices from each of these outputs were compared in order to find the solution of best fit (Costello & Osborne, 2005). A solution of best fit is a conceptually sound factor structure, with strong factor loadings and few item crossloadings (Costello & Osborne, 2005; Fabrigar & Wegener, 2012). Comparison of the pattern matrices suggested that the 5 factor structure was the cleanest and most conceptually sensible representation of the data. This result was supported by a separate parallel analysis which was carried out in conjunction with the main factor analysis. The parallel analysis involved conducting a Principal Component Analysis (PCA) and comparing the resulting eigenvalues with random eigenvalues generated using Watkins (2006) Monte Carlo PCA for parallel analysis. As explained by Pallant (2013), if the eigenvalue from the PCA is larger than the criterion value from parallel analysis, then this factor is retained. If the eigenvalue is smaller than the criterion value, then the factor is rejected. The results of the parallel analysis, displayed in Table 9, supported the decision to retain five factors.

Table 9

Comparison of eigenvalues from PCA and criterion values from parallel analysis

<i>Component Number</i>	<i>Actual eigenvalue from PCA</i>	<i>Criterion value from parallel analysis</i>	<i>Decision</i>
1	5.777	1.4541	Accept
2	2.869	1.3877	Accept
2	2.013	1.3414	Accept
4	1.623	1.2953	Accept
5	1.338	1.2542	Accept
6	1.164	1.2212	Reject

Five-Factor Model

As illustrated in Table 10, the five-factor EFA result accounted for 40% of the cumulative variance before rotation. After rotation, the cumulative variance accounted for remained the same but this variance was redistributed across the factors (Cohen et al., 2012; Fabrigar & Wegener, 2012). SPSS does not report this redistributed percentage variance for EFA. It does, however, provide some data on the rotation sums of square loadings and these figures give a sense of the explanatory power of the factors after rotation. The five factors extracted from the data differed from Brown's (2006) original 4-factor model except for the school accountability factor which was fully recovered. As presented in Table 11, the five factors extracted from the data were named as follows: (1) *Assessment is a diagnostic and formative tool*; (2) *Assessment is irrelevant (bad, ignored and inaccurate)*; (3) *Assessment is a school accountability tool*; (4) *Assessment is a measurement and categorisation tool*; (5) *Assessment is a valid grading tool*. In the case of factors three and four, the substantive loadings were negative and were reversed for ease of interpretation. This is an accepted practice within exploratory factor analysis. As explained by Thompson (2004), while referring to (Gorsuch, 1983, p.181), "because most people find it easier to think in positive terms, if the larger pattern and structure coefficients have negative signs on a given

factor, it is completely appropriate for the analyst to "reflect" the factor by reversing the signs of the coefficients on any given factor" (p.96). Fabrigar and Wegener (2012) also refer to this practice by explaining that "the scaling direction of common factors is arbitrary. Programs simply scale solutions on the basis of computational convenience rather than as a function of some fundamental conceptual property of the common factors. Thus, for any factor solution, it is permissible to reverse the signs of all factor loadings in a given column of the factor loading matrix. Such reversals (as long as they are applied to all elements in the column) do not affect the fit of the model or the communalities of the measured variables" (p.79).

In advance of discussing each factor individually, due regard must be given to the communality values provided for each variable (see table 12). The communality values after extraction represent the "variance in items accounted for by the extracted factors" (Fabrigar & Wegener, 2012, p.131). These values are important in that they provide information on the degree of association between variables whilst also alerting the researcher to variables which may prove problematic in the interpretation of factors. Pallant (2013) notes that "low values (e.g. less than .3) could indicate that the item does not fit well with the other items in its component" (p.206). Five of the twenty seven items in this study had communality values of less than 0.3. The items in question are Assessment has little impact on teaching (.135), Assessment is assigning a grade or level to student work (.169), Assessment places students into categories (.218), Teachers should take into account the error and imprecision in all assessment (.222), and Assessment determines if students meet qualifications standards (.244). These values will be considered as part of the discussion on each individual factor.

Factor 1: Assessment is a diagnostic and formative tool.

The first factor contains seven items with a minimum factor loading of 0.394.

All of the items which make up this factor were part of Brown's (2006) original *Improvement* factor. The relevant items and their factor loadings are as follows:

- Assessment feeds back to students their learning needs (0.746)
- Assessment information modifies ongoing teaching of students (0.664)
- Assessment helps students improve their learning (0.606)
- Assessment allows different students to get different instruction (0.586)
- Assessment measures students' higher order thinking skills (0.466)
- Assessment establishes what students have learned (0.433)
- Assessment results can be depended upon (0.394)

The items which load on this factor represent a strongly formative approach to assessment. Words such as “learning needs”, “modifies” and “improve” all point to the use of assessment for pedagogical purposes. Central to the strongest loading item is the concept of feedback which is a key aspect of formative assessment and which Hattie and Timperley (2007) describe as one of the most powerful influences on learning and achievement. In addition to this, as noted in the introduction, the adaptive expertise of the teacher is paramount to the facilitation of meaningful learning through formative assessment. This point is illustrated by two items which point to the ongoing modification of teaching as well as the differentiation of instruction. Another key tenet of formative assessment which is present in this factor is the use of assessment to activate students' higher order thinking skills, thereby helping them to become critical thinkers.

Table 10

Total Variance Explained for Exploratory Factor Analysis with Oblique Rotation

							Rotation
							Sums of
							Squared
Initial Eigenvalues				Extraction Sums of Squared			Loadings
				Loadings			
		% of	Cumulative		% of	Cumulative	
Factor	Total	Variance	%	Total	Variance	%	Total
1	5.777	21.398	21.398	5.129	18.995	18.995	3.906
2	2.869	10.625	32.023	2.288	8.476	27.471	3.239
3	2.013	7.455	39.479	1.472	5.452	32.923	2.623
4	1.623	6.011	45.489	1.114	4.124	37.047	2.725
5	1.338	4.957	50.446	.799	2.959	40.006	1.414
6	1.164	4.310	54.755				
7	1.086	4.023	58.779				
8	1.047	3.878	62.657				
9	.974	3.609	66.266				
10	.837	3.100	69.365				
11	.758	2.807	72.172				
12	.737	2.729	74.901				
13	.660	2.445	77.345				
14	.637	2.361	79.706				
15	.615	2.279	81.986				
16	.539	1.995	83.980				
17	.505	1.870	85.850				
18	.495	1.833	87.683				
19	.459	1.701	89.384				
20	.444	1.643	91.027				
21	.402	1.491	92.518				
22	.389	1.439	93.957				
23	.382	1.414	95.370				
24	.345	1.277	96.647				
25	.330	1.221	97.868				
26	.292	1.081	98.949				
27	.284	1.051	100.000				

Table 11

Factor Loadings for Exploratory Factor Analysis with Oblique Rotation of TCoAI- IIIA

Items	Factor				
	1	2	3	4	5
Assessment feeds back to students their learning needs	.746	.085	.061	.054	.000
Assessment information modifies ongoing teaching of students	.664	.077	.075	.150	-.043
Assessment helps students improve their learning	.606	-.095	.050	.060	-.029
Assessment allows different students to get different instruction	.586	-.074	.052	-.086	-.007
Assessment measures students' higher order thinking skills	.466	-.100	.107	-.008	.147
Assessment establishes what students have learned	.433	-.016	.160	.114	.068
Assessment determines if students meet qualifications standards	.260	-.029	.185	.075	.232
Assessment results should be treated cautiously because of measurement error	.093	.609	-.050	.059	-.171
Assessment forces teachers to teach against their beliefs	-.051	.594	-.067	-.026	.119
Teachers conduct assessments but make little use of the results	.018	.554	.113	-.075	.033
Assessment results are filed and ignored	-.053	.544	.037	-.074	.021
Assessment is unfair to students	-.170	.519	-.012	-.032	.093
Assessment interferes with teaching	-.104	.513	-.105	-.049	.275
Assessment is an imprecise process	-.114	.469	-.117	.085	-.068
Teachers should take into account the error and imprecision in all assessment	.167	.423	.022	.099	-.208
Assessment has little impact on teaching	-.143	.213	-.006	-.116	.177
Assessment is a good way to evaluate a school	.126	.039	.869	-.143	-.041
Assessment is an accurate indicator of a school's quality	-.003	-.019	.785	-.018	-.022
Assessment provides information on how well schools are doing	-.126	-.046	.503	.249	.049
Assessment provides feedback to students about their performance	.034	-.052	-.031	.827	-.011
Assessment is a way to determine how much students have learned from teaching	.079	-.052	.040	.694	.069
Assessment is integrated with teaching practice	.126	-.017	-.050	.650	-.033
Assessment places students into categories	-.152	.198	.161	.332	.076
Assessment results are consistent	.315	-.045	.064	.004	.473
Assessment results can be depended upon	.394	-.283	.043	-.029	.451
Assessment results are trustworthy	.097	-.193	.065	.368	.435
Assessment is assigning a grade or level to student work	-.040	.097	.082	.078	.355

Note. Factor loadings greater than .30 are shown in boldface.

Table 12

Communality Values for the TCoA-III variables following Exploratory Factor Analysis with Oblique Rotation

Items	Communalities	
	Initial	Extraction
Assessment provides information on how well schools are doing	.352	.377
Assessment places students into categories	.220	.218
Assessment is a way to determine how much students have learned from teaching	.517	.561
Assessment provides feedback to students about their performance	.541	.692
Assessment is integrated with teaching practice	.444	.469
Assessment results are trustworthy	.426	.485
Assessment forces teachers to teach against their beliefs	.411	.392
Teachers conduct assessments but make little use of the results	.407	.304
Assessment results should be treated cautiously because of measurement error	.382	.394
Assessment is an accurate indicator of a school's quality	.497	.600
Assessment is assigning a grade or level to student work	.235	.169
Assessment establishes what students have learned	.400	.319
Assessment feeds back to students their learning needs	.449	.526
Assessment information modifies ongoing teaching of students	.447	.470
Assessment results are consistent	.372	.379
Assessment is unfair to students	.367	.371
Assessment results are filed and ignored	.422	.323
Teachers should take into account the error and imprecision in all assessment	.281	.222
Assessment is a good way to evaluate a school	.538	.731
Assessment determines if students meet qualifications standards	.268	.244
Assessment measures students' higher order thinking skills	.394	.332
Assessment helps students improve their learning	.481	.462
Assessment allows different students to get different instruction	.412	.372
Assessment results can be depended upon	.487	.562
Assessment interferes with teaching	.364	.383
Assessment has little impact on teaching	.231	.135
Assessment is an imprecise process	.313	.310

Note. Communality values after extraction of less than .3 are shown in boldface.

While the factor as a whole has moderate to strong loading items which are easy to interpret, the lowest loading item, *Assessment results can be depended upon* (0.394), is not as clearly related to the factor as the other items. Justification for its inclusion will now be provided. As can be seen from the pattern matrix (Table 11), *Assessment results can be depended upon* loaded on two factors, the first factor and the fifth factor. An item is seen as cross-loading if it loads at 0.32 or higher on more than two factors (Costello & Osborne, 2005). While one might ideally like to achieve a solution where all variables load strongly on one factor only, the existence of double-loading items is not necessarily problematic. Fabrigar and Wegener (2012) note that cross-loadings “can often be quite interpretable if the measured variable can be plausibly interpreted to be influenced by more than one construct” (p.66). They also point out that “an interpretation that can account for single-loading and double-loading items would be more compelling than an interpretation that can only account for single-loading items” (pp.138-199). In this case, *Assessment results can be depended upon* loads at 0.451 on the fifth factor and at 0.394 on the first factor which is currently under discussion. It was noted in the literature review that teachers can often hold multiple and sometimes conflicting conceptions of assessment. It seems quite plausible, therefore, in the context of this post-primary research, that teachers could at once consider formative assessment and the grading of students to be dependable forms of assessment. Exclusion of this variable from the first factor would be to ignore the fact that teachers attached a sense of validity to formative assessment. While one could argue that a loading of 0.394 is low, it is still, according to Stevens’s (2002) table of critical values, a statistically significant loading with a sample size of 489. In addition to this, 0.394 lies within a reasonably close numerical distance from the next lowest loading item in the factor which loads at 0.433. Bearing all these points in mind, the variable in

question was considered meaningful and was therefore included in the first factor. It should be noted that there was one variable (*Assessment determines if students meet qualifications standards*) which loaded at .260 on factor one but which was not included in the naming of the factor. While a loading of .260 is just about statistically significant in this case, the substantive importance of the variable to the factor was quite low. An estimate of such importance is found by squaring the value of the loading (Field, 2009; Stevens, 2002). In this instance, the item in question shared less than 10% variance with the factor. This is not surprising given that the variable had a low communality value and does not at all appear conceptually aligned with the higher loading items.

Factor 2: Assessment is irrelevant (bad, ignored & inaccurate).

The second factor contains eight items with a minimum factor loading of 0.423. All but one of the variables in this factor correspond exactly to Brown's second-order *Irrelevance* factor which was made up of three first-order factors (*Assessment is bad*, *Assessment is ignored* and *Assessment is inaccurate*). The one item which did not load adequately enough to be included in the factor was *Assessment has little impact on teaching*. The relevant items and their factor loadings are as follows:

- Assessment results should be treated cautiously because of measurement error (0.609)
- Assessment forces teachers to teach against their beliefs (0.594)
- Teachers conduct assessments but make little use of the results (0.554)
- Assessment results are filed and ignored (0.544)
- Assessment is unfair to students (0.519)
- Assessment interferes with teaching (0.513)

- Assessment is an imprecise process (0.469)
- Teachers should take into account the error and imprecision in all assessment (0.423)

Firstly, the one item which was not included in this factor (*Assessment has little impact on teaching*) is worthy of discussion. As was noted earlier in this chapter, this item had a very low communality value (0.135) which alerted the researcher to a potentially weak relationship between the item and the extracted factors. It was not surprising, therefore, that the item loaded at 0.213 on the second factor, which, according to Stevens's (2002) table of critical values, is not a statistically significant loading with a sample size of 489. It is noteworthy also that this item had the lowest mean (2.09) of all the items in this factor as well as having the lowest mean of all 27 items on the questionnaire. This finding indicated strong disagreement amongst Irish post-primary teachers with the idea of assessment has having little impact on teaching. The extent of disassociation between this item and the factor at large is particularly apparent when one compares the strength of the effect size between the mean of this item and the mean of the strongest loading item in the factor using Cohen's *d*. The strength of the effect size ($d = 1.35$) indicates a large amount of divergence in how participants responded to the two items. In light of all this information, a decision was made not to include the item in the factor.

A second variable of note in this factor is the lowest loading item to be included in the factor (*Teachers should take into account the error and imprecision in all assessment*). This too was one of the items flagged earlier through the communality values as potentially problematic. In addition, this variable was also previously

highlighted in the discussion on percentage frequencies due to it being the most strongly endorsed item in the factor. It would appear, therefore, that the comparatively lower loading for this variable was due to participants' positive interpretation of its meaning. In other words, while participants interpreted most of the statements in this factor as representing a negative conception of assessment, they believed for the most part that teachers should take stock of error and imprecision, interpreting this positively as an important part of the assessment process.

Factor 3: Assessment is a school accountability tool.

The third factor contains three items with a minimum factor loading of 0.503. The three items which make up the factor are the exact same items which combined to form this factor in Brown's (2006) original scale. The relevant items and their factor loadings are as follows:

- Assessment is a good way to evaluate a school (0.869)
- Assessment is an accurate indicator of a school's quality (0.785)
- Assessment provides information on how well schools are doing (0.503)

As was previously noted in the earlier description of Brown's (2006) TCoA-III A, this school accountability factor refers to assessment as a representation of how a school is performing overall. Within this view of assessment, assessment results are used to publicly demonstrate the quality of instruction being delivered by teachers and schools and may be used to invoke consequences for schools not reaching required standards (Brown, 2008). All of the variables in this factor loaded quite strongly. The comparatively lower loading of the third item aligns with the earlier discussion on around percentage frequencies which pointed to participants' clear distinction between

assessment as a *good/accurate* means of school evaluation and assessment as a means of school evaluation.

Factor 4: Assessment is a measurement and categorisation tool.

The fourth factor contains five items with a minimum factor loading of 0.332. The four highest loading items in this factor were part of Brown's (2006) *Improvement* factor and the lowest loading item was part of his *Student Accountability* factor. The relevant items and their factor loadings are as follows:

- Assessment provides feedback to students about their performance (0.827)
- Assessment is a way to determine how much students have learned from teaching (0.694)
- Assessment is integrated with teaching practice (0.650)
- Assessment results are trustworthy (0.368)
- Assessment places students into categories (0.332)

The items which load on this factor appear to represent a summative perspective on assessment as it relates to both students and teachers. As was discussed in the introductory chapter, summative assessment can be used for multiple purposes and this is reflected in this factor. The highest loading item in the factor refers to assessment as a means of providing feedback to students about their performance. While this item would appear to point to the use of summative assessment in a formative way, it is interesting to note the distinction made within the data between feedback about *performance* and feedback directed towards the improvement of *learning* as illustrated in the first factor. The separation of feedback purposes in this way would suggest a nuanced understanding of assessment amongst the participants. The second item in the

factor points to the use of assessment as a tool for gauging the amount or the extent of learning that has taken place after teaching. This is one of the main purposes for which teachers use summative assessment. Another purpose which is included in this factor is the use of assessment to categorise learners. This item was the lowest loading item in the factor (0.332) and was highlighted earlier as having a low communality value. A possible reason for this is that the item was originally part of Brown's *Student Accountability* factor which no longer exists within this five-factor model. It seems quite plausible in the context of this new factor, however, that assessment could separate students into different categories based on the extent of their learning. In addition to this, given the sample size in question, a loading of 0.332 is still statistically significant and so for these reasons the item was deemed worthy of inclusion. Worthy of discussion also in this factor is the integration of assessment with teaching practice. The appearance of this item within the factor would suggest that the participants viewed the type of summative assessment in this factor as existing throughout the teaching and learning process rather than simply at the end of it. This viewpoint logically aligns with the previously discussed item which referred to assessment as a mechanism for feedback about performance. Finally, the last remaining item (*Assessment results are trustworthy*), cross-loaded (statistically significantly) onto this factor from the fifth factor in a comparable fashion to the validity item which cross-loaded onto factor one. As previously noted, teachers can hold multiple conceptions of assessment at any one time and so it would seem reasonable that teachers could at once attach a sense of validity to ongoing summative assessment (factor 4), formative assessment (factor 1) and the grading of students which will be discussed in the fifth factor. The multi-faceted understanding within the sample of what constitutes *valid* assessment, is,

arguably, indicative of a complex multi-purpose assessment environment at post-primary level in the Republic of Ireland.

Factor 5: Assessment is a valid grading tool.

The fifth factor contains four items with a minimum factor loading of 0.355.

The relevant items and their factor loadings are as follows:

- Assessment results are consistent (0.473)
- Assessment results can be depended upon (0.451)
- Assessment results are trustworthy (0.435)
- Assessment is assigning a grade or level to student work (0.355)

Three of the items in this factor made up Brown's (2006) first-order *Assessment is valid* factor which was subsumed under his second-order *Improvement* factor. The fourth item in this new factor was part of Brown's (2006) *Student Accountability* factor. This item was flagged earlier as having a low communality value which is possibly due to it now belonging to a completely new factor. Given that post-primary teachers in the Republic of Ireland have been working in a high-stakes grade-orientated system for a long time, it's not surprising that the grading of students has emerged from the data as a particularly valid form of assessment. It is noteworthy, however, that grading has appeared as a discrete category delineated clearly from summative assessment for other purposes as outlined in the fourth factor. This separation would suggest that teachers in this sample saw grading as having a very particular role which is valid in its own right but which is distinct from summative assessment which gauges the extent of learning and provides feedback throughout the teaching and learning process. Similar to the distinction made earlier between the feedback types, this distinction in summative

assessment purposes once again points to quite a developed and multi-dimensional understanding of assessment within the sample.

Additional Analyses

As illustrated in Table 10, the internal consistency for the factors ranged from minimally reliable (.61) to highly reliable (.80). The rank order of endorsement of each factor is also listed in Table 10. Participants agreed most strongly with the factor describing assessment as a measurement and categorisation tool. They next strongly rated the conception of assessment for formative use, followed thirdly by assessment for the purpose of grading. The two factors which participants agreed with the least were the negative irrelevance factor and the factor describing assessment as a school accountability tool. Both of these mean scores fell between the moderately disagree and slightly agree points on the scale.

The factor correlations outlined in Table 11 point to some noteworthy relationships. The highest correlation was a medium negative correlation (-.369) between factor one and factor two. This seems like a logical result as one would expect that the more a teacher endorses the idea of formative assessment, the less likely he/she is to believe that assessment is bad, inaccurate or ignored. A medium positive correlation (.3) emerged between factor 1 (assessment as a diagnostic and formative tool) and factor 4 (assessment as a measurement and categorisation tool). This result again appears to be a rational outcome as the items in both of these factors pointed to both types of assessment as being integrated with the teaching and learning process. It seems quite likely, therefore, that the learning advice provided to a student through the formative assessment process could influence their performance in ongoing summative

assessment tasks which in turn could lead to their categorisation. A third medium positive correlation (.318) occurred between assessment as a grading tool and assessment for school accountability purposes.

Table 13

Mean Factor Endorsement Rankings and Cronbach Alpha Reliabilities

<i>No.</i>	<i>Factor</i>	<i>n Items</i>	<i>α</i>	<i>M</i>	<i>SD</i>	<i>Rank</i>
1.	Assessment is a diagnostic and formative tool	7	.8	27.00	5.77	2
2.	Assessment is irrelevant (bad, ignored & inaccurate)	8	.76	23.74	6.04	4
3.	Assessment is a school accountability tool	3	.75	8.42	3.07	5
4.	Assessment is a measurement and categorisation tool	5	.73	21.45	4.44	1
5.	Assessment is a valid grading tool	4	.61	13.82	3.37	3

Table 14

Factor Correlation Matrix

<i>No.</i>	<i>Factor</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
1.	Assessment is a diagnostic and formative tool	1.000	-.369	.231	.300	.024
2.	Assessment is irrelevant (bad, ignored & inaccurate)	-.369	1.000	-.084	-.007	-.063
3.	Assessment is a school accountability tool	.231	-.084	1.000	.288	.318
4.	Assessment is a measurement and categorisation tool	.300	-.007	.288	1.000	.073
5.	Assessment is a valid grading tool	.024	-.063	.318	.073	1.000

There were no significant differences between males and females on the majority of the factors (see Table 15). There was a statistically significant difference, however, in the mean level of endorsement between male and female participants in relation to the fourth factor which describes assessment as a measurement and categorisation tool. Female participants endorsed this factor more strongly than males. Notwithstanding the fact that this difference reached statistical significance, the

magnitude of the difference was very small ($d = -0.19$) and is therefore of little practical significance.

Table 15

Factor Mean Differences by Gender

	<i>Factor</i>	Male <i>M (SD)</i>	Female <i>M (SD)</i>	<i>t</i> (489)	Effect (Cohen's <i>d</i>)
1.	Assessment is a diagnostic and formative tool	26.28 (5.75)	27.34 (5.75)	-1.89	-0.18
2.	Assessment is irrelevant (bad, ignored & inaccurate)	24.27 (6.28)	23.50 (5.92)	1.30	0.13
3.	Assessment is a school accountability tool	8.49 (3.02)	8.38 (3.09)	0.37	0.03
4.	Assessment is a measurement and categorisation tool	20.85 (4.87)	21.72 (4.20)	-2.02*	- 0.19
5.	Assessment is a valid grading tool	13.99 (3.65)	13.74 (3.23)	0.78	0.07

* $p < .05$ two-tailed

A one-way between groups analysis of variance (ANOVA) was conducted to explore the impact of participants' teaching experience on the mean score for each of the five factors. Participants were divided into four groups according to the number of years teaching experience they had (0-5 years, 6-10 years, 11-20 years, 20+ years). No statistically significant differences were found in the factor mean scores for the four groups in the case of the first four factors. In the case of the fifth factor, however, there was a statistically significant difference at the $p < .05$ level for the mean factor scores between the groups: $F(3, 485) = 7.1, p = .00$. Post-hoc comparisons using the Tukey test indicated that the mean score for group 1 (0-5yrs) [$M = 12.61, SD = 3.33$] was statistically different from group 3 (11-20yrs) [$M = 13.96, SD = 3.19$] and group 4 (20+yrs) [$M = 14.54, SD = 3.51$]. The magnitude of the difference in the mean scores was close to a medium effect size in the case of group 1 and group 3 ($d = -0.41$) and was of a medium effect size in the case of group 1 and group 4 ($d = -0.57$).

Additional ANOVA tests were run to investigate the impact of school type and role on the mean score for each of the five factors. No statistically significant differences were found in factor mean scores for any of the school types explored. In relation to differences by role, only one statistically significant difference was found and this occurred with the school accountability factor, $F(7, 481) = 1.916, p = .05$. The Tukey test indicated that the mean level of endorsement of this factor by principals ($M = 9.25, SD = 3.25$) differed statistically significantly from that of retired teachers/teachers on a career break ($M = 6.36, SD = 2.90$). The magnitude of the difference in the means was large ($d = 0.94$).

A series of independent-samples t-tests were conducted to compare the mean factor scores for specific subject teacher groups and the remainder of the sample. Given the large volume of subject groups in the sample, only those which differed statistically significantly from the remainder of the sample are reported. As outlined in Table 16, physics and science teachers endorsed the first factor more strongly than the remainder of the sample. In the case of the science teachers, the effect size was small ($d = 0.3$) and in the case of the physics teachers, the effect size was medium ($d = 0.54$). A statistically significant difference was also found for this factor in relation to religion teachers. This group of teachers disagreed more than others that assessment is a diagnostic and formative tool. The effect size for this difference was small ($d = -0.29$).

Table 16

Factor 1: Statistically significant mean differences by subject

<i>Subject Name</i>	<i>Teachers in Subject Group M (SD)</i>	<i>Sample (less subject group) M (SD)</i>	<i>t (489)</i>	<i>Effect (Cohen's d)</i>
Religion	25.60 (5.51)	27.23 (5.78)	2.158*	-0.29
Physics	29.78 (4.83)	26.92 (5.78)	-2.087*	0.54
Science	28.47 (5.42)	26.78 (5.79)	-2.188*	0.30

* $p < .05$ two-tailed

Turning to factor 2, as illustrated in Table 17, CSPE, English and French teachers endorsed this factor more than others whereas business, mathematics and accounting teachers endorsed it less than others. The magnitude of the differences in each case was small ($d < 0.5$).

Table 17

Factor 2: Statistically significant mean differences by subject

<i>Subject Name</i>	<i>Teachers in Subject Group M (SD)</i>	<i>Sample (less subject group) M (SD)</i>	<i>t (489)</i>	<i>Effect (Cohen's d)</i>
Mathematics	22.67 (5.96)	24.10 (6.03)	2.260*	-0.24
CSPE	24.86 (6.26)	23.47 (5.96)	-2.034*	0.23
English	24.81 (5.27)	23.45 (6.21)	-2.263*	0.24
French	25.60 (5.62)	23.50 (6.05)	-2.477*	0.36
Accounting	21.70 (6.99)	23.88 (5.96)	2.153*	-0.42
Mathematics	22.67 (5.96)	24.10 (6.03)	2.260*	-0.24

* $p < .05$ two-tailed

Only one statistically significant difference emerged in relation to factor 3 (see Table 18). Geography teachers agreed more than others with the idea of assessment for school accountability purposes. Once again the effect size for this difference was small ($d = 0.28$).

Table 18

Factor 3: Statistically significant mean differences by subject

<i>Subject Name</i>	<i>Teachers in Subject Group M (SD)</i>	<i>Sample (less subject group) M (SD)</i>	<i>t (489)</i>	<i>Effect (Cohen's d)</i>
Geography	9.13 (2.82)	8.31 (3.09)	-1.988*	0.28

* $p < .05$ two-tailed

In the case of the fourth factor, German, history and French teachers endorsed this summative conception of assessment more strongly than others, with the effect sizes ranging from small to almost moderate ($0.26 > d < 0.47$). SPHE teachers on the other hand endorsed this factor less than others ($d = -0.40$) (see Table 19).

Table 19

Factor 4: Statistically significant mean differences by subject

<i>Subject Name</i>	<i>Teachers in Subject Group M (SD)</i>	<i>Sample (less subject group) M (SD)</i>	<i>t (489)</i>	<i>Effect (Cohen's d)</i>
French	22.68 (3.41)	21.28 (3.41)	-2.79*	-0.35
History	22.38 (3.91)	21.27 (4.51)	-2.007*	0.26
German	23.31 (4.01)	21.34 (4.44)	-2.208*	0.47
SPHE	19.77 (4.61)	21.56 (4.41)	2.176*	-0.40

* $p < .05$ two-tailed

Finally, in relation to the fifth factor, home economics teachers disagreed moderately more than others with the idea of grading ($d = 0.64$). This mean difference had the largest effect size of all the statistically significant subject mean differences (see Table 20).

Table 20

Factor 5: Statistically significant mean differences by subject

<i>Subject Name</i>	<i>Teachers in Subject Group M (SD)</i>	<i>Sample (less subject group) M (SD)</i>	<i>t (489)</i>	<i>Effect (Cohen's d)</i>
Home Economics	11.83 (3.21)	13.92 (3.35)	2.929**	0.64

** $p < .01$ two-tailed

Chapter 5: Conclusions

Introduction

This study set out to investigate how Irish post-primary teachers currently conceive of assessment. At present, the international assessment landscape is characterised by the interplay of multiple assessment purposes which are underpinned by the priorities of economic competitiveness, performativity and contemporary learning theory. Harmonisation of these conflicting purposes for assessment presents education systems internationally with a great challenge (Wyse et al., 2016). Junior Cycle education in Ireland is currently responding to this challenge. A new policy for reform, *Framework for Junior Cycle* (2015), introduces school-based assessment, for the first time, to an educational context that has heretofore been dominated by externally assessed examinations. Teachers are now centrally positioned in the assessment process and are faced with the challenge of reconciling the new policy mandates with their current assessment beliefs and practices. While it is acknowledged that teacher assessment capacity has many constituent parts (Looney et al., 2017; Xu & Brown, 2016), beliefs have been shown to play a particularly significant role in times of educational change (Fullan, 2007; Gardner & Galanouli, 2016). Research has shown that teachers' beliefs act as an interpretive screen which filters content entering the cognitive domain, frames particular educational situations or problems and guides teachers' intentions and actions. Hence, it was argued, that in the current context of policy change in Ireland, an understanding of how teachers conceptualise assessment is an important starting point for the planning of professional development support. The purpose of this study, therefore, was to elicit large-scale base-line data in order to answer the question of how Irish post-primary teachers conceptualise assessment.

Adopting a non-experimental cross-sectional design, a large sample ($n=489$) of post-primary teachers were surveyed using the abridged version of Brown's (2006) Teachers' Conceptions of Assessment Inventory (TCoA-IIIa). The structure, strength and interconnectedness of the teachers' conceptions were analysed using a mixture of descriptive statistics, exploratory factor analysis, independent samples t-tests and one-way analysis of variance.

This chapter provides a critical review of the study, its findings and its implications, in light of previous research and literature in the field, and current policy reform in the Irish context. The chapter begins by presenting an analysis of the findings in relation to factor structure, factor endorsement, factor correlations and factor endorsement by subject domain. Consideration is then given to the overall conclusions and implications of the research. Finally, the chapter acknowledges the limitations of the study and the potential for future research.

Analysis of Findings

Factor Structure

The results from this study paint a picture of how post-primary teachers in Ireland conceive of assessment amidst curriculum and assessment reform at Junior Cycle. While Brown's (2006) original factor structure suggested that conceptions of assessment are best represented as a hierarchical structure of four intercorrelated assessment purposes, this four-factor framework proved insufficient in capturing the full range of teachers' conceptions of assessment in the Irish context. Exploratory factor analysis recovered a five-factor structure from the data which included two of the factors (*school accountability and irrelevance*) from Brown's (2006) original model.

The five factors, in line with previously discussed conception research (Barnes et al., 2015; Brown & Gao, 2015; Gardner & Galanouli, 2016; Remesal, 2011), map readily onto a continuum of assessment purposes ranging from assessment for improvement purposes to assessment for grading and accountability purposes. The range of factors extracted from the data appear to reflect the multifaceted reality of assessment in post-primary schools in Ireland, thereby supporting the contention by Brown and Harris (2009) that teachers' conceptions of assessment are ecologically rational. The education system at post-primary in Ireland simultaneously uses assessment to guide improvements in teaching and learning (formative assessment), to grade the quality of student learning as part of the national qualifications system at Junior Cycle (Junior Certificate / Junior Cycle Examination) and at Senior Cycle (Leaving Certificate Examination), to hold schools accountable through the publication of national league tables, and with the implementation of the current reform movement, to partially assess student learning during Junior Cycle through a school-based model of assessment. The emergence of four factors, corresponding broadly to these four purposes, suggests an understanding in the minds of the participants about the multiplicity of assessment purposes at play in the Irish educational context. Furthermore, the attachment of validity items (*assessment results are consistent/can be depended upon /are trustworthy*) to three of the factors (*Assessment is a valid grading tool*, *Assessment is a measurement and categorisation tool* and *Assessment is a diagnostic and formative tool*) points to a particular perspective in this sample in relation to dependable and meaningful assessment types. Consideration of the factor endorsement levels and factor correlations provides further insight into this perspective.

Factor Endorsement Levels and Factor Correlations

The teachers in this study most strongly endorsed a summative conception of assessment as a measurement and categorisation tool which provides feedback to students about their performance (factor 4). This factor, which was not part of Brown's (2006) original structure, reflects an understanding of summative assessment as integrated with the teaching and learning process. Consistent with the demands of an examinations system, which, heretofore, was completely externally assessed, regular summative assessment is a key element of post-primary teachers' work in Irish classrooms. The participants' perception of it as a valid form of assessment is, therefore, unsurprising. What is surprising, however, is the dichotomous distinction in the data between classroom-based summative assessment (factor 4) and summative assessment through grading (factor 5). The assignment of discrete functions to these two assessment types would suggest that the categorisation of learning (*assessment places students into categories*) referred to in factor 4 is conducted by some means other than grading. As mentioned in the introductory chapter, the recent introduction of school-based assessment at Junior Cycle includes two classroom-based assessments which are assessed in line with quality descriptors in place of grades (DES, 2015a). The quality descriptors in question are exceptional, above expectations, in line with expectations and yet to meet expectations (www.juniorcycle.ie). Notwithstanding the fact that implementation of these measures is ongoing, and that English, Science, Business, Gaeilge and Modern Foreign language teachers are the only ones to have received training in the assessment procedures thus far (www.jct.ie), perhaps the emergence of this factor points to the general ecological influence of the new assessment rhetoric on Irish teachers' conceptions of assessment? Such a finding would align with the literature which points to the influence of contextual imperatives on

teachers' beliefs generally (Buehl & Beck, 2015; Fives & Buehl, 2012) and teachers' conceptions of assessment more specifically (Brown et al., 2011; Brown & Harris, 2009). Furthermore, it is worth noting at this point that the appearance of this summative factor in the Irish data is consistent with the recovery of a very similar factor in the New Zealand study (Yates & Johnston, 2017) investigating the impact of school-based assessment for qualifications on teachers' conceptions of assessment. The factor, which consisted of the following items, (*assessment is integrated with teaching practice, assessment places students into categories and assessment provides feedback to students about their performance*), was labelled *Assessment is for qualifications*. Yates and Johnston (2017) interpreted this factor as representing the school-based assessment component which accounts for around 60% of students' grades in the National Certificate of Educational Achievement (NCEA) in New Zealand. While the school-based assessment component in the Irish context is somewhat different in that it is not combined with the students' results from external assessment, the overall model of internal assessment for certification is quite similar. From a policy perspective, the endorsement of this factor is a welcome finding as it may be indicative of an openness to classroom-based assessment, a topic which has been at the heart of industrial relations unrest among the ASTI until recently. While this study did not gather information on whether participants were or were not affiliated with a particular teaching union (due to the delicate nature of this topic at the time of data collection), it is noteworthy that no statistically significant differences occurred in the mean endorsement scores for this factor by role, school type or years of experience. Such a finding would point to a widespread level of endorsement for the factor.

The moderate positive correlation ($r = .30$) between factor 4 and the second most strongly endorsed factor, *Assessment is a diagnostic and formative tool* (factor 1), is

worthy of discussion at this point. The association teachers made between these two factors could be interpreted in numerous ways. It could be that the teachers conceive of classroom-based summative assessment as having a dual purpose of gauging and guiding learning. On the other hand, it is possible that the teachers conceive of formative assessment as playing a role in guiding students to successful outcomes in the classroom-based assessment tasks. When one looks at the factor items more closely, however, the latter of the two options above, arguably, appears more probable. This suggestion is presented in light of the fact that the highest loading items on each of these factors both relate to feedback but are clearly distinguished from each other in terms of purpose (feedback about *learning needs* (factor 1) and feedback about *performance* (factor 4)). It would appear, therefore, that participants believe that internal summative assessment feeds back information on performance, but that this does not equate to feedback which helps learners to know “where they need to go and how best to get there” (Assessment Reform Group, 2002). The moderate level of endorsement for the formative assessment factor would suggest that the participants in general value its worth. This is a positive finding which aligns well with the emphasis on formative assessment set out in the Junior Cycle reform measures. However, despite the fact that some studies have shown positive endorsement of formative conceptions of assessment to be a predictor of diagnostic assessment practices (Brown et al., 2015; Brown et al., 2012; Panadero et al., 2014) one must approach the above finding with caution as the belief-practice relationship is not always linear in nature (Buehl & Beck, 2015; Thompson, 1992). As was noted in the literature review, as formative assessment is a process which requires teachers to “reconstruct the teaching contract” (Perrenoud, 1991, p.92), deep learning (Marshall & Drummond, 2006) and understanding of underlying pedagogical principles (Hopfenbeck et al., 2015) is

required. Therefore, drawing on Xu and Brown (2016) and Looney et al. (2017), while one might endorse a formative conception of assessment, successful engagement with the “spirit” (Marshall & Drummond, 2006) of the process could be mediated by such aspects as inadequate pedagogical knowledge, lack of confidence, accountability demands and negative school culture to name but a few. The extent to which the participants’ conception of formative assessment is in line with their practice is worthy of exploration. This is an issue which will be further discussed at a later point in the context of future research.

Another noteworthy finding relating to factor 1 (*Assessment is a diagnostic and formative tool*) was its lack of correlation with the fifth factor (*Assessment is a valid grading tool*). Studies conducted using Brown’s (2006) TCoA-IIIa in high-stakes examination contexts in Hong Kong (Brown et al., 2009) and China (Brown et al., 2011) revealed strong correlations ($r = .91$ / $r = .80$) between the improvement conception of assessment and assessment for the purposes of grading. Such strong correlations speak to the conviction of teachers in Chinese contexts that examining students is the best way to improve their learning. As post-primary teachers in Ireland have been working for a long time in a high-stakes system at Junior and Senior Cycle, and continue to work in a high-stakes system at Senior Cycle, it is somewhat surprising that the correlation between grading and improvement of learning (factor 1) isn’t higher. The grading item (*assessment is assigning a grade or level to student work*) in factor 5 sits alongside all three of the validity items, suggesting that the teachers view grading as a valid form of assessment but not for improvement purposes. Perhaps the teachers view grading as a valid way of assessing learning for the purposes of reporting and third level opportunities? This would reflect the well-established Junior and Leaving Certificate Examination processes in Ireland. It is important at this point,

however, to draw attention to a statistically significant difference which occurred in the mean endorsement levels by years of experience for assessment as a grading tool. As highlighted in Chapter 4, teachers with 11-20 and 20+ years of teaching experience agreed moderately more than teachers with 0-5 years of experience with this factor. This is a logical finding when it is positioned in the realm of belief theory which argues that one's beliefs, actions and environment are all shaped by and shape each other (Bandura, 1997; Fives & Buehl, 2012). The teachers with more experience have spent longer working in a high-stakes system which has held them accountable through the Junior and Leaving Cert Examination system. The pressures which such a system can exert, as well as the intense school cultures which can develop as a result of it, can influence teachers' beliefs. Furthermore, a lot of the older teachers in the sample are likely to have formed deep-seated beliefs about assessment, through their own "apprenticeship of observation" (Lortie, 1975), at a time when the mantra of formative assessment was not centre stage in educational debate. On the other hand, while the teachers with 0-5 years of experience would also have been educated through the Junior and Leaving Certificate systems, their schooling occurred at a time when factors such as formative assessment, 21st century skills and digital technology were shaping the international educational narrative. There is a chance, therefore, that such developments may have influenced their everyday experience of assessment. Furthermore, 14 of the 95 teachers within this category classified themselves as trainee teachers and as such would be aware of the drive in Irish education towards a more balanced approach to assessment.

While the participants in this study did not associate grading with improving learning, they did moderately associate grading with school accountability ($r = .32$). This is not a surprising finding given the high-stakes qualification system in place in

Ireland. Schools, and by default, teachers, are judged on the results that their students obtain in the Junior and Leaving Certificate Examinations. Rates of progression to third level from each school are published nationally in league tables. These league tables rank schools regionally and nationally and hold considerable sway in the public domain. This type of practice narrows the focus of educational success by defining the quality of students' development and educational experience solely in terms of grades. The participants' rejection (i.e. less than slight agreement) of the school accountability factor suggests that they do not accept the legitimacy of the accountability mechanisms currently in place in the post-primary system. When the endorsement level for this factor was broken down further by role, one statistically significant difference did occur between principals who endorsed the factor the most and retired/career break teachers who endorsed it the least. It stands to reason that as the managers of schools, principals would be particularly concerned with how accountability measures influence the image and success of their schools. On the other hand, those who are retired or on career break would most likely be looking at education through a much broader external lens, detached somewhat from the running of schools. Overall, the disagreement with this factor would indicate that in the minds of the teachers in this sample, school accountability, as it currently stands, only paints part of the educational picture. This finding suggests that the broader reporting style included as part of Junior Cycle reform could be widely welcomed. The new certificate, the Junior Cycle Profile of Achievement (JCPA), will report not only on the students' grades in their terminal exam but on their achievement in their classroom-based assessments as well as their achievement in other areas of learning such as wellbeing. This change is a step towards a broader perspective on school accountability where more than grades and results are valued.

In relation to the *Irrelevance* factor, a particular finding emerged from the data which is not unique to the Irish context and is worthy of further discussion. While teachers tended to disagree with the idea of assessment as bad or ignored, they agreed with the notion of measurement error in assessment. The items they agreed with the most from this factor were *Teachers should take into account the error and imprecision in all assessment* (86% overall agreement) and *Assessment results should be treated cautiously because of measurement error* (80% overall agreement). In the recent New Zealand study by Yates and Johnston (2017), the first item mentioned above was also highlighted for high levels of agreement. The authors explained that “the data showed relatively little response variability for this item with 82% of respondents either slightly, moderately or mostly agreeing” (Yates & Johnston, 2017, p.9). Agreement with the idea of inaccuracy in assessment also emerged from other studies with pre-service teachers in Spain (Brown & Remesal, 2012) and Canada (Daniels et al., 2014) and with primary and secondary teachers in New Zealand (Brown, 2011). Daniels et al. (2014) noted that the pre-service teachers in the Canadian study were enrolled on a mandatory assessment course learning about imprecision and error in assessment and may therefore have been sensitive to these issues. Brown (2011) noted in relation to the New Zealand data that “it would be interesting to discover whether this level of agreement applied equally to formal, externally sourced tests and examinations and to teacher-made observations, checklists, and judgements” (p.11). It would also be interesting, in the case of the Irish data, to further explore the types of assessment that the teachers consider to be inaccurate. Surmising on the evidence available from this study, particularly drawing on the position of the validity items, perhaps participants associate measurement error with the assessment systems in place for school accountability?

Conclusions and Implications

So what conclusions can be drawn about Irish secondary teachers' conceptions of assessment? The teachers in this study displayed broad and multidimensional conceptions of assessment which correspond to the heterogeneous assessment landscape at post-primary in Ireland and in education internationally. The extraction of five factors, and the intercorrelations between them, speak to the complexity of assessment purposes in the Irish context and demonstrates that post-primary teachers in Ireland "do not exhibit the simplistic notion of formative assessment good, summative assessment bad" (Brown et al., 2011, p.217). The somewhat unexpected distinction that teachers made between summative assessment for grading purposes and in-class summative assessment for the purpose of measurement, categorisation and feedback, illustrates a nuanced understanding of summative assessment purposes and is possibly suggestive of engagement with recent Junior Cycle reform measures which include classroom-based assessment. While a clear rank order of endorsement for the factors emerged from the data, the lack of strong agreement with any of the factor points to the potential for fluidity within these conceptions – a finding which could have beneficial or detrimental implications for policy change. On the one hand, in relation to the filtering role of conceptions in times of reform (Fives & Buehl, 2012; Fullan, 2007), a lack of strong commitment to any particular conception could result in teachers being more open and receptive to change and development (Green, 1971; Rokeach, 1968). On the other hand, in the absence of any clear and decisive stance, teachers may be more susceptible to contextual constraints and accountability pressures and may allow these to supersede any other beliefs they may have. That said, the teachers in this sample would appear to have the courage of their convictions given their endorsement

of formative assessment and their rejection of school accountability in a system which has been dominated mainly by high-stakes examinations. Consistent with studies in New Zealand (Brown, 2004) and Canada (Daniels et al., 2014), the analysis also revealed that in spite of a small number of statistically significant differences, the teachers' conceptions of assessment appear to be relatively homogenous regardless of gender, school type, experience, role or subject.

The literature states that one of the key first steps in teacher belief change or development is awareness (Cabaroglu & Roberts, 2000; Fives & Buehl, 2012; Kagan, 1992). If teachers are to engage in deeper learning about assessment and ultimately re-construct their identity as assessors (Xu & Brown, 2016), then they must be initially supported to reflect upon their own beliefs and assumptions about assessment. Participation in this study has allowed teachers to do just that. By engaging with and completing the TCoA-III inventory, teachers were furnished with an opportunity to make their conceptions of assessment explicit. Cognisant of the fact that this is but a first step in the belief extraction process, the information gleaned in this study provides an important insight into the lens through which post-primary teachers might translate policy into practice. Confirmation of this conception structure, as well as further insight into its origin and potential future trajectory, is needed. The potential for future research to support the interpretations offered here is outlined below following a discussion of the limitations of this study.

Limitations

As with all research in the social sciences, the limitations of the study here should be borne in mind when interpreting the findings. Firstly, as it was not feasible to utilise simple random sampling in this study (see chapter 3), generalisability to the post-primary teaching population in Ireland cannot be asserted. However, despite the reliance on volunteer sampling as a data collection method for this study, a large sample was obtained ($n=489$) with a gender breakdown which closely represents the post-primary teaching population in Ireland. While under- and over-representation in relation to other categorical variables such as role was identified, the extent of the divergence was not so extreme to cast doubt on the overall thrust of the findings presented here.

That said, it is acknowledged that the under-representation of classroom teachers in this study may have been influenced by the contentious industrial relations climate which existed at the time of data collection for this study. With the Junior Cycle reform measures now adopted by all teaching unions, were a similar study to be conducted again in the future, it is possible that classroom teachers might be better represented.

Thirdly, bearing in mind that research on the efficacy of using self-report instruments points to the potential for response bias and inaccuracy in reporting (Hoffman & Seidel, 2015), validation of the findings, through other means, is recommended.

Recommendations for Future Research

Firstly, the findings from this study could be used to inform the design of professional development support for Junior Cycle reform in Ireland. As has been outlined in this study, while teacher assessment competence is comprised of many constituent parts, beliefs appear to play a particularly significant role in times of reform. The importance of this role could be reflected in professional development design which initially steps back from assessment practice in order to engage with teachers' conceptions of assessment as an important point of departure in their own right. The findings from this study, in terms of factor structure, endorsement levels and correlations, could act as a useful starting point for discussion, thereby encouraging teachers to engage with their own assessment beliefs, and with the reform agenda at large, in a deeper way.

This study opens doors for further investigation of teachers' conceptions of assessment in the Irish context. Firstly, the five-factor structure extracted from this study points to the potential for modification and refinement of Brown's (2006) TCoA-III A to better suit the Irish context. Suggestions include the amendment of the *Improvement* factor so that it better reflects a truly diagnostic and formative use of assessment. It is possible that one or two of the items currently in this factor (e.g. *assessment is a way to determine how much students have learned from teaching/assessment provides feedback to students about their performance*) would be better positioned as part of an additional summative purpose for assessment, particularly given the introduction of school-based assessment. Secondly, given that the item (*assessment has little impact on teaching*) did not load adequately on any factor, had a very low communality value, and that 77% of participants either mostly or strongly disagreed with it, perhaps this item should be dropped from the questionnaire

for the Irish context. Thirdly, it is interesting that the item (*assessment determines if students meet qualifications standards*) did not load adequately enough on any factor despite 83% of participants agreeing with it in total. It is possible that extra items related to the same purpose need to be added to the instrument so that this purpose for assessment is extracted as a factor should it exist among a sample. Finally, it would appear from the findings in this study that participants interpreted the first-order *inaccuracy* factor in a different way to the other items in the overarching *Irrelevance* factor. Given that participants agreed a lot more with the items related to error and inaccuracy, it could be argued that this agreement inflated the overall result for a factor that participants otherwise conceived of quite negatively. For this reason, it is suggested that the removal of the *inaccuracy* sub-factor from the overarching *Irrelevance* factor might be warranted. A future national study using a modified version of the TCoA-IIIa and simple random sampling has the potential to shine further light on the findings from this study.

In terms of the evidenced-based data gathered in this study, follow-up research through focus groups discussions could be a very useful way of further exploring the more contextual and emotional aspects of the conceptions uncovered. Possible frameworks for such discussions include the exploration of the internal and external factors which have contributed to shaping the teachers' conceptions of assessment (Buehl & Beck, 2015, Xu & Brown, 2016); the degree to which school culture specifically supports or hinders the translation of these conceptions into practice (Goddard & Goddard, 2001); the extent to which the teachers' sense of self-efficacy influences how they act upon their conceptions (Looney et al., 2017; Yan, 2014; Yang & Cheng, 2015); the practices which teachers associate with their conceptions; and further exploration of their rejection of the school accountability factor. A focus group

discussion could also provide teachers with the opportunity to further expand upon and clarify the factor correlations which emerged in this study. A lot of potential therefore exists for follow-up qualitative research to further inform the findings uncovered in this study.

Another possible route of further research would be to quantitatively or qualitatively examine the relationship between the teachers' conceptions of assessment and their assessment practices. Information on the alignment between the two would provide a clearer picture as to the impact of internal and external mediating factors on the belief-practice relationship in the Irish context.

Examining how teachers' conceptions of assessment develop or change over time would also be a possible angle of further inquiry. Given that the Junior Cycle policy reform is still in the initial stages of implementation, a longitudinal approach to investigation could track the impact of the reform on teachers' conceptions of assessment throughout the implementation process.

Another area meriting future research is the degree of congruence between teachers' and students' conceptions of assessment in the Irish post-primary context. Alignment of these two sets of conceptions would appear to be particularly essential given the required mutual commitment to formative assessment underpinning the policy reform.

Epilogue

This study presents a key message to policy makers in the context of Junior Cycle assessment reform in Ireland. If teachers are to implement valid and meaningful school-based assessment, then they need to believe in its worth. Teachers interpret educational change through the prism of their own beliefs. This lens is powerful and must not be ignored. Existing as a complex interconnected web, teachers' beliefs are multifaceted, sometimes conflicting, and open to contextual influence. Teachers build up a set of beliefs overtime which become a guiding philosophy for their everyday practice. The introduction of change, which does not align with this philosophy, can be unsettling. This is normal. Teaching, after all, is a human endeavour, and the human condition is often resistant to change. If change is to proceed in any form, however, it is during this unsettling period that teachers need to be encouraged and supported to critically engage with their underlying conceptions. The pace of school life does not often allow for the unearthing of beliefs and the questioning of assumptions. This, however, must change. If policy makers fail to create a space for deep mutual deliberation about how teachers' conceive of policy change, then surface-level change is all that is likely to occur.

So, let that space be created. Let us see teachers being encouraged to interrogate assessment policy and to question their assumptions for the betterment of education in Ireland.

This is essential at the current juncture because teachers' conceptions of assessment really do matter.

References

- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32(4), 665-683. doi:10.1111/j.1559-1816.2002.tb00236.x
- Ashton, P. T. (2015). Historical overview and theoretical perspectives of research on teachers' beliefs. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 31-47). New York: Routledge.
- Assessment Reform Group. (2002). *Assessment for Learning: 10 Principles*. Retrieved from <http://cdn.aiaa.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf>
- Association of Secondary Teachers in Ireland. (2014). Teachers' key concerns. *Nuacht* (Issue 4). Retrieved from http://www.asti.ie/uploads/media/Nuacht_press_ready.pdf
- Association of Secondary Teachers in Ireland. (2017a). *Junior Cycle Reform Not Harm*. Retrieved from http://www.asti.ie/fileadmin/user_upload/Documents/Campaigns/ASTI_Junior_Cycle_A5.pdf
- Association of Secondary Teachers in Ireland. (2017b) ASTI suspends industrial action [Press release]. Retrieved from <http://www.asti.ie/news/latest-news/news-article/article/asti-suspends-industrial-action-1/>
- Ball, S. J. (2013). *The education debate* (2nd ed.). Great Britain: The Policy Press.

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.
- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 284-300). New York: Routledge.
- Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16*(Series B), 296-298.
- Beijaard, D., Meijer, P., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*, 107-119.
- Beswick, K. (2007). Influencing teachers' beliefs about teaching mathematics for numeracy to students with mathematics learning difficulties. *Mathematics Teacher Education & Development, 9*, 3-20.
- Black, P. (2015). Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice, 22*(1), 161-177.
doi:10.1080/0969594X.2014.999643
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.
doi:10.1007/s11092-008-9068-5
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.
doi:10.1080/0969595980050102
- Black, T. R. (1999). *Doing quantitative research in the social sciences*. London: Sage.

- Broadfoot, P. M. (2002). *Learnacy or lunacy?* Paper Presented to the International Association for Educational Assessment Annual Conference, Hong Kong.
- Bronfenbrenner, U. (1989). Ecological systems theory. In R. Vasta (Ed.), *Annals of child development* (Vol. 6, pp. 187-251). Greenwich, CT: JAI Press.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318. doi:10.1080/0969594042000304609
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports*, 99, 166-170.
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.
- Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70.
- Brown, G. T. L., Chaudhry, H., & Dhamija, R. (2015). The impact of an assessment policy upon teachers' self-reported assessment beliefs and practices: A quasi-experimental study of Indian teachers in private schools. *International Journal of Educational Research*, 71, 50-64.
doi:<http://doi.org.dcu.idm.oclc.org/10.1016/j.ijer.2015.03.001>
- Brown, G. T. L., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2(1), 993836. doi:10.1080/2331186X.2014.993836

- Brown, G.T.L., & Harris, L.R. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of MultiDisciplinary Evaluation*, 6(12), 68-91.
- Brown, G. T. L., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28(7), 968-978.
doi:<http://dx.doi.org.remote.library.dcu.ie/10.1016/j.tate.2012.05.003>
- Brown, G. T. L., Hui, S. K. F., Yu, F. W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5–6), 307-320.
doi:<http://dx.doi.org.remote.library.dcu.ie/10.1016/j.ijer.2011.10.003>
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363. doi:10.1080/09695940903319737
- Brown, G. T. L., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210-220. doi:10.1016/j.tate.2010.08.003
- Brown, G. T. L., & Michaelides, M. P. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand.

- European Journal of Psychology of Education*, 26(3), 319-337.
- doi:10.1007/s10212-010-0052-3
- Brown, G. T. L., & Remesal, A. (2012). Prospective teachers' conceptions of assessment: A cross-cultural comparison. *The Spanish Journal of Psychology*, 15(01), 75-89. doi:10.5209/rev_SJOP.2012.v15.n1.37286
- Brownlee, J. (2003). Changes in primary school teachers' beliefs about knowing: A longitudinal study. *Asia-Pacific Journal of Teacher Education*, 31(1), 87-96.
- Buehl, M. M., & Beck, J. S. (2015). The relationship between teachers' beliefs and teachers' practices. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 66-84). New York: Routledge.
- Buehl, M. M., & Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? Does it change? *Journal of Experimental Education*, 77(4), 367-408.
- Bullough, R. V. (2015). Methods for studying beliefs: Teacher writing, scenarios, and metaphor analysis. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 150-169). New York: Routledge.
- Cabaroglu, N., & Roberts, J. (2000). Development in student teachers' pre-existing beliefs during a 1-year PGCE programme. *System*, 28(3), 387-402.
- doi:[http://dx.doi.org.dcu.idm.oclc.org/10.1016/S0346-251X\(00\)00019-1](http://dx.doi.org.dcu.idm.oclc.org/10.1016/S0346-251X(00)00019-1)
- Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255-296). New York: Macmillan.

- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). USA: Routledge.
- Cornett, J. W. (1990). Teacher thinking about curriculum and instruction: A case study of a secondary social studies teacher. *Theory & Research in Social Education*, 18(3), 248-273. doi:10.1080/00933104.1990.10505617
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment*, 10(7), 1-9.
- Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44(4), 613-642. doi:10.1002/tea.20157
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. London: Sage.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th, International student ed.). London: SAGE.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Crooks, T. J. (2003). *Some criteria for intelligent accountability applied to accountability in New Zealand*. Paper presented to the Annual Conference of the American Education Research Association, Chicago. Retrieved from <https://www.fairtest.org/some-criteria-intelligent-accountability-applied-a>

- Crooks, T. J. (2011). Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation, 37*, 71-77.
- Daniels, L. M., Poth, C., Papile, C., & Hutchison, M. (2014). Validating the conceptions of assessment-III scale in Canadian preservice teachers. *Educational Assessment, 19*(2), 139-158. doi:10.1080/10627197.2014.903654
- Darling-Hammond, L. (2003). Standards and assessments: Where we are and what we we need. *Teachers College Record*.
- Davis, D. S., & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *Journal of Educational Research, 104*(3), 202-215. doi:10.1080/00220671003690148
- DeLuca, C., & Bellara, A. (2013). The current state of assessment education. *Journal of Teacher Education, 64*(4), 356-372. doi:10.1177/0022487113488144
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability, 28*(3), 251-272. doi:10.1007/s11092-015-9233-6
- DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy & Practice, 24*(2), 121-126. doi:10.1080/0969594X.2017.1297010
- Department of Education and Skills. (1995). *Charting our Education Future – White Paper on Education (Launch Copy)*. Retrieved from

<http://www.education.ie/en/Publications/Policy-Reports/Charting-Our-Education-Future-White-Paper-On-Education-Launch-Copy-1995-.pdf>

Department of Education and Skills. (2011). *Literacy and Numeracy for Learning and Life: The National Strategy to Improve Literacy and Numeracy among Children and Young People 2011-2020*. Dublin: Author. Retrieved from https://www.education.ie/en/Publications/Policy-Reports/lit_num_strategy_full.pdf

Department of Education and Skills. (2012). *A Framework for Junior Cycle*. Retrieved from <http://www.education.ie/en/Publications/Policy-Reports/A-Framework-for-Junior-Cycle-Full-Report.pdf>

Department of Education and Skills. (2013). *Chief Inspector's Report 2010-2012*. Retrieved from <http://www.education.ie/en/Publications/Inspection-Reports-Publications/Evaluation-Reports-Guidelines/Chief-Inspector's-Report-2010-2012-Main-Report.pdf>

Department of Education and Skills. (2014). *Looking at Mathematics: Teaching and Learning in Post-Primary Schools*. Retrieved from <http://www.education.ie/en/Publications/Inspection-Reports-Publications/Evaluation-Reports-Guidelines/Looking-at-Mathematics-A-composite-report-based-on-findings-of-subject-inspections-in-Mathematics-conducted-in-the-2009-2010-school-year.pdf>

Department of Education and Skills. (2015a). *Framework for Junior Cycle*. Dublin: Author. Retrieved from <https://www.education.ie/en/Publications/Policy-Reports/Framework-for-Junior-Cycle-2015.pdf>

- Department of Education and Skills. (2015b). *Appendix to the Joint Statement on Principles and Implementation*. Retrieved from <http://education.ie/en/Schools-Colleges/Information/Curriculum-and-Syllabus/Junior-Cycle-/Junior-Cycle-Reform-Appendix-to-Joint-Statement-on-Principles-and-Implementation.pdf>
- Department of Education and Skills. (2016a). *Key Statistics 2015-2016*. Retrieved from <http://www.education.ie/en/Publications/Statistics/Key-Statistics/Key-Statistics-2015-2016.pdf>
- Department of Education and Skills. (2016b). *Post Primary Staffing Arrangements for 2016/2017*. Retrieved from http://www.education.ie/en/Circulars-and-Forms/Active-Circulars/cl0003_2016.pdf
- Devine, D., Fahie, D., & McGillicuddy, D. (2013). What is 'good' teaching? Teacher beliefs and practices about their teaching. *Irish Educational Studies*, 32(1), 83-108. doi:10.1080/03323315.2013.773228
- Dillman, D. A. (2007). *Mail and internet surveys: The tailored design method* (2nd ed.). Hoboken, N.J.: Wiley.
- Elwood, J., & Murphy, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182-192. doi:10.1080/0969594X.2015.1021568
- Ertmer, P. A., Ottenbreit-Leftwich, A. T., & Tondeur, J. (2015). Teachers' beliefs and uses of technology to support 21st-century teaching and learning. In H. Fives, &

- M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 403-418). New York: Routledge.
- Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford: Oxford University Press.
- Field, A. P. (2009). *Discovering statistics using SPSS: And sex and drugs and rock 'n' roll* (3rd ed.). London: SAGE.
- Fives, H., & Alexander, P. A. (2004). How schools shape teacher efficacy and commitment: Another piece in the achievement puzzle. In D. M. McInerney, & S. Van Etten (Eds.), *Big theories revisited* (pp. 329-359). Greenwich, CT: Information Age Publishing.
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham & T. Urda (Eds.), *APA educational psychology handbook: Individual differences and cultural and contextual factors* (Vol. 2, pp. 471-499). Washington, DC: American Psychological Association.
- Freire, P. (1972). *Pedagogy of the oppressed*. Great Britain: Penguin.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York: Teachers College Press.
- Gaitas, S., & Alves Martins, M. (2015). Relationships between primary teachers' beliefs and their practices in relation to writing instruction. *Research Papers in Education*, 30(4), 492-505. doi:10.1080/02671522.2014.908406

- Gardner, J., & Galanouli, D. (2016). Teachers' perceptions of assessment. In D. Wyse, L. Hayward & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy, and assessment* (Vol. 2, pp. 710-724). London: Sage.
- Gebril, A., & Brown, G. T. L. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 21(1), 16-33.
doi:10.1080/0969594X.2013.831030
- Gipps, C. (2002). *Discussant Presentation – Round table on school-based assessment*. Paper Presented to the International Association for Educational Assessment Annual Conference, Hong Kong.
- Gipps, C., Brown, M., McCallum, B., & McAlister, S. (1995). *Intuition or evidence? Teachers and national assessment of seven year olds (assessing assessment)*. Buckingham, UK: Open University Press.
- Glackin, M. (2016). 'Risky fun' or 'Authentic science'? How teachers' beliefs influence their practice during a professional development programme on outdoor learning. *International Journal of Science Education*, 38(3), 409-433.
doi:10.1080/09500693.2016.1145368
- Goddard, R. D., & Goddard, Y. L. (2001). A multilevel analysis of the relationship between teacher and collective efficacy in urban schools. *Teaching and Teacher Education*, 17, 1-12. doi:[http://dx.doi.org/10.1016/S0742-051X\(01\)00032-4](http://dx.doi.org/10.1016/S0742-051X(01)00032-4)

- Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher*, 33(3), 3-13. doi:10.3102/0013189X033003003
- Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.
- Green, T. F. (1971). *The activities of teaching*. New York: McGraw-Hill.
- Grek, S., Lawn, M., Lingard, B., Ozga, J., Rinne, R., Segerholm, C., & Simola, H. (2009). National Policy brokering and the construction of the European Education Space in England, Sweden, Finland and Scotland. *Comparative Education*, 45(1), 5-21. Doi: 10.1080/03050060802661378
- Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. E. Farstrup, & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 370-291). Newark, DE: International Reading Association.
- Hall, C. (2000). National certificate of educational achievement: Issues of reliability, validity and manageability. *New Zealand Annual Review of Education*, 9, 173-196.
- Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365-381. doi:10.1080/09695940903319745
- Hartas, D. (Ed.). (2010). *Educational research: Qualitative and quantitative approaches*. London: Continuum International Publishing Group.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi:10.3102/003465430298487
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 27-43.
doi:10.1080/0969594X.2014.984656
- Hoffman, B. H., & Seidel, K. (2015). Measuring teachers' beliefs: For what purpose? In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 106-127). New York: Routledge.
- Hopfenbeck, T. N., Flórez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of assessment for learning in Norway. *Assessment in Education: Principles, Policy & Practice*, 22(1), 44-60.
doi:10.1080/0969594X.2014.996524
- Hutchinson, C. & Young, M. (2011). Assessment for learning in the accountability era: Empirical evidence from Scotland. *Studies in Educational Evaluation*, 37, 62-70.
- Irving, S., Harris, L., & Peterson, E. (2011). One assessment doesn't serve all the purposes or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review*, 12(3), 413-426.
- James, M., & Lewis, J. (2012). Assessment in harmony with our understanding of learning: Problems and possibilities. In J. Gardner (Ed.), *Assessment and learning* (2nd ed., pp. 187-205). London: Sage.

- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
doi:10.1080/09585170600792712
- Kagan, D. M. (1992). Implication of research on teacher belief. *Educational Psychologist*, 27(1), 65-90.
- Kahn, E. A. (2000). A case study of assessment in a grade 10 English course. *The Journal of Educational Research*, 93, 276-286.
- Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36.
- Karp, G. G., & Woods, M. L. (2008). Preservice teachers' perceptions about assessment and its implementation. *Journal of Teaching in Physical Education*, 27(3), 327-346.
- Kiely, M. T., Brownell, M. T., Lauterbach, A. A., & Benedict, A. E. (2015). Teachers' beliefs about students with special needs and inclusion. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 475-491). New York: Routledge.
- Klenowski, V., & Carter, M. (2016) Curriculum reform in testing and accountability contexts. In D. Wyse, L. Hayward & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy and assessment* (Vol. 2, pp. 790-804). London: Sage.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

- Lejeune, A., Poth, C., & Daniels, L. M. (2010). *Examining pre-service teachers' knowledge, skills, attributes, experiences, and goals related to formative and summative classroom assessment*. Paper Presented at the Canadian Society for the Study of Education, Montreal, Canada.
- Levin, B. (2015). The development of teachers' beliefs. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 48-65). New York: Routledge.
- Levin, B., & He, Y. (2008). Investigating the content and sources of teacher candidates' personal practical theories (PPTs). *Journal of Teacher Education*, 59(1), 55-68.
doi:10.1177/0022487107310749
- Levin, B. B., He, Y., & Allen, M. H. (2013). Teacher beliefs in action: A cross-sectional, longitudinal follow-up study of teachers' personal practical theories. *The Teacher Educator*, 48(3), 201-217. doi:10.1080/08878730.2013.796029
- Lieberman, A., & Miller, L. (2008). *Teachers in professional communities: Improving teaching and learning*. New York: Teachers College Press.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in Public Service*. New York: Russell Sage Foundation.
- Looney, A., Cumming, J., Van Der Kleij, F., & Harris, K. (2017). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in*

- Education: Principles, Policy & Practice*, 1-26.
doi:10.1080/0969594X.2016.1268090
- Lortie, D. (1975). *School teacher: A sociological perspective*. Chicago: University of Chicago.
- Lysaght, Z., & O'Leary, M. (2013). An instrument to audit teachers' use of assessment for learning. *Irish Educational Studies*, 32(2), 217-232.
doi:10.1080/03323315.2013.784636
- Mansour, N. (2009). Science teachers' beliefs and practices: Issues, implications and research agenda. *International Journal of Environmental & Science Education*, 4(1), 25-48.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133-149. doi:10.1080/02671520600615638
- McCormick, R., & Murphy, P. (2008). Curriculum: The case for a focus on learning. In P. Murphy, & K. Hall (Eds.), *Learning and practice: Agency and identities* (pp. 3 - 18). London: Sage.
- McMillan, J. H. (2001). *Classroom assessment: Principles and practice for effective instruction* (2nd ed.). Boston, MA: Allyn and Bacon.
- Murchan, Damian. (2017, May 27). Can a dual-currency solution deliver genuine Junior Cert reform. *The Irish Independent*. Retrieved from <http://www.independent.ie/irish-news/education/can-a-dualcurrency-solution-deliver-genuine-junior-cert-reform-31255854.html>

National Council for Curriculum and Assessment. (1999). *The Junior Cycle Review Progress Report: Issues and Options for Development*. Dublin: Author.

National Council for Curriculum and Assessment. (2010). *Innovation and Identity: Ideas for Junior Cycle*. Retrieved from
http://www.juniorcycle.ie/NCCA_JuniorCycle/media/NCCA/Curriculum/Research/Background/Innovation_and_Identity_Ideas_for_a_new_junior_cycle.pdf

National Council for Curriculum and Assessment. (2011). *Towards a Framework for Junior Cycle*. Dublin: Author. Retrieved from
http://www.juniorcycle.ie/NCCA_JuniorCycle/media/NCCA/Curriculum/Research/NCCA_TowardsaFramework_JuniorCycle.pdf

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19(4), 317-328. doi:10.1080/0022027870190403

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 149-170.
doi:10.1080/09695940701478321

Nusche, D. (2016). Student assessment and its relationship with curriculum, teaching and learning in the twenty-first century. In D. Wyse, L. Hayward & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy and assessment* (Vol. 2, pp. 838-852). London: Sage.

O’Keeffe, B. (2009, June 17). Speech by the Minister for Education and Science at the NCCA conference in the Hogan Stand Conference Centre, Croke Park, Dublin.

- Retrieved from <http://www.education.ie/en/Press-Events/Speeches/2009-Speeches/SP09-06-17.html>
- O'Sullivan, J. (2015). Foreword from the Minister. In Department of Education and Skills, *Framework for Junior Cycle*. Dublin: DES. Retrieved from <https://www.education.ie/en/Publications/Policy-Reports/Framework-for-Junior-Cycle-2015.pdf>
- Organisation for Economic Co-operation and Development. (2016). *Education at a glance 2016: OECD indicators*. London; Paris: Organization for Economic Cooperation & Development. doi:10.1787/eag-2016-en
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th ed.). Maidenhead: McGraw Hill Education.
- Panadero, E., Brown, G., & Courtney, M. (2014). Teachers' reasons for using self-assessment: A survey self-report of Spanish teachers. *Assessment in Education: Principles, Policy & Practice*, 21(4), 365-383.
doi:10.1080/0969594X.2014.919247
- Perrenoud, P. (1991). Toward a pragmatic approach to formative evaluation. In P. Weston (Ed.), *Assessment of pupils' achievement: Motivation and school success* (pp. 77-101). Amsterdam: Swets and Zeitlinger.

- Phipps, S., & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System*, 37(3), 380-390.
doi:<http://dx.doi.org.dcu.idm.oclc.org/10.1016/j.system.2009.03.002>
- Pintrich, P. (1990). Implications of psychological research on student learning and college teaching for teacher education. In W. R. Houston, M. Haberman & J. Sikula (Eds.), *Handbook of research on teacher education* (pp. 826-857). New York: Macmillan.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (6th ed.). Boston, MA: Allyn and Bacon.
- Pryor, J., & Crossouard, B. (2010). Challenging formative assessment: Disciplinary spaces and identities. *Assessment and Evaluation in Higher Education*, 35, 37-41.
- Reips, U. D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, 49(4), 243-256.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
doi:<http://dx.doi.org.remote.library.dcu.ie/10.1016/j.tate.2010.09.017>
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (pp. 102-119). New York: Simon & Schuster.
- Rieskamp, J., & Reimer, T. (2007). Ecological rationality. In R. F. Baumeister, & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (Vol. 1, pp. 273-275). Thousand Oaks, California: Sage Publications.

- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rokeach, M. (1968). *Beliefs, attitudes and values: A theory of organization and change*. London: Jossey-Bass.
- Rushton, G., Lotter, C., & Singer, J. (2011). Chemistry teachers' emerging expertise in inquiry teaching: The effect of a professional development model on beliefs and practice. *Journal of Science Teacher Education*, 22(1), 23-52. doi:10.1007/s10972-010-9224-x
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Schraw, G., & Olafson, L. (2015). Assessing teachers' beliefs: Challenges and solutions. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 87-105). New York: Routledge.
- Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. McLaughlin, & D. C. Phillips (Eds.), *Evaluation and education: At quarter century* (Part II, pp. 19-64). Chicago: NSSE.
- Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49-54. doi:<http://dx.doi.org/remote.library.dcu.ie/10.1016/j.stueduc.2011.03.008>
- Sellar, S., & Lingard, B. (2013). The OECD and global governance in education. *Journal of Education Policy*, 28(5), 710-725. doi:10.1080/02680939.2013.779791

- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. doi:10.3102/0013189X015002004
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-23. doi:10.17763/haer.57.1.j463w79r56455411
- Skott, J. (2009). Contextualising the notion of ‘belief enactment’. *Journal of Mathematics Teacher Education*, 12(1), 27-46. doi:10.1007/s10857-008-9093-9
- Skott, J. (2015). The promises, problems, and prospects of research on teachers’ beliefs. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers’ beliefs* (pp. 13-30). New York: Routledge.
- Smith, K. (2016). Assessment for learning: A pedagogical tool. In D. Wyse, L. Hayward & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy, and assessment* (Vol. 2, pp. 740-755). London: Sage.
- Smyth, E., Dunne, A., Darmody, M. & McCoy, S. (2006). *Pathways through the Junior Cycle*. Dublin: ESRI and NCCA in association with Liffey Press.
- Song, Y., & Looi, C. (2012). Linking teacher beliefs, practices and student inquiry-based learning in a CSCL environment: A tale of two teachers. *International Journal of Computer-Supported Collaborative Learning*, 7(1), 129-159. doi:10.1007/s11412-011-9133-9
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). New Jersey: Lawrence Erlbaum Associates Inc.
- Stiggins, R. J. (1991). Assessment literacy. *The Phi Delta Kappan*, 72(7), 534-539.

- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.
- Taber, K. S., Riga, F., Brindley, S., Winterbottom, M., Finney, J., & Fisher, L. G. (2011). Formative conceptions of assessment: Trainee teachers' thinking about assessment issues in English secondary schools. *Teacher Development*, 15(2), 171-186. doi:10.1080/13664530.2011.571500
- Tam, A. C. F. (2015a). Exploring teachers' beliefs about teacher learning in professional learning communities and their influence on collegial activities in two departments. *Compare: A Journal of Comparative and International Education*, 45(3), 422-444. doi:10.1080/03057925.2013.872025
- Tam, A. C. F. (2015b). The role of a professional learning community in teacher change: A perspective from beliefs and practices. *Teachers and Teaching*, 21(1), 22-43. doi:10.1080/13540602.2014.928122
- Teng, L. S. (2016). Changes in teachers' beliefs after a professional development project for teaching writing: Two Chinese cases. *Journal of Education for Teaching*, 42(1), 106-109. doi:10.1080/02607476.2015.1135228
- Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: Macmillan.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham, UK: Open University Press.
- Tschannen-Moran, M., Salloum, S. J., & Goddard, R. D. (2015). Context matters: The influence of collective beliefs and shared norms. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 301-316). New York: Routledge.
- Volante, L., & Earl, L. (2013). Standards, accountability, and student assessment systems: An international overview. *Education Canada*, 53(4). Retrieved from: <https://www.edcan.ca/articles/standards-accountability-and-student-assessment-systems/>
- Wallace, C. S., & Priestley, M. (2011). Teacher beliefs and the mediation of curriculum innovation in Scotland: A socio-cultural perspective on professional development and change. *Journal of Curriculum Studies*, 43(3), 357-381.
doi:10.1080/00220272.2011.563447
- Wang, J., Kao, H., & Lin, S. (2010). Preservice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. *Teaching and Teacher Education*, 26(3), 522-529.
doi:<http://dx.doi.org.dcu.idm.oclc.org/10.1016/j.tate.2009.06.014>
- Warwick, P., Shaw, S., & Johnson, M. (2015). Assessment for learning in international contexts: Exploring shared and divergent dimensions in teacher values and practices. *The Curriculum Journal*, 26(1), 39-69.
doi:10.1080/09585176.2014.975732

- Watkins, M. W. (2006). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*, 5(2), 344-346.
- Winter, C. (2012). School curriculum, globalisation and the constitution of policy problems and solutions. *Journal of Education Policy*, 27(3), 295-314.
doi:10.1080/02680939.2011.609911
- Winterbottom, M., Brindley, S., Taber, K. S., Fisher, L. G., Finney, J., & Riga, F. (2008). Conceptions of assessment: Trainee teachers' practice and values. *The Curriculum Journal*, 19(3), 193-213. doi:10.1080/09585170802357504
- Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). Chapter 2: To use their minds well: Investigating new forms of student assessment. *Review of Research in Education*, 17(1), 31-74. doi:10.3102/0091732X017001031
- Wyse, D., Hayward, L., & Pandya, J. (Eds.) (2016). *The SAGE handbook of curriculum, pedagogy and assessment* (Vol. 1). London: Sage.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.
doi:<http://doi.org.dcu.idm.oclc.org/10.1016/j.tate.2016.05.010>
- Yan, Z. (2014). Predicting teachers' intentions to implement school-based assessment using the theory of planned behaviour. *Educational Research and Evaluation*, 20(2), 83-97. doi:10.1080/13803611.2013.877394
- Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45, 128-136.
doi:<http://dx.doi.org.remote.library.dcu.ie/10.1016/j.tate.2014.10.002>

- Yates, L., & Young, M. (2010). Globalisation, knowledge and the curriculum Wiley-Blackwell. *European Journal of Education*. 45(1), 4-10. doi:10.1111/j.1465-3435.2009.01412.x
- Yates, A., & Johnston, M. (2017). The impact of school-based assessment for qualifications on teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 1-17. doi:10.1080/0969594X.2017.1295020
- Yuan, R., & Lee, I. (2014). Pre-service teachers' changing beliefs in the teaching practicum: Three cases in an EFL context. *System*, 44, 1-12.
doi:<http://dx.doi.org.dcu.idm.oclc.org/10.1016/j.system.2014.02.002>

Appendices

Appendix A: Correspondence from the Teaching Council

----- Forwarded message -----

From: **Maria Fitzgerald** <mfitzgerald@teachingcouncil.ie>

Date: 13 June 2017 at 16:17

Subject: FW: FW: Doctoral Research Request for National Statistics

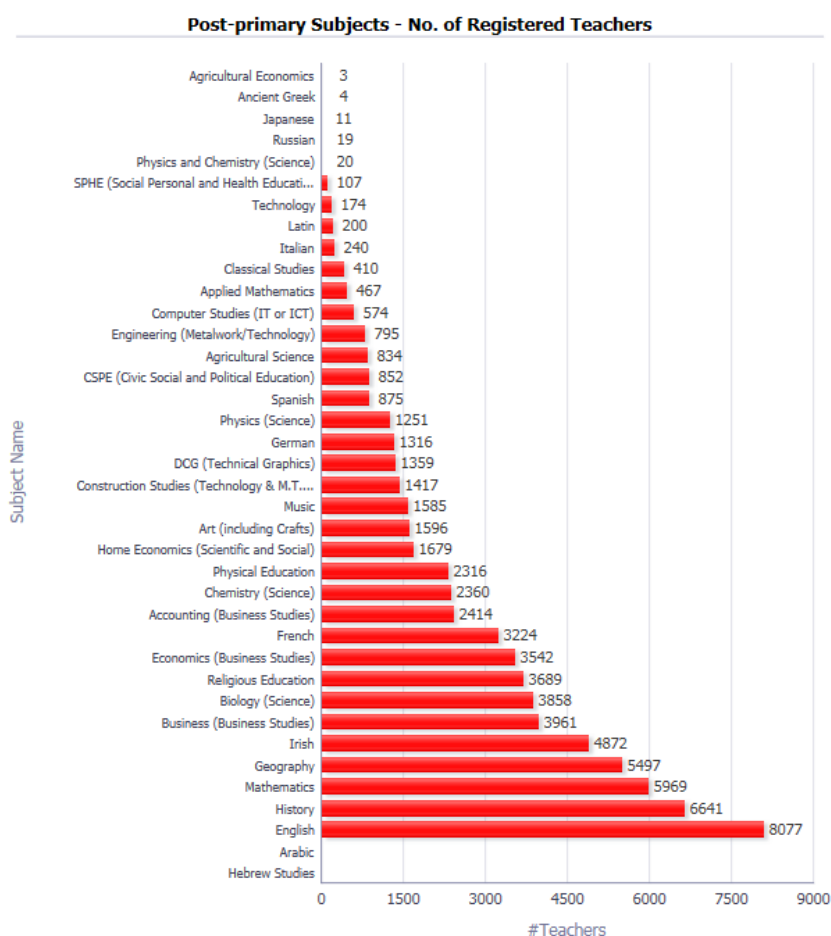
To: Marie Darmody <marie.darmody3@mail.dcu.ie>

Hi Marie,

Please see the information below. This reflects the number of teachers on the register with the subjects attached. Please note teachers may be registered with more than one subject attached. Also it is important to note that many teachers may be teaching subjects for which they are not registered but have studied at undergraduate level (they may not have adequate ECTS in order to register for that subject).

Regards

Maria



From: **Elaine McDermott** <emcdermott@teachingcouncil.ie>
Date: 28 February 2017 at 16:26
Subject: Doctoral research request for national statistics [STATISTICS]
To: "marie.darmody3@mail.dcu.ie" <marie.darmody3@mail.dcu.ie>

Dear Marie

Thank you for your email enquiry below.

The Council is the regulatory body for teachers in Ireland and holds a register of teachers in Ireland in the primary, post-primary and further education sector. I will be able to assist you in answering some questions below, however the Council has not remit in the areas of employment, this falls under the school as their employer and the Department of Education and Skills. For ease of reference I have listed your queries below in the order raised:

- a) the number of post-primary teachers currently registered as classroom teachers*
- b) the number of post-primary teachers currently registered as principals*
- c) the number of post-primary teachers currently registered as deputy principals*

The Council holds information in relation to the number of teachers that are registered in each sector i.e. Primary, Post-Primary and Further Education. The query you have raised specifically crosses over into the employment area for which we would not have the details. Teachers are not obliged to confirm their employment details with the Council and many teachers such as substitutes change their employment on a regular basis. For your information, I have provided the sector breakdown for the Register as it stands on 21 February 2017:

Sector	Number of Teachers
Primary	44,575
Post-Primary	42,542
Further Education	9,486
Other (includes Montessori)	1,228

- d) the number of male post-primary teachers*

The number of male post-primary teachers on the Register currently stand at 13,298

- e) the number of female post-primary teachers*

The number of female post-primary teachers on the Register currently stand at 29,244

f) the number of post-primary teachers who hold a master's degree or higher
Approximately 2562 post primary teachers hold masters degrees or higher. Please note here whilst the Council encourages teachers to update their registration with additional qualifications, some teachers choose not to record additional qualifications.

I hope this information is of assistance to you.

Kind Regards

Elaine McDermott
Higher Executive Officer
Registration Section

The Teaching Council | Block A, Maynooth Business Campus | Maynooth | Co.
Kildare

T: 01 6517900

F: 01 6517901

W: www.teachingcouncil.ie

Follow us on [@TeachingCouncil](https://twitter.com/TeachingCouncil) and [@ FÉILTE](https://twitter.com/FÉILTE)

Appendix B: Example Page of Online Survey Instrument

Irish Post-Primary Teachers' Conceptions of Assessment

33 %

Conceptions of Assessment

Please respond to each statement by choosing the response that best reflects your beliefs on assessment, whatever that term means to you. Please note: you are asked to focus on your assessment beliefs not your assessment practices. Further, you need to provide a response to each statement in order to progress through the survey.

1. Assessment provides information on how well schools are doing *

- ☐ Strongly Disagree
- ☐ Mostly Disagree
- ☐ Slightly Agree
- ☐ Moderately Agree
- ☐ Mostly Agree
- ☐ Strongly Agree

2. Assessment places students into categories *

- ☐ Strongly Disagree
- ☐ Mostly Disagree
- ☐ Slightly Agree
- ☐ Moderately Agree
- ☐ Mostly Agree
- ☐ Strongly Agree

3. Assessment is a way to determine how much students have learned from teaching *

- ☐ Strongly Disagree
- ☐ Mostly Disagree
- ☐ Slightly Agree
- ☐ Moderately Agree
- ☐ Mostly Agree
- ☐ Strongly Agree

Appendix C: Feedback Sheet

Questionnaire Feedback

1. Approximately how long did it take you to complete the survey and the biographical data?

2. Was there any aspect of the survey that was unclear?

3. Did you find the survey to be user-friendly?

4. Any further comments?

Appendix D: Correspondence from Professor Gavin Brown

From: **Gavin Brown** <gt.brown@auckland.ac.nz>
Date: 18 April 2015 at 02:40
Subject: RE: TCoA-III inventory
To: Marie Darmody <marie.darmody3@mail.dcu.ie>

Hi Marie

There is no problem in your using the TCoA in your doctoral research

I am attaching the long and short versions for your consideration. The spread sheet shows the relationship of items to factors.

My only requests are that (a) you send me any adaptations you make and (b) you send me a copy of your research results using the inventory if you use it once the thesis or reports are approved for release.

You may want to consider factors or items from the variation inventories used in the Chinese and Indian studies, especially if your policy context has high stakes testing like UK or USA

- Brown, G. T. L., Chaudhry, H., & Dhamija, R. (2015). The impact of an assessment policy upon teachers' self-reported assessment beliefs and practices: A quasi-experimental study of Indian teachers in private schools. *International Journal of Educational Research*, 71, 50-64.doi: 10.1016/j.ijer.2015.03.001
- Brown, G. T. L., Hui, S. K. F., Yu, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320.doi: 10.1016/j.ijer.2011.10.003

Best wishes with your studies
sincerely

Gavin T L Brown, PhD
January to June 2015
Visiting Associate Professor The University of Hong Kong
Office: HOC 105; Telephone no.: 2219-4841
Email Address: gtlbrown@hku.hk

Director Quantitative Data Analysis and Research Unit (Quant-DARE)
School of Learning, Development & Professional Practice
Faculty of Education, The University of Auckland
Private Bag 92019, Auckland, 1142, New Zealand
Tel: [+649 623 8899](tel:+6496238899) ext. 48602; Fax: [+649 623 8827](tel:+6496238827)
<http://www.education.auckland.ac.nz/uoa/gavin-brown>

Appendix E: Plain Language Statement



COLÁISTE PHÁDRAIG
ST PATRICK'S COLLEGE
DROIM CONRACH | DRUMCONDRA

Plain Language Statement

Irish Post-Primary Teachers' Conceptions of Assessment is a study by Marie Darmody under the supervision of Dr. Zita Lysaght and Professor Michael O'Leary. The research is being conducted in partial fulfilment of the Doctorate in Education at St. Patrick's College, Drumcondra and is financially supported under the NCCA's Assessment Research and Development Programme 2015/16. The purpose of this research is to elicit baseline data about Irish post-primary teachers' conceptions of assessment. Post-primary education in Ireland is in the midst of much curriculum and assessment change and teachers are at the coalface of this change. With the proposed introduction of new forms of classroom assessment, teachers will have an increased role in the assessment process. However, teachers' beliefs provide a lens for the translation of policy into practice and so when considering the implementation of assessment reform one needs to know what teachers believe about the nature and purpose of assessment.

Involvement in this research study will require you to provide some background information and to complete an anonymous survey regarding your beliefs about assessment. You will be asked to respond to twenty-seven statements and to indicate your level of agreement with each one on a scale ranging from strongly disagree to strongly agree. You will also be asked to indicate whether you would or would not be willing to take part in a focus group discussion on this topic which may take place as part of a future study.

Involvement in a study such as this presents a unique opportunity to reflect upon your own beliefs and assumptions about assessment. It is hoped that this study will contribute significantly to national understanding of assessment at post-primary level.

Participants who complete the survey will not be identified or named in any way. While the researcher cannot offer anonymity to those opting in to a potential future focus group, every effort will be made to ensure confidentiality.

Data will only be used for the purposes of the research as well as possible future conferences and publications.

Involvement in this research study is voluntary. Participants may withdraw from the study at any point. It must be noted, however, that given the anonymous nature of the survey, it will not be possible to extract or omit data from this source should a participant elect to withdraw at some point in the study.

Please note the status bar displayed on the upper right-hand side of your screen indicates your progress thus far through the survey.

If you have any questions that need clarification, please do not hesitate to ask. Your co-operation with this research project is much appreciated.

Marie Darmody
marie.darmody3@mail.dcu.ie

If participants have concerns about this study and wish to contact an independent person, please contact:

REC Administration,
Research Office,
St Patrick's College,
Drumcondra,
Dublin 9.
Tel +353-(0)1-884 2149
research@spd.dcu.ie

Appendix F: Informed Consent Form



COLÁISTE PHÁDRAIG
ST PATRICK'S COLLEGE
DROIM CONRACH | DRUMCONDRA

Informed Consent Form

Irish Post-Primary Teachers' Conceptions of Assessment is a study by Marie Darmody under the supervision of Dr. Zita Lysaght and Professor Michael O'Leary. The research is being conducted in partial fulfilment of the Doctorate in Education at St. Patrick's College, Drumcondra. The purpose of this research is to elicit base-line data about Irish post-primary teachers' conceptions of assessment.

Involvement in this research study will require you to complete an anonymous online questionnaire regarding your beliefs about assessment.

A potential follow-up focus group on this topic may take place as part of a future study. Please indicate below (by circling yes or no) whether you would or would not be willing to partake in a focus group discussion:

Would you be willing to partake in a focus group discussion? Yes / No

If yes, please provide an email address:

I am aware that if I agree to take part in this study, I can withdraw from participation at any stage. There will be no penalty for withdrawing before all stages of the research have been completed.

Please complete the following. **Circle Yes or No in the case of each question.**

Have you read the Plain Language Statement?	Yes / No
Do you understand the information provided?	Yes / No

Have you had an opportunity to ask questions and discuss this study?	Yes / No
Have you received satisfactory answers to all your questions?	Yes / No
Do you understand that the researcher cannot extract or omit data provided in the questionnaire should you elect to withdraw from the study at some point?	Yes / No
Do you understand that the researcher cannot offer anonymity to those opting in to a focus group, but will ensure, as far as possible, confidentiality?	Yes / No

I have read and understood the information in this form. The researcher has answered my questions and concerns, and I have a copy of this consent form. Therefore, I consent to take part in this research project.

Appendix G: Range of Subjects Taught by Participants

Subject	Participants (n=489)	% of total sample	Population Value (%)
Accounting	24	4.9	5.7
Agricultural Science	5	1	2
Applied Maths	6	1.2	1.1
Art	15	3.1	3.8
Biology	39	8	9.1
Business	65	13.3	9.3
Chemistry	22	4.5	5.5
CSPE (Civic, Social & Political Education)	96	19.6	2
Classical Studies	3	0.6	1
Construction Studies	9	1.8	3.3
Design and Comm Graphics	15	3.1	3.2
Economics	21	4.3	8.3
Engineering	10	2	1.9
English	106	21.7	19
Environmental and Social Studies	3	0.6	–
French	57	11.7	7.6
Geography	64	13.1	12.9
German	26	5.3	3.1
History	76	15.5	15.6
Home Economics	23	4.7	3.9
Irish	96	19.6	11.5
Italian	3	0.6	0.56
Japanese	2	0.4	0.03
Materials Technology	9	1.8	–
Mathematics	120	24.5	14
Metalwork	8	1.6	–
Music	17	3.5	3.7
Physics	18	3.7	2.9
Physics/Chemistry	2	0.4	0.05
Religious Education	67	13.7	8.7
Russian	1	0.2	0.04
Science	64	13.1	–
Spanish	17	3.5	2.1
Technical Graphics	15	3.1	–
Technology	15	3.1	0.4
Typewriting	1	0.2	–
Physical Education	16	3.3	5.4
SPHE (Social, Personal & Health Education)	31	6.3	0.25
Other	61	12.5	–

Note. Population values represented by “–” were not available to the Researcher.

Appendix H: Mean and Standard Deviation for each statement on the TCoA-III A

	N	Min.	Max.	Mean	Std. Dev.
Assessment provides feedback to students about their performance	489	1	6	4.55	1.297
Assessment is integrated with teaching practice	489	1	6	4.52	1.352
Assessment is a way to determine how much students have learned from teaching	489	1	6	4.41	1.255
Assessment information modifies ongoing teaching of students	489	1	6	4.30	1.144
Assessment places students into categories	489	1	6	4.21	1.330
Assessment establishes what students have learned	489	1	6	4.19	1.107
Assessment helps students improve their learning	489	1	6	4.15	1.176
Assessment feeds back to students their learning needs	489	1	6	4.03	1.247
Teachers should take into account the error and imprecision in all assessment	489	1	6	3.97	1.293
Assessment determines if students meet qualifications standards	489	1	6	3.83	1.284
Assessment results are trustworthy	489	1	6	3.75	1.139
Assessment results should be treated cautiously because of measurement error	489	1	6	3.61	1.225
Assessment results can be depended upon	489	1	6	3.60	1.155
Assessment provides information on how well schools are doing	489	1	6	3.49	1.288
Assessment measures students' higher order thinking skills	489	1	6	3.38	1.276
Assessment allows different students to get different instruction	489	1	6	3.36	1.374
Assessment results are consistent	489	1	6	3.30	1.206
Assessment is assigning a grade or level to student work	489	1	6	3.16	1.448
Assessment forces teachers to teach against their beliefs	489	1	6	3.09	1.398
Assessment is an imprecise process	489	1	6	3.01	1.184
Teachers conduct assessments but make little use of the results	489	1	6	2.87	1.266
Assessment results are filed and ignored	489	1	6	2.53	1.240
Assessment interferes with teaching	489	1	6	2.52	1.285
Assessment is an accurate indicator of a school's quality	489	1	6	2.47	1.221
Assessment is a good way to evaluate a school	489	1	6	2.46	1.238
Assessment is unfair to students	489	1	6	2.14	.968
Assessment has little impact on teaching	489	1	6	2.09	1.063
Valid N (listwise)	489				