

Exploring the Impact of Standardised Assessment in the Primary School Classroom

**Michael Mc Namara
B.Ed., PD PES.**

**Thesis submitted to the National University of Ireland, Galway
in partial fulfilment of the requirements for the degree of
Master of Education**

**School: School of Education
Head of School: Dr. Tony Hall
Supervisor: Dr. Veronica Mc Cauley**

August 2010

Dedicated to:

My parents,

Maura and Michael Mc Namara,

for all their encouragement and support.

Abstract

Standardised testing has become a widely-debated issue in recent international forums, yet research into this form of assessment has been limited from the Irish perspective. The purpose of this study was to investigate teachers' attitudes towards standardised testing, thereby gaining a valuable insight into the impacts of this form of assessment in the Irish primary school classroom. This exploration also allowed the researcher to locate the development of standardised testing in Ireland on an international scale, comparing findings in the Irish context with those derived from the international perspective.

The research undertaken involved a small-scale survey of 30 teachers from 15 primary schools in Galway City and County, as well as conducting a semi-structured interview with a Department of Education Inspector. The data collected was analysed thematically using some of the techniques associated with the Grounded Theory approach to data analysis. It was found that the numerous benefits that were experienced by teachers when utilising standardised testing contributed to the overall positive attitude towards this form of assessment. However, the data collected also indicated that the majority of teachers experienced pressure when using standardised assessments from a number of internal and external sources. Furthermore, it was found that, in many cases, this pressure resulted in teachers preparing their students for the tests in a variety of manners, both ethical and unethical. Subsequently, the data collected indicated that this test preparation impacted on classroom practice, as the time spent preparing for the standardised tests was time that could have been spent engaging in more meaningful forms of learning.

The findings of this study strongly indicated that teachers require support in this area as the data collected suggested that teachers were self-constructing their own 'best practice' guidelines autonomously, resulting in the many diverse approaches currently undertaken when implementing standardised testing in the Irish primary school context.

Acknowledgements

There are many people who contributed to the production of this dissertation through their encouragement, support and participation.

I would like to express my sincere gratitude to all those who participated in this study, from the teachers who filled out the questionnaires to the Inspector who gave so generously of their time.

I would like to express my sincere appreciation to Dr. Veronica Mc Cauley, my supervisor, for all her guidance, encouragement, time and support throughout this process.

I also wish to thank all the staff of the School of Education for their encouragement and support throughout the Masters of Education programme, and in particular Bonnie Long for all her time and assistance.

I also wish to acknowledge and express my sincere gratitude for the research bursary received from the Irish National Teachers' Organisation which certainly aided in conducting this research.

Finally, I wish to thank my family and friends for all their understanding and moral support throughout the course of this journey.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents	iv
List of Tables	viii
List of Figures.....	ix
List of Acronyms	x
Chapter 1: Introduction	1
1.1 Project Aim	1
1.2 Rationale Behind Project Choice	2
1.3 Professional Context	4
1.4 Structure of the dissertation	5
Chapter 2: Literature Review	7
2.0 Introduction	7
2.1 Defining Standardised Assessment	8
2.2 The Key Benefits Associated with Standardised Testing	9
2.3 The Irish Context.....	11
2.3.1 <i>Recurring themes in Irish Literature</i>	11
2.3.2 <i>The Theme of 'Accountability'</i>	11
2.3.3 <i>The Theme of 'Teaching to the Test'</i>	13
2.3.4 <i>The Theme of 'Impact on Classroom Practice'</i>	15
2.4 The International Context	16
2.4.1 <i>Recurring Themes in International Literature</i>	16
2.4.2 <i>The Theme of 'Accountability'</i>	17
2.4.3 <i>The Theme of 'Teaching to the Test'</i>	19
2.4.4 <i>The Theme of 'Impact on Classroom Practice'</i>	21
2.5 Conclusion	25
Chapter 3: Methodology.....	27
3.1 Introduction	27
3.2 The Research Design.....	27
3.3 Locating the Research within a Paradigm.....	28
3.4 The Research Strategy.....	29
3.5 Research Methods Employed.....	31

3.5.1	<i>Self-Completion Qualitative Questionnaires as a Method of Research..</i>	31
3.5.2	<i>Semi-Structured Interview as a Method of Research.....</i>	33
3.6	The Sampling Process	35
3.7	Ethical Issues.....	37
3.8	The Research Instruments: From Inception to Analysis	39
3.8.1	<i>The Self-Administered Questionnaire</i>	39
3.8.2	<i>The Interview Schedule</i>	46
Chapter 4: Data Findings, Analysis and Discussion		51
4.1	Introduction	51
4.2	Grounded Theory Approach to Data Analysis.....	53
4.3	Standardised Testing at Primary Level	55
4.3.1	<i>Standardised Test Instruments Selected.....</i>	55
4.3.2	<i>Perceived Effectiveness of Standardised Testing.....</i>	56
4.4	Theme 1: Benefits Associated with Standardised Testing.....	57
4.4.1	<i>Indicating Achievement.....</i>	58
4.4.2	<i>National Comparisons</i>	59
4.4.3	<i>Diagnostic/ Screening</i>	60
4.4.4	<i>Providing Evidence for Parents</i>	61
4.4.5	<i>Other Noteworthy Benefits Identified.....</i>	61
4.5	Theme 2: Accountability.....	62
4.5.1	<i>The Presence of Pressure.....</i>	62
4.5.2	<i>Parental Pressure</i>	63
4.5.3	<i>Reflection of Own Teaching</i>	65
4.5.4	<i>Peer-Teacher Pressure.....</i>	66
4.5.5	<i>Curriculum Overload.....</i>	67
4.5.6	<i>Educational Management</i>	68
4.6	Theme 3: Preparation	70
4.6.1	<i>Time Spent Preparing for the Tests:</i>	70
4.6.2	<i>Procedural Instruction.....</i>	72
4.6.3	<i>Directed Revision.....</i>	73
4.6.4	<i>General Revision.....</i>	74
4.7	Theme 4: Classroom Practice	75
4.7.1	<i>Ascertaining if Classroom Practice is Affected</i>	75
4.7.2	<i>No Impact Identified.....</i>	76
4.7.3	<i>Pre-Testing Impacts</i>	78

4.7.4	<i>Post-Testing Impacts</i>	79
4.8	Conclusion	81
Chapter 5: Conclusions and Recommendations		82
5.1	Conclusions	82
5.1.1	<i>Standardised Testing in Ireland</i>	82
5.1.2	<i>Theme 1: The Benefits Associated with Standardised Testing</i>	83
5.1.3	<i>Theme 2: Accountability</i>	84
5.1.4	<i>Theme 3: Preparation</i>	86
5.1.5	<i>Theme 4: Impact on Classroom Practice</i>	88
5.2	Recommendations	90
5.2.1	<i>Standardised Testing In Ireland</i>	90
5.2.2	<i>Theme 1: Benefits Associated with Standardised Testing</i>	91
5.2.3	<i>Theme 2: Accountability</i>	92
5.2.4	<i>Theme 3: Preparation</i>	94
5.2.5	<i>Theme 4: Impact on Classroom Practice</i>	95
5.3	Further Research	96
References		99

Appendices:

Appendix 1: The Development of Standardised Testing in the Irish Context ...	Error! Bookmark not defined.
Appendix 2: The Development of Standardised Testing in the International Context.	Error! Bookmark not defined.
Appendix 3: Letter to Principal	Error! Bookmark not defined.
Appendix 4: Letter to Teachers	Error! Bookmark not defined.
Appendix 5: Consent Form	Error! Bookmark not defined.
Appendix 6: Pilot Questionnaire	Error! Bookmark not defined.
Appendix 7: Final Questionnaire	Error! Bookmark not defined.
Appendix 8: Codes Used when Analysing Data	Error! Bookmark not defined.
Appendix 9: Sample Category	Error! Bookmark not defined.
Appendix 10: Interview Schedule	Error! Bookmark not defined.
Appendix 11: Layout of Interview Transcription	Error! Bookmark not defined.

(A separate ‘Dossier of Evidence’ has been constructed to accompany this dissertation, containing both the questionnaire and interview transcriptions)

List of Tables

Table 1.1: Finalised Research Questions.....	1
Table 2.1: Initial Research Questions	26
Table 3.1: Relevance of Each Question on the Pilot Questionnaire	39
Table 3.2: Outline of Final Questionnaire Research Instrument.....	44
Table 3.3: Questions contained on the Pilot Interview Schedule.....	47
Table 4.1: Outlining the Manner in which Each Research Question was Addressed.....	51

List of Figures

Figure 4.1: Test Instruments Used	55
Figure 4.2: Perceived Effectiveness of Standardised Testing.....	56
Figure 4.3: Benefits Associated with Standardised Testing.....	58
Figure 4.4: Origins of Pressure Associated with Standardised Assessment	63
Figure 4.5: Time Spent Preparing for Tests	70
Figure 4.6: Activities Undertaken in Preparation for the Test	71
Figure 4.7: The Perceived Impact of Standardised Testing on Classroom Practice	76

List of Acronyms

AoL	Assessment of Learning
AfL	Assessment for Learning
BERA	British Educational Research Association
ERC	Educational Research Centre
INTO	Irish National Teachers' Organisation
NCCA	National Council for Curriculum and Assessment
NCLB	'No Child Left Behind' Act, 2002

Chapter 1: Introduction

1.1 Project Aim

The main purpose of this dissertation is to explore teachers' attitudes towards standardised testing, thereby ascertaining the impacts of this form of assessment in the primary school classroom. To achieve this aim, a survey approach was adopted whereby teachers' views were obtained using a qualitative questionnaire, and were supplemented with the information acquired from interviewing a Department of Education Inspector. The data collected and analysed was then informed and contextualised with literature relating to the impacts of this form of assessment. This process aimed to identify key areas, later embodied in the research questions, that would benefit from further research. Thus, to achieve this aim the research questions outlined in table 1.1 below were carefully devised, refined and explored throughout the research process in order to effectively determine the impacts of standardised testing in the primary school classroom:

Table 1.1: Finalised Research Questions

Questions:	Question Content:
Question 1	What are the benefits associated with standardised testing?
Question 2	How does accountability affect teachers' attitudes towards standardised testing?
Question 3	In what ways are children prepared for standardised testing?
Question 4	What are the impacts on classroom practice as a result of the adoption of standardised testing in our schools?

Each area highlighted in the research questions above was thus explored in this study in order to identify the benefits associated with this form of assessment, as well as allowing the researcher the opportunity to discover where improvements could be

made. By undertaking such an approach, this study aims to enhance the contribution that standardised testing can make within the Irish education system, thus supporting teachers' professional practice in the primary school classroom.

1.2 Rationale Behind Project Choice

In a world where measurement, evaluation and accountability are becoming an ever-increasing part of public life, the 'measurement' of education comes in for increasing scrutiny and the object of growing debate.

(NCCA, 2005, p.12)

When one considers the on-going debate surrounding standardised testing, both within the Irish context and internationally, the sentiment embodied in the above statement certainly rings true. As assessment is seen as a powerful tool in the 'measurement' of education, as alluded to above, it is vital that the procedures governing this activity are themselves assessed, ensuring that such practices serve to enhance the quality of education in our primary schools, not undermine it. It is with this objective in mind that it was deemed vital to explore the impacts created by standardised testing from the Irish primary school perspective. As this form of assessment is universally utilised in primary schools nationwide, it should thus be monitored to ensure that it contributes to a teacher's professional practice in a constructive manner.

Discussion on the merits of standardised testing has undoubtedly grown in recent years as a result of this form of assessment becoming a compulsory feature in our primary school education system from 2007 (Ireland. Department of Education and Science, 2006). Thus, a further underlying principle for undertaking this study was the researcher's belief that any such recent innovations should be judiciously monitored and their effectiveness ascertained at regular intervals. This process serves to ensure that all aspects of our current education system serve to meaningfully influence the realisation of the aims and objectives embodied in the revised primary school curriculum (Government of Ireland, 1999). Thus, the findings of this study can be seen as a timely appraisal of this relatively recent addition to our education

system, providing a valuable insight into the current state of standardised testing by those most experienced in utilising this form of assessment.

Furthermore, worrying results have emerged in international literature based on standardised testing. Here, unintentional negative impacts have been identified which can manifest as a result of the misuse of this form of assessment. One such study in the United States advocates that “standardised testing has swelled and mutated, like a creature in one of those old horror movies, to the point that it now threatens to swallow our schools whole” (2000, p.1). Thus, one cannot but agree that we must dutifully look to our own context to identify if such maltreatment of standardised assessment is occurring, and explore ways of negating any undesirable impacts. This approach aims to “improve educational policy and practice, by informing pedagogic, curricular and other educational judgements and decisions” (BERA, 2003, p.2). Therefore, it is with this objective in mind that the research questions, as previously outlined, were strategically constructed to explore key areas relating to standardised testing from the perspective of both mainstream primary school teachers and the Department of Education Inspector.

Also, this dissertation can be viewed as a contributing factor that aims to aid in the successful implementation of this form of assessment in our national education system. Here, conclusions and recommendations noted in this study can be seen to support and develop the current practices that define standardised testing, promoting the effective use of this form of assessment, and thereby enhancing teaching and learning in the primary school classroom. As the research focuses on teachers’ classroom practices, as well as their socially and/or self-constructed attitudes towards this form of assessment, this study takes a practical look at standardised testing in our current educational climate and highlights areas that require attention. Therefore, this study can be viewed as a catalyst in promoting the initiation of further research on standardised assessment, challenging the passive acceptance this form of assessment has enjoyed to date, and ultimately aiming to introduce positive changes in the primary school curriculum (Government of Ireland, 1999) which is itself considered fluid in nature, open to revisions and enhancements.

Furthermore, this research may prove useful for regulatory bodies, such as the National Council for Curriculum and Assessment (NCCA), when devising guidelines and policies in relation to standardised assessment. As both the NCCA (2005) and the Irish National Teachers' Organisation (INTO) (2010) have noted, it is timely that an evaluation of standardised testing within our education system be carried out. This evaluation will undoubtedly result in the formulation of policy in this regard as currently there is a lack in the literature available aimed at supporting standardised testing in the Irish primary school context. Here, the findings of this study would certainly add a valuable perspective for those undertaking such an evaluation. Likewise, the recommendations proposed in this dissertation embody practical steps that could be undertaken to effectively support teachers when utilising standardised testing in their classrooms, ensuring that the assessments are being used appropriately, thereby maximising the benefits experienced by teachers in this regard.

1.3 Professional Context

Since completing my Bachelor of Education degree in Mary Immaculate College, Limerick in 2004, I have been working as a mainstream teacher, gaining experience in both single-grade and multi-grade contexts. Currently, I am teaching fifth class in a medium-sized school in County Galway. As it has been designated that the standardised test results of the children in my class are to be sent home to parents annually, this topic has become of particular interest to me in recent years. Together with my prior observations relating to standardised testing, this aspect has contributed in forming my initial interest in the area of compulsory assessment in primary schools.

The personal experience gained as a result of utilising standardised testing over the past number of years has also developed my keen interest in this area. Some of the pressures and unintentional by-products associated with this form of assessment, as outlined later in this study, were personally experienced within my own context, thus further inspiring to me to undertake research in this area to ascertain the extent to which such features were experienced by other practitioners. Furthermore, when

speaking with other professionals about such features in the past, it became apparent that teachers have strong unvoiced opinions on this topic. I found it interesting that such opinions were socially and/or self constructed, and thought it would be fascinating to delve into the reasons behind the formulation of such views. This aim, stemming from past observations, has certainly contributed to the undertaking of this current study.

Furthermore, my professional context has also allowed me to become a reflective practitioner, identifying areas of our current education system that have become unconstructively diluted. Here, owing to my experience utilising standardised assessment in a number of educational sites, I became aware that the practices defining this activity were not undertaken in a uniform manner by all teachers. The varying extent to which teachers employed ethical and unethical practices in this regard was duly noted, and I began to consider ways in which this situation could be appropriately addressed, an interest that has also certainly contributed in the formation of this dissertation.

1.4 Structure of the dissertation

This dissertation is divided into five chapters.

Chapter 1: Introduction

This chapter outlines the aims of the dissertation project whereby the researcher clearly indicates the objectives that this study aspires to achieve. A rationale is also provided which validates the reasons behind conducting a study of this nature. Furthermore, the researcher dutifully describes the professional context from which he derives, showing the reader how personal motivations stemming from experience in this area have influenced the formation of this study. The structure of this dissertation is also outlined when drawing this chapter to a close.

Chapter 2: Literature Review

This chapter aims to explore a variety of literature relevant to the topic of standardised testing. This form of assessment is clearly defined and the benefits associated with standardised assessment documented in the literature are examined.

Prevalent themes are then explored from both the Irish and international perspective, allowing the researcher to ascertain the extent to which such features exist in our own context.

Chapter 3: Methodology

This chapter outlines the research methodology used to conduct this study. The reasons validating the adoption of a survey approach are discussed, as well as commenting upon the appropriateness of utilising a self-completion qualitative questionnaire and semi-structured expert interview to achieve this aim. Ethical considerations are also dutifully addressed and discussed in this chapter. Furthermore, the process of data collection is comprehensively described, allowing the reader to gain an appreciation of the careful planning that such procedures entail.

Chapter 4: Findings, Analysis and Discussion

The findings of this study, derived from careful application of the Grounded Theory approach to data analysis (Glaser and Strauss, 1967), are clearly outlined within the confines of this chapter. Logically, discussion based on these findings ensues, whereby the researcher explores the reasons behind the presence of such features in the data, as well as examining the relationship of these features to those identified in relevant literature devised on the topic of standardised testing.

Chapter 5: Conclusions and Recommendations

The main conclusions derived from the data collected and analysed are outlined in this chapter. Furthermore, based upon these conclusions, informed recommendations are also posed in this chapter that aim to maximise the potential contribution that standardised assessment can make in the primary school context. Areas in which future research could be carried out to further improve this area of education are also proposed in this chapter.

Chapter 2: Literature Review

2.0 Introduction

This chapter provides a broad overview of standardised assessment, exploring the associated benefits as well as the often unintended by-products frequently documented in both the Irish and international literature pertaining to this form of assessment. In this manner, “Ireland can learn from the developments elsewhere in foraging an approach to assessment that meets the needs of Irish students and schools and the Irish education system” (NCCA, 2005, p.18).

In order to effectively explore the aforementioned features, this chapter has been strategically constructed using clear and logical sections. Section one focuses on outlining a clear definition of standardised testing. Section two then explores the widely-documented benefits associated with standardised assessment usage. Section three examines the Irish perspective in relation to standardised assessment. Here, reoccurring themes in the limited body of Irish literature are explored in order to ascertain the key impacts standardised testing exerts upon our education system. The literature in this section then parallels that presented in section four which focuses on reoccurring themes from the international perspective in relation to standardised testing, allowing us to place the features present in our own education system within a wider international context.

When exploring the international context it was decided to focus attention on the body of literature deriving from both the United States and Great Britain. This approach was strategically adopted as it is these two countries that have most strongly influenced Irish advancements in standardised assessment, largely due to their close proximity to Ireland. This close relationship has resulted in the frequent exchange of educational innovation between personnel in these countries and our own, a commonly held belief also featured in the limited number of Irish studies in this area (Madaus, Airasian and Kellaghan, 1971; Kellaghan and Fontes, 1989; INTO, 1997; NCCA, 2005; Mac Ruairc, 2009). Furthermore, it is from these two countries that most of the literature generated on standardised testing has originated, a point reiterated by Madaus, Airasian and Kellaghan who note that “since about

1960 the controversy over testing has been particularly active in both the United States and Great Britain” (1971, p.70), therefore both countries were deemed most suitable for inclusion in this study.

2.1 Defining Standardised Assessment

When entering into discussion on standardised assessment it is vital that we clearly define what this form of assessment essentially entails, as many opinions are offered in this regard (INTO, 1997, 2010; Government of Ireland, 1999; Kohn, 2000; Santrock, 2001a; NCCA, 2005, 2007; Wall and Burke, 2004, 2007; Educational Research Centre (ERC), 2006, 2007; Mac Ruairc, 2009). Therefore, this section aims to clearly outline the definition of standardised testing upon which this dissertation is based, and explore what distinguishes this type of assessment from others.

A standardised test is an instrument of assessment that contains standardised procedures for its administration and scoring and for the interpretation of its results. In practice, the term ‘standardised test’ is most often applied to assessment instruments that contain objectively scored items that are produced commercially by a test agency and that are norm-referenced.

(NCCA, 2005, p.2)

This quotation shall embody the working definition of standardised assessment employed throughout the course of this study. From this clear definition we can ascertain two key features of standardised assessment that promotes its widespread usage in primary schools in Ireland; namely its inherent ‘objectivity’ and the fact that these assessments are ‘norm referenced’, whereby allowing teachers to grade their students on a national scale based on class-level or age. These tests are also said to be high in reliability, purporting that a child will perform consistently if the test was “carried out on the same learner by another assessor/ marker or if it was carried out on the same learner at a different time” (Boland, 2008, p.45). This beneficial feature of standardised assessment undeniably promotes its incorporation in our classrooms nationwide as it enhances the confidence that users have in the test; namely teachers, schools, Inspectors and parents. They can also provide valuable information to teachers in relation to a pupil’s general attainment level in English and Mathematics, when used in conjunction with other sources of assessment.

Validity is also an important feature associated with standardised testing. ‘Content validity’ refers to the ability of the standardised test to assess, and effectively measure, the content that it claims to examine (Santrock, 2001a). In this regard, the standardised test creators provide reassurances that the tests are specifically constructed with this issue in mind, thereby ensuring that appropriate curricular context is carefully chosen and comprehensively assessed (Wall and Burke, 2004, 2007; ERC, 2006, 2007). Likewise, ‘construct validity’ is also a vital feature promoted by the test creators. This refers to the “degree to which the test actually measures that which it purports to measure” (ERC, 2007, p.43). In this manner, as these types of assessments measure achievement in a specific domain, for example literacy, they are therefore constructed with this sole objective in mind. Furthermore, the scores obtained from standardised tests are measured and interpreted using carefully constructed scales to ensure that these tests are measuring precisely what they aim to measure.

2.2 The Key Benefits Associated with Standardised Testing

This form of assessment has been the subject of much debate in recent years, and thus extensive research, more so in the international context, has been undertaken to ascertain the key benefits of adopting a standardised approach to assessment in the primary school classroom.

Firstly, the most obvious benefit associated with standardised testing is that the results of this form of assessment provide teachers with a complimentary source of information regarding student achievement. In this manner, “standardised test results contribute to the accuracy of the teacher’s monitoring, and assist in identifying the needs of individual children” (NCCA, 2005, p.60). Thus, when used in conjunction with teacher-designed tests and tasks, standardised test results can facilitate the teacher in effectively monitoring and situating a student’s learning on a national scale, a widely acknowledged beneficial feature within its own right (Wall and Burke, 2004, 2007; NCCA, 2005, 2007; ERC, 2006, 2007, INTO, 2010).

Also, the results of standardised assessments are essential when schools are evaluating their resource provisions and time allocation for students requiring learning support. Ex-Minister for Education Noel Dempsey explicitly promoted this beneficial aspect in his press release regarding the implementation of standardised testing, in which he stated that “policy decisions regarding the allocation of resources to support special needs must be underpinned by up-to-date and reliable data” (2004). Here, the test results are deemed an objective marker of how a child is performing in relation to his/her peers nationally, and so can contribute to validating a teacher’s request for additional support for those scoring exceptionally low, or indeed for those on the opposite end of the spectrum who are gifted and consistently scoring exceptionally well. Also, in this regard, the results can provide a teacher with a valuable objective source of information when reporting to parents about their child’s achievement. Here, the result obtained on the test can support a teacher’s professional judgement in relation to pupil performance.

Furthermore, standardised testing can be viewed as a catalyst that results in effectively enhancing teacher professionalism. On this topic, Wall and Burke note that “by analysing the results of an entire class, a teacher can gain useful feedback on which areas or aspects of the [...] curriculum have been mastered by pupils and which ones are posing problems” (2007, p.2). Here, as well as identifying the pupils’ strengths, when analysing the test results a teacher may notice common areas of difficulty experienced by the majority of the pupils in the class. This may point to the fact that the class may benefit from consolidation in this area and thus the teacher can dutifully construct a programme of work to ensure this area is effectively addressed, thereby enhancing the quality of their instruction and avoiding future difficulties in the area.

Finally, the information obtained from standardised assessments can be used to track educational trends over time, thus effectively monitoring our educational system and maintaining favourable standards in primary schools nationwide. In compliance with the Education Act (Government of Ireland, 1998), Inspectors require such test results from each school in their jurisdiction to ascertain the general level of pupil achievement in that particular site, subsequently comparing current scores to past standardised testing results and probing reasons behind any fluctuations. This feature

is also highlighted by the NCCA, when stating that “standardised tests contribute to the evaluation of schools by the inspectorate of the Department of Education and Science” (2005, p.4). In this manner, standardised testing contributes towards quality control in our education system, when used in conjunction with other related data collected from each school.

2.3 The Irish Context

2.3.1 Recurring themes in Irish Literature

Upon close examination of the limited Irish literature available in this area, it became evident that three particular themes frequently reoccur, each of which are outlined and examined below.

2.3.2 The Theme of ‘Accountability’ in Irish Literature

The theme of accountability is perhaps the most significant ‘by-product’ of standardised testing identified in Irish literature. The limited depository of Irish research in this area would suggest that ‘accountability’ acts as a catalyst in promoting anti-educational practices, such as ‘teaching to the test’ and other result-enhancing exercises, explored more comprehensively later. Accountability arises when the results of standardised tests are used for means other than solely assessing student achievement, their true purpose. As the INTO note, there is “a generalised acceptance that a range of interest groups such as policy makers, parents and the wider public have an entitlement to information and data on how effectively teachers, schools and the education system are performing” (1997, p.5). This public belief, coupled with the current legal obligation that teachers must administer standardised assessments and report the results to parents (Government of Ireland, 1998, Act 22) can create pressure on teachers to provide favourable results. Researchers in Ireland resoundingly concur that this increased pressure can influence teachers to adopt questionable practices aimed at enhancing their pupils’ performance on the standardised test (INTO, 1997, 2010; NCCA, 2005; Mac Ruairc, 2009).

Such circumstances are addressed directly in the NCCA’s insightful report which plainly warns against adopting “measurement-driven instruction [...] resulting in a

narrowing of the curriculum” (NCCA, 2005, p.8). Here, the NCCA is advising teachers to learn from the mistakes observed internationally, particularly in the United States, where accountability has raised the stakes involved in using standardised assessments. Although we have not reached a ‘high-stakes’ situation in Ireland, whereby schools are pitted against each other and marketed on the strength of their test scores, as is the case in both the United States and Great Britain, Irish-based literature would imply that we certainly seem to be moving in this direction. Though still a relatively new phenomenon to the Irish market, we have already progressed into what the NCCA deem a ‘medium-stakes’ situation (the relatively recent development of which is clearly outlined in appendix one), as by sharing the standardised assessment scores with parents, other teachers, Principals and Inspectors, “information other than the assessment results may now be inferred about teachers, about classes or about schools” (NCCA, 2005, p.8). On this matter Corbett and Wilson further clarify that “people may attach a level of stakes to a test that is out of character with the formal consequences associated with it” (1991, p.26), as often the degree to which such ‘stakes’ are experienced is defined by local perception as opposed to official statute. Here we can identify the root of the accountability problem; the ‘Big Brother’ type effect, whereby teachers feel their ability is being appraised by those who define these stakes, based on the results the pupils in their class obtain in the standardised assessments. This point was also highlighted in the proceedings of the Consultative Conference on Education 2008 (INTO, 2010).

Worryingly, the Green Paper on Education (Ireland. Department of Education and Science, 1992) originally supported a move to high-stakes assessment, as outlined in appendix one. Clearly those involved in the construction of the Green Paper were misinterpreting the original purpose of standardised assessment, creating a situation whereby the outcomes of these assessments were far more wide reaching than was intended by their creators. Thankfully, with the introduction of the White Paper (Ireland. Department of Education and Science, 1995), any such notions were abated, yet here we can see how the misuse of standardised testing could have led to the ‘marketability’ of our primary schools, resulting in the unwanted ‘high-stakes’ dilemma that the United States and Britain now find themselves in. Mac Ruairc’s related study also highlights this accountability debate, when stating “the arguably

inevitable link between testing, once mandatory, and externally imposed accountability requirements is the source of considerable tension at school level in many countries” (2009, p.48). It is therefore vital that we learn from the experiences of other countries when implementing appropriate standardised testing procedures in our own context, not allowing accountability to rule assessment but rather vice versa.

From the above account it is evident that accountability is a vital issue to consider when incorporating standardised testing into an education system. Informed literature would suggest that a more holistic approach to teacher appraisal should be adopted, rather than centring any judgement of a teacher’s ability on a sole result, itself open to influence from uncontrollable factors such as stress felt by those undertaking the assessment (Wall and Burke, 2004). Here, the literature examined suggests that if we do not monitor this situation carefully, accountability may start to define the purpose of assessment, which is arguably already the case in both the United States and Great Britain.

2.3.3 The theme of ‘Teaching to the Test’ in Irish Literature

When reading the limited body of literature relating to standardised assessment in the Irish context, it was noted that ‘teaching to the test’ was commonly identified as a major issue, impeding the effective use of standardised testing in primary schools. Wall and Burke clarify that ‘teaching to the test’ is evident where pupils are “selectively taught those areas of the [...] programme which feature in the tests” (2007, p.5), a clear explanation that shall constitute the working definition adopted throughout the course of this study when referring to this feature. This questionable activity is also widely documented in the international literature, and as we have followed in the footsteps of other countries, more established in the use of standardised testing, it stands to reason that we have also inherited such problematic features. Irish research highlights that the adoption of standardised assessment in the primary school classroom can lead to a ‘teaching to the test culture’ (Mc Neil 2000; Lam and Bordinon, 2001; Anagnostopoulos, 2005; Mac Ruairc, 2009; INTO, 2010). Here, teachers explicitly cover material that is to be examined on the test in a bid to raise the likelihood that their students will do well, thus reflecting favourably on their own teaching ability. This unconstructive activity is clearly noted in research carried out by the INTO, who aptly state that:

Some teachers feel pressurised to go to considerable lengths in order to boost results– even if this means that the results are invalid, inflated measures of childrens’ real achievement levels and even if the achievement of higher scores requires the adoption of teaching approaches which the teachers believe are of dubious pedagogical value or even downright anti-educational.

(INTO, 1997, p.19)

Based on this insight one can conclude that this particular practice is intrinsically linked to the issue of accountability, as it is evident that increased pressure on teachers to be accountable for their pupils’ performance can lead to such practices infiltrating the Irish education system in such a manner. Interestingly, in 2004 when the then-Minister for Education, Noel Dempsey, released a statement promoting the use of standardised assessment in primary schools, it explicitly referred to this negative feature whereby the minister simply states “I am confident that teachers in Irish schools will not fall into the trap of ‘teaching to the test’” (2004). Here, overtly addressing this complex issue is seen as a satisfactory technique in abating its infiltration into our education system. In stark contrast, those who constructed the Irish standardised test instruments deem it more plausible to take a stronger stance on the matter, stating that “it cannot be emphasised too strongly that pupils should not be selectively taught any of the material in the test, as this would inflate scores and invalidate the results” (Wall and Burke, 2004, p. 1). Having researched the effects of directed revision, or ‘teaching to the test’, in other countries more experienced in using standardised testing, the Irish test creators are issuing a clear warning against the adoption of this practice in our own context. However, having been issued with clear warnings to the contrary, research would suggest that such practices still continue to pervade our education system (NCCA, 2005, 2007; Mac Ruairc, 2009; INTO, 2010). When one considers the current consequences of having pupils that score below par on the standardised test, “It is not surprising that both teachers and students have been found to regard examination success and examination-oriented activities as being very important” (Fontes *et al*, 1980, p.54). However, no directed investigation has been carried out to ascertain the true extent to which this activity currently exists in Ireland, a situation this dissertation aims to address within the confines of this study.

2.3.4 The Theme ‘Impact on Classroom Practice’ in the Irish Literature

This prevalent theme can be seen to link with the two previous themes outlined in the Irish literature. As issues relating to accountability can result in some teachers ‘teaching to the test’ or employing other result-boosting methods, it stands to reason that classroom practice is altered to facilitate the inclusion of such activities. Here, Haladyna, Bobbit-Nolen and Hass (1991), as cited in INTO (1997, p.19) report that “a great deal of time is devoted to preparing for the assessment in order to try and maximise test performance rather than promoting the ‘official’ goals of education”. This astute observation suggests that some teachers narrow the curriculum being delivered in their classrooms to ensure assessment success. Thus, by implication, other non-tested meaningful forms of learning are suffering. This neglectful pre-testing feature is highlighted by the NCCA who report that pressure emanating from the over-importance placed on standardised test scores can create “pressure on teachers to coach children for the tests, and to focus teaching and learning on test practice” (2005, p.9). Thus, logically one can deduce that time spent preparing for the test comes at the expense of time that could have been spent delivering a broad and balanced curriculum which “affords flexibility to the school and the teacher in planning the learning experiences that are useful to the individual child at the various stages of his or her development”, as advocated by our national curriculum (Government of Ireland, 1999, p.18). Yet, research would suggest that the aforementioned ‘flexibility’ of the curriculum is greatly diminished when teachers feel constrained to cover as much of the curriculum as possible before standardised testing begins, adopting methodologies that aim to maximise concept acquisition of the material that is to be assessed in the tests (INTO, 1997, 2010; NCCA, 2005; Mac Ruairc, 2009) rather than methodologies that aim to creatively develop the child’s learning experiences, as espoused in the primary school curriculum guidelines (Government of Ireland, 1999).

Interestingly, it was noted that literature documenting post-testing impacts on classroom practice was lacking from the Irish perspective. Here, the limited number of studies conducted to date in Ireland have referred to pre-testing impacts, as outlined above, yet have not focused on the post-testing aspects, an area this dissertation aims to explore.

From the Irish literature explored above we can identify that classroom practice can certainly be influenced by standardised assessment usage, yet no research on this topic has been carried out to ascertain to what extent this is happening. However, one can plainly observe that Irish bodies, such as the NCCA and INTO, refer to U.S. and British research in this regard, aptly applying the features and concepts identified in this literature to the Irish context, due to the closely related procedures shared with these countries in the administration of standardised testing. However, it should be noted that the stakes in the United States and Great Britain are higher than those involved in the Irish situation. Yet, by comparing our situation to that of the United States expert bodies are implying that our context has the potential to mirror that of the United States and inherit many of the same problems highlighted by leading theorists in American literature, a point further reiterated by the NCCA when stating “What is intended as ‘low stakes’ could quickly become ‘high stakes’ in the absence of other data on student progress and system effectiveness and quality” (2005, p.25). Therefore, one cannot but recognise the urgent need for pre-emptive measures in this regard. When raising awareness in relation to the correct usage of standardised assessments and in outlining the negative features linked to their misuse, teachers and policy-makers in Ireland can effectively support and enhance assessment usage, thus maximising test efficacy and avoiding entering into a detrimental ‘high-stakes’ approach to assessment.

2.4 The International Context

2.4.1 Recurring Themes in International Literature

This section provides an account of the prevalent themes identified in international literature, focusing attention particularly on those deriving from the United States and Great Britain for reasons previously discussed. It is vital that international literature based on standardised testing is examined to allow us to identify and avoid obvious pitfalls experienced by other countries and compare features outlined in international literature with those identified in our own. This allows the researcher to ascertain the extent to which such features are present in the Irish primary school context and to hypothesise what developments may lie ahead. To further contextualise the literature presented in this section, appendix two contains an

overview of the historical development of standardised testing from both the American and British perspective.

2.4.2 The Theme of ‘Accountability’ in International Literature

2.4.2.1 The American Context

When considering standardised assessment, Haladyna, Bobbit-Nolen and Haas note “the considerable increase in the use of these test scores might be attributed to the onset of the ‘age of accountability’ and an increased perceived need to evaluate education at virtually all units of analysis” (1991, p.2). Here one can infer clear undertones in the American literature that standardised assessment scores are being used far beyond their means in the appraisal of teachers and their teaching. At present in the United States the stakes involved in standardised testing are deemed ‘high’, yet in reality researchers concur that it would be more appropriate to label these stakes as ‘dangerous’, owing to the fact that from a single result, a number of important decisions can be made (Kohn, 2000; Abrams, Pedulla and Madaus, 2003; Hursh, 2005). In the United States, for example, it has been found that increased pressure is placed on teachers for their pupils to perform well “stemming from the publishing of test results in newspapers and on television and the subsequent comparison of scores among schools and school districts” (Cimbricz, 2002, p.7). Clearly, standardised test scores are being used out of context here and over-emphasised in order to ‘market’ schools. When faced with such serious repercussions, it would seem that teachers in such situations may engage in result-boosting activities to ensure that their class’s scores remain high, and thus reflect favourably on both themselves and their school’s status in the marketplace. Literature in this area proposes that in such cases “the test becomes the curriculum” (Neill and Medina, 1989, p.694) as teachers focus solely on the material upon which their pupils will be tested, to the detriment of other meaningful forms of learning. Irish research on the American situation has further concluded that

A range of effects of ‘high-stakes’ tests has also been identified for schools, teachers, parents and the system at large, although it is debatable whether these can be classed as ‘unintended’ since they are usually associated with the publication or dissemination of results on a school-by-school basis.
(NCCA, 2005, p.9)

Here, the NCCA are clearly stating that some effects, such as the marketing of American schools in league tables, can hardly be seen as unintentional when each school knowingly releases its results yearly to the media in a bid to out-perform neighbouring institutions. As Kohn, a leading researcher in American educational practice highlights “only a person ignorant or dishonest would present a ranking of schools’ test results as though it told us about the quality of teaching” (2000, p.7), yet, as we can see, this questionable practice is a reality in the international context.

Abrams, Pedulla and Madaus further state that “central to the current state accountability models is the need for steady increases in test scores as indicators of improved student achievement and, in turn, school effectiveness” (2003, p.25), proposing that this added pressure can act as a catalyst in promoting other negative practices to ensure that students perform well, such as ‘teaching to the test’ and altering instruction to guarantee student success.

2.4.2.2 The British Context

In a similar fashion, developments in accountability in Great Britain can be seen to largely mirror those outlined in the United States, as shown in appendix two. Bartlett astutely recognises that “the desire for accountability, standardisation and control of the work of teachers has taken precedence over [...] the quest for raising standards” (2000, p.25). The original purpose of standardised testing was to measure student attainment in a specific subject area, yet literature would suggest this focus seems to be lost in countries such as Great Britain, where pupils’ scores are seen as an easy and immediate tool for teacher appraisal without recognising the fact that each student and their needs are unique, developing at different stages (Bartlett, 2000). Therefore, no single test score could possibly indicate a student’s yearly progress, a feature that the Irish test creators stress unequivocally (Wall and Burke, 2004, 2007; ERC, 2006, 2007). However, this is a point that requires some clarification amongst the general British public as results from such assessments can often be mistaken as an indication of general ability. This point is further validated in research conducted by Zitlow, who concludes that “no one ‘assessment’ can cover all of what we want students to know and be able to do” (2001, p.113). Yet, Zitlow advocates that, worryingly, this seems to have become the case. Furthermore, with increased pressure on British teachers to produce high scores, anti-educational activities have

seeped into classroom practice in order to inflate pupil-performance. Such activities include ‘teaching to the test’, explored in the following subsection, and deviating from specified administration procedures. Thus, upon close examination of the British literature, it is plain to see that many subversive and anti-educational practices can arise due to pressures directly arising from increased accountability in the British context.

2.4.3 The Theme of ‘Teaching to the Test’ in International Literature

Upon close examination of literature from the international perspective, one can clearly ascertain that ‘teaching to the test’ constitutes a key feature that presents itself in many studies (Gipps, 1998; Cizek, 2001; Cimbricz, 2002; Abrams, Pedulla and Madaus, 2003; Bond, 2004; Posner, 2004).

2.4.3.1 The American Context

This practice is highlighted in Brown’s study on standardised testing in the United States when stating that “teachers reported altering the scope and sequence of the curriculum and eliminating concepts that were not covered on state tests” (1992, p.13). Here, Brown has clearly identified that teachers are narrowing the curriculum, selectively teaching material that will appear on the test and neglecting other areas included in the curriculum. This practice undoubtedly serves to hinder the holistic education envisaged for each child, in favour of raising their scores on standardised assessments.

An interesting incongruity that presents itself in literature from the United States is the belief by teachers that they are not engaging in this practice, yet when outlining the activities they incorporate into their teaching, it is found they are indeed ‘teaching to the test’. This anomaly can be observed in a study undertaken by Cimbricz who states that “testing led to a 40% increase in such activities as drills, coaching, and practice on testing. It is interesting to note [...] such activity was not necessarily viewed by teachers as a narrowing of the curriculum” (2002, p.10). Here, activities directly related to ‘teaching to the test’ are not viewed as such by all teachers and so we can infer that some teachers deem these practices as valid test preparation activities, even though they are proven to inflate a child’s test score beyond their actual capability.

Cizek further explores this ‘unintentional’ aspect and found that “in California, 36 percent of teachers thought it appropriate to practice with current test forms” (2001, p.3). The literature here would suggest that it is vital that acceptable practice be clearly defined so that unintentional ‘test polluting’ practices, such as that outlined by Cizek above, are avoided, as the results from the tests taken in this manner are of no educational value.

Furthermore, when discussing ‘teaching to the test’, Dave Posner, a leading researcher in assessment in the United States, strongly believes that “pressure causes teachers to devote virtually all classroom time and resources to preparing students for the standardized test” (2004, p.749). In such a ‘high-stakes’ testing climate, Posner argues that most of the classroom time is spent preparing students to score well on tests so that teachers and their teaching will not become the undue focus of attention, the consequences of which can be quite serious, as previously explored.

An interesting feature in American literature is the debate over what ‘teaching to the test’ essentially entails. In his directed study in this area, Bond defines ‘teaching to the test’ as a continuum.

At one extreme, some teachers examine the achievement objectives as described in their state’s curriculum and then design instructional activities around those objectives [...] At the other extreme is the unsavoury and simply dishonest practice of drilling students on the actual items that will appear on the tests.

(2004, p.1)

Here, we can clearly infer that there are less severe and more severe forms of ‘teaching to the test’ in which teachers have been found to engage. ‘Item teaching’ has been deemed the most serious deviation from good practice, as when a child is actually taught the items on a test it completely invalidates the score they obtain, as we cannot know if the child would have arrived at this score unaided. Thus American literature in this area is in strong agreement, whereby it has been found that ‘teaching to the test’ is an activity that results solely in rendering the scores of standardised assessments useless and limits the learning experiences offered to children, in favour of more test-oriented activities. Neill and Medina aptly summarise the American stance on this unsavoury activity stating “as teaching becomes ‘coaching for the test’ in too many schools, real learning and real thinking are crowded out” (1989, p. 694).

2.4.3.2 The British Context

Many of the same observations were also made in relation to the existence of a 'teaching to the test' culture in the British context. As the stakes involved in standardised testing in Great Britain are also deemed 'high', some teachers adopt this activity to enhance their class's performance and to ensure that their teaching is not the object of scrutiny. Tymms (2004) recognises this feature in his study in which he contemplates why standardised testing has failed to monitor standards. He argues that "pressure makes it hard to interpret the data. Teaching test technique must surely have contributed to some of the rise, as must teaching to the test" (2004, p.492). Here, Tymms notes that the results of standardised tests are difficult to interpret due to 'test pollution' practices, such as 'teaching to the test' artificially boosting childrens' performances.

Some theorists, such as Pierson, as quoted in Gipps (1988), denounce opposition to this feature. When faced with the reality that schools are 'teaching to the test' in order to boost scores, Pierson believes "that doesn't mean they're cheating...but they are moulding their curriculum to fit what the CAP tests" (1988, p.31). Here, Pierson advocates that this activity does not constitute cheating at all, merely altering the curriculum. However, on this topic, research undertaken by Zitlow concludes that "increased efforts to raise standardized test scores come at the expense of more meaningful forms of learning" (2001, p.112). Moreover, Turner and Clift further support Zitlow's stance on this issue, stating that "teachers pressured to ensure their students pass [...] spend huge amounts of time drilling minimal competency skills" (1988, p.314). Here, one can detect a certain irony, whereby the children who are meant to benefit from standardised testing are the very ones losing out. Thus, the original focus of this form of assessment in the British context is clearly lost in the battle to secure high scores using unconstructive and questionable means.

2.4.4 The Theme of 'Impact on Classroom Practice' in International Literature

There has been a great deal documented in international literature regarding the impact of standardised testing on classroom practice, again predominantly from a pre-testing perspective. This is unsurprising, as given the arguments surrounding the high-stakes involved in the standardised testing process in these countries, positive post-testing impacts are not the object of focus in many cases. Abrams, Pedulla and

Madaus further support this stance, stating that “much of the research in state testing programs addresses their effects on what is taught” (2003, p.19), neglecting to address how such programmes may benefit classroom practice/ teacher instruction.

2.4.4.1 The American Context

Many American theorists concur that this form of assessment impacts teachers and their teaching negatively (Brown, 1992; Kohn, 2000, Crimbricz, 2002; Abrams, Pedulla and Madaus, 2003; Kraemer, 2005). Crimbricz explores some of the aforementioned ‘negative’ impacts in her insightful study on the topic. These include;

- 1) narrowing of the curriculum and instruction, 2) fostering of anxiety, confusion, fear, shame, anger, and/or mistrust, 3) deskilling of teachers and/or a perception of powerlessness, 4) the invalidity and inadequacies of these tests as accurate measurements of what is taught and learned, and 5) loss of instructional time due to test preparation and testing.

(2002, p.6)

Here, Cimbricz clearly outlines the unconstructive impacts standardised testing can exert on classroom practice. This is but a sample from an extensive list of such practices that can infiltrate our classrooms owing to pressures relating to raising the childrens’ test scores on standardised assessments. In such instances, childrens’ education can be seen to suffer as they are not being provided with as holistic a curriculum as intended.

Interestingly, Abrams, Pedulla and Madaus conducted an extensive survey of teachers’ opinions in the United States towards state-testing, and found that “76% of high-stakes teachers and 63% of low-stakes teachers [...] reported that their state testing programme has lead them to teach in ways that contradict their own notions of sound educational practice” (2003, p.23). Perhaps most significant here is the admission by sixty-three per cent of low-stakes teachers that their practice is negatively affected. This is interesting as in our own context we are what is commonly considered ‘medium-stakes’ (NCCA, 2005). Therefore applying this logic to our own context, more than half the teachers in Irish primary schools are possibly adopting practices that undermine the quality of the educational experiences being offered to children in our education system. Clearly, the international perspective on this matter can inform Irish policy formation, whereby we must dutifully install

safeguards in our education system to identify the extent to which this problem may have pervaded current practice, subsequently alleviating any such sinister aspects.

One key reoccurring feature in literature emanating from the United States that must be explored when considering how classroom practice has been affected by standardised testing is the use of time. Zitlow closely examines this area, focusing particularly on the work of Kohn (2000), consequently concluding that “Time spent preparing students to succeed on such tests is time that could have been spent helping them become critical, creative, curious thinkers” (2001, p.113). Here, Zitlow notes that teachers spend varying amounts of time preparing their students for standardised assessments, time that could have been spent more productively. Generally, when preparing the students for the tests, teachers can engage in both ethical and unethical practices. Ethical procedures can include such activities as practising filling in the answer sheet properly if a different format to what they are used to is encountered. Unethical practices may include ‘teaching to the test’, as previously explored. Both types of activity demand time and a focus generally on lower-order thinking skills. Zitlow (2001) therefore supports the belief that the time spent conducting these activities is time that could have been spent developing the childrens’ critical thinking skills and engaging in the more complex, higher-order skills that are not valued or assessed in the standardised assessments.

2.4.4.2 The British Context

The British literature in this area again generally parallels that deriving from the American perspective. Once-more ‘time’ was the key focus of a British study conducted by Green (1992) in which it was found that teachers estimated they spent between ten to fifteen per cent of their time preparing for the tests. When we consider these percentages in the context of an average working day in primary schools we can clearly distinguish that a significant amount of time is spent directly focusing on testing. Here, a clear impact on classroom practice is identified, as the time that is spent on testing and test preparation is time that the children could have spent engaged in more meaningful forms of learning that standardised testing cannot assess. Broekhoff further supports this stance, stating that “for some types of learning, especially in the affective domain, experience is more important than competence: doing it at all is what counts, not doing it well” (1978, p.36). Here,

Broekhoff outlines that there are some important types of learning that cannot be measured or assessed, and in the current British educational climate which values and prioritises those areas of the curriculum most suitable for formal assessment, these forms of learning are largely marginalised.

Again, literature deriving from the British context overtly concurs with that originating from the United States in its belief that standardised testing impacts on classroom practice in a definitively negative manner. Kraemer (2005, p.88) notes that “standardized tests narrow the curriculum and distort authentic learning”. When focusing on areas of the curriculum that will be assessed, other features are neglected, as the end goal of this type of teaching is the raising of scores as opposed to the holistic increasing of knowledge and experience. Here, research clearly advocates that this type of results-driven practice will ‘distort authentic learning’ in a very direct way, as children are not being provided with the range of learning experiences embodied in the carefully constructed British curriculum. This ‘narrowing’ of the curriculum is a feature often highlighted in the international literature in this area by those who are opposed to standardised testing.

Furthermore on this topic, Gipps (1988, 1994) outlines how some teachers in Britain focus their instruction on subjects being assessed in the weeks leading up to the standardised tests. In her study, attention was also drawn to the practice whereby those in managerial positions within a school may also insist on teachers undertaking this questionable practice yearly, in a bid to ensure their school’s continued success in the all-important annual school ‘league tables’. Clearly, this reality is one that no researcher would expect to find documented in any professionally run school’s assessment policy. This perceived ‘high standard’ in education can thus be used to publicise the school in a favourable light. Gipps astutely notes “the stimulus (testing) is applied and the outcome (improved test performance) hoped for, but the process linking the two remains largely undiscussed; it is the ‘black box’ metaphor” (1988, p.30). Such research would suggest that we must dutifully ask ourselves, if a school improves its standardised testing scores, or if they are persistently scoring higher than other schools in the same district, what is causing this ‘enhanced’ performance? Such questions need to be addressed if those in the British context are to ensure that

‘black box’ activities, such as those outlined in this chapter, do not undermine the carefully constructed education system in place.

2.5 Conclusion

In conclusion, from the above exploration and comparison of literature surrounding standardised assessment, it is plain to see that practices and features identified in the Irish context closely align themselves with those found in the international literature, though more conservative in nature. Thus, we are able to situate our progression on an international scale and can therefore hypothesise what developments we may encounter in the future. Furthermore, in learning from the experiences of both countries explored above, we can dutifully avoid the documented pitfalls associated with this form of assessment that have become evident in such ‘high-stakes’ testing nations.

Clearly, as well as noting the benefits associated with this form of assessment, the themes of ‘accountability’, ‘teaching to the test’ and ‘impact on classroom practice’ also frequently present themselves in literature based on this topic, predominantly from an international perspective. However, as noted previously, such features are frequently alluded to in the Irish-based literature, yet seldom researched explicitly. As Mac Ruairc notes “the absence of debate in Ireland in relation to the impact of a policy of mandatory testing on children [...] is deeply regrettable” (2009, p. 52). On this matter, the INTO notes that a large-scale revision of the current practices and policies defining the use of standardised testing in the Irish education system is now required (INTO, 2010). This vital requirement was of paramount importance when devising the research questions at the heart of this study. In this regard, it was decided to directly align each research question with one of the three specific reoccurring themes examined in the literature in order to explore the extent to which these practices are currently present in our primary school classrooms. It was also decided to engage in a general examination of the benefits associated with this form of assessment as this may prove fruitful when analysing the data collected later in the research, facilitating the presentation of a more balanced view of standardised

testing. Table 2.1 below documents the initial research questions formulated, each linking to one of the three main themes explored in the literature above.

Table 2.2: Initial Research Questions

Research Questions:	Linking to:
1 How does accountability affect teachers' attitudes towards standardised testing?	Literature based on 'Accountability'.
2 In what ways are children prepared for standardised testing?	Literature based on 'Teaching to the Test'.
3 What are the impacts on classroom practice as a result of the adoption of standardised testing in our schools?	Literature based on the impact standardised testing has on classroom practice

When considering that the phenomenon of standardised testing has become so infused in our education system, it is understandable that a passive acceptance of this form of assessment has occurred, somewhat explaining the limited Irish research conducted in this area. Having previously been 'successfully' established in international education systems, it stands to reason that Ireland would follow suit and adopt this recent educational innovation, yet as Spolsky notes "tests should be labelled just like dangerous drugs: 'Use with Care!'" (1981, p. 20). When standardised tests are used for the purpose intended, these tests can provide the teacher with valuable information relating to a child's attainment in a specific subject area. Yet, aforementioned literature would suggest that in the case of 'medium' and 'high' stakes testing, results are being used out of context and the initial purpose of standardised testing is being manipulated and subsequently utilised to influence decisions far beyond its intended means. In reference to this situation, Turner and Clift surmise that "the testing 'tail' too often wags the curriculum 'dog'" (1988, p.313), a point also supported by the NCCA (2005). Thus, it is apparent that an evaluation of our national policy on assessment is undoubtedly due (INTO, 2010), a requirement that has resulted in the formation of this dissertation. The means by which such an evaluation was carried out within the confines of this study are clearly outlined in the following chapter.

Chapter 3: Methodology

3.1 Introduction

This chapter focuses on outlining the research process undertaken to ensure the effective collection of qualitative data, required to answer the research questions documented in the previous chapter.

To comprehensively outline the research process employed, this chapter is split into clear sections, each examining separate facets of the methodological process.

Following this introduction, section two explores the research design, justifying the qualitative approach adopted in this study. In the third section, the researcher clearly outlines the paradigm within which this research is situated, thus allowing the reader to place this study within the correct context. Following this, the fourth section then documents and validates the research strategy employed. This leads to the fifth section, which explores the methods of data collection, outlining their suitability for use, as well as their limitations. Section six dutifully examines the sampling technique used and links with section seven which looks at ethical issues in relation to the research, such as those experienced when gaining access to participants in the sample. Section then eight explicitly explores the creation, piloting, administration and analysis of each research instrument. As there are two research instruments, the self-administration questionnaire and the interview schedule, each is dealt with separately in order to effectively outline the differing methodological and analytical procedures employed in each case.

3.2 The Research Design

As the focus of the research is on gaining and analysing teachers' attitudes and experiences, a qualitative approach is justly employed. As Sherman and Webb note "Qualitative research has the aim of understanding experience as nearly as possible as its participants feel it or live it" (1988, p.7). Likewise, the aim of this research is to ascertain teachers' attitudes towards, and experiences of, standardised testing, thus understanding this form of assessment from the viewpoint of those most experienced in its administration. Miles and Huberman further support the qualitative focus of

this research on “naturally occurring, ordinary events in natural settings, so that we have a strong handle on what ‘real life’ is like” (1994, p.10). As standardised assessment is a ‘naturally occurring event’ in primary schools nationwide, by delving into practitioners’ experiences utilising this widely adopted form of assessment the researcher aims to investigate the forces present in these social contexts that result in certain attitudes being formed, thereby exposing what ‘real life’ consists of for teachers in primary school settings. This qualitative exploration will result in the generation of data upon which conclusions and recommendations can be based, thus supporting teachers’ practice in this area

In stark contrast, a quantitative research approach would entail a focus on statistical data, the emphasis of which is more centred on numerical information, as opposed to verbal accounts. Blaxter, Hughes and Tight (1996) would further support this claim, stating that “quantitative research is, as the term suggests, concerned with the collection and analysis of data in numeric form” (p.60). Thus, a quantitative research approach would have been wholly unsuitable in this case, as it is unquestionably the rich *verbal* accounts embodied in teachers’ attitudes and experiences that constitute the focus of this study.

3.3 Locating the Research within a Paradigm

In order to outline one’s methodology comprehensively, it is necessary to understand in which paradigm their research is situated as “it is the choice of paradigm that sets down the intent, motivation and expectations for the research” (Mac Kenzie and Knipe, 2006, no page). Considering the emphasis on exploring teachers’ socially-constructed attitudes regarding standardised assessment in this study, it was found that traits of a ‘constructivist’ approach were present, whereby “the ways in which we understand the world and the things we consider true are not just natural ways of understanding reality [...] they are constructed between people as they go about their everyday lives” (Miell, Phoenix and Thomas, 2002, p.69). Thus, teachers form their own opinions about standardised testing based on their experience of utilising this form of assessment and their interactions with other practitioners on the subject. Here, the researcher believes that by investigating such views, it is possible to

identify the underlying reasons behind the formulation of attitudes on the topic of standardised testing.

Furthermore, it also became clear that the research was undertaken with an ‘interpretive’ approach in mind. Bassey (1995, p.13) states that “Interpretive researchers recognise that by asking questions or by observing that they may change the situation which they are studying”. In this regard, the researcher believes that by probing teachers’ attitudes it will be possible to highlight positive aspects, as well as problematic features, associated with standardised testing, thus raising awareness amongst those surveyed about such features in order to effect positive change, thereby improving current practice.

Interestingly, as both approaches outlined above are closely interrelated, they are often embodied in one paradigm; that of ‘Interpretivist/ Constructivist’ (Cohen and Manion, 1994; Creswell, 2003; Mertens, 2005; Mac Kenzie and Knipe, 2006), which is unquestionably the framework being employed in this research. Here, Mac Kenzie and Knipe note that “the interpretivist/ constructivist researcher tends to rely upon participants’ views of the situation being studied and recognises the impact on the research of their own background and experiences” (2006, no page number).

Applying this logic, the qualitative approach adopted in this study facilitates the collection of teachers’ views relating to standardised testing; the ‘situation being studied’. Also, as the researcher is a primary school teacher, their own experience of utilising standardised assessment has resulted in personal opinions being formed on the topic, thus added measures in the research design had to be taken to ensure that researcher bias did not influence any of the qualitative data collected, as explored in later sections.

3.4 The Research Strategy

When developing an appropriate research strategy, the research questions were carefully considered throughout the entire process. After much deliberation it was decided that the effective use of a survey strategy would best suit the aims of this research, as this approach to data collection would reach a wider sample of

participants than other research strategies, resulting in a richer pool of data. Other strategies, such as case studies, ethnography or action research would have been unsuitable here as the research questions demand that the focus of enquiry is based upon multiple teachers' attitudes. In order to derive such data, a flexible research strategy was required that would allow the researcher to broaden their sample to include enough of the target population so as to be able to make claims upon the data found. As the aforementioned strategies are limited in the populations they can encompass, such approaches to data collection were deemed unsuitable for use in this particular study.

It was further felt that the use of a survey strategy was most suitable here as this structured approach to data collection is most often used when collecting data based on attitudes and opinions. As Denscombe notes, "Surveys are associated with getting information 'straight from the horses mouth'" (1998, p.27). As this research is based on real-life experiences and opinions, a survey approach allowed the researcher to obtain data documenting teachers' attitudes towards standardised testing in their own words. Furthermore, due to limitations in relation to the time permitted to conduct the research, surveying participants produced a considerable amount of data in relatively little time (Denscombe 1998; Bryman, 2001, 2008). Other strategies do not offer this beneficial aspect. For example, ethnography demands a lot of time in the field observing participants and action research can be quite time-consuming when setting up an innovative project and monitoring its effectiveness. However, a noted limitation of this survey design can be identified here, whereby "generally, a relatively small amount of information is collected from any one individual" (Robson, 1993, p.49) in contrast to other approaches, as outlined above. To address this limitation research instruments must be scrutinised and piloted to ensure that information addressing each research question is effectively derived from each participant, an approach adopted in this study and outlined later in this chapter.

Denscombe justly states "The survey approach is a research strategy, not a research method. Many methods can be incorporated in the use of a social survey" (1998, p.7). Here, we can identify another noted benefit when adopting a 'Survey' strategy, whereby it acts as an umbrella approach, under which various symbiotic methods of data collection can be effectively combined. This aspect was of particular relevance

here as it was decided that anonymous self-completion questionnaires would constitute the main method by which data would be collected, the content of which was then complimented with the data derived from a semi-structured interview with a Department of Education Inspector.

When obtaining data from a number of sources using a survey approach, as outlined above, triangulation in the research can be seen to occur which “increases scope, depth and consistency in methodological proceedings” (Flick, 1998, p.230). Here, data collected by one method (the questionnaires) is cross-checked with that obtained from other viable sources (the interview and the in-depth literature review). This methodological feature enhances the validity of the research (Denzin and Lincon, 1994; Denscombe, 1998; Flick, 1998; O’Leary, 2004; Bell, 2005; Cohen, Manion and Morrison, 2007; Bryman, 2008). Denscombe believes that “seeing things from a different perspective and the opportunity to corroborate findings can enhance the validity of the data” (1998, p.85) and thus enhance the quality of the research by allowing the researcher to gain “a more holistic view of the setting” (Denzin and Lincon, 1994, p.224). Therefore, the employment of a survey strategy was further validated as it facilitated triangulation to occur, thus allowing the researcher to obtain a more informed and comprehensive overview of the opinions and attitudes associated with standardised assessment in the Irish primary school context.

3.5 Research Methods Employed

3.5.1 Self-Completion Qualitative Questionnaires as a Method of Research

When considering the nature of the research it was decided that anonymous self-completion questionnaires were best suited as the main research instrument in obtaining teachers’ attitudes towards standardised testing, for a number of reasons. Firstly, the research explicitly aims to investigate the impacts of standardised assessment from both a positive and negative perspective. As clearly outlined in the previous chapter, these attitudes can result in practices of an educationally questionable nature, such as ‘teaching to the test’, and therefore the nature of the research is quite sensitive. As the questionnaires are anonymous in nature, it was felt that this would promote the honest answering of questions as participants in the study

could not be linked to the responses given in any particular questionnaire.

Interviewing, as a main research method, was thus deemed unsuitable as when engaging face-to-face with the interviewer participants may be less likely to answer truthfully as it may reflect negatively on their professional practice.

Secondly, given that the research aims to encompass a relatively large sample for a small-scale research study, a data collection method had to be chosen that would allow for the easy gathering of data from participants that were geographically quite dispersed. Thus, self-completion questionnaires were deemed favourable in such a situation. A further beneficial feature is highlighted here by Bryman who states that “the cheapness of the self-completion questionnaire is especially advantageous if you have a sample that is geographically widely dispersed” (2001, p.129). When compared to costs involved in interviewing participants individually, such as travelling expenses, the cost of sending a self-completion questionnaire is significantly lower and thus more favourable.

Thirdly, having worked as a primary school teacher, it stands to reason that the researcher has socially constructed their own views and attitudes toward standardised testing. Thus, in order to ensure the validity of data gained from participants, it was vital that the researcher’s own opinions did not influence that of the participants in any way. As Bryman justly states, “since there is no interviewer present when a self-completion questionnaire is being completed, interviewer effects are eliminated” (2001, p.130). In this manner, the researcher ensured that their own background knowledge did not influence the data derived from participants, and thus the adoption of this method of data collection enhanced the validity of the research.

Also, as each participant is presented with a standardised questionnaire, variations in answers “are very unlikely to be contaminated through variations in the wording of the questions or the manner in which the question is asked” (Denscombe, 1998, p.105). Thus, responses given on the questionnaire have resulted from the attitude of the participant alone, a key feature required in a study of this nature. Furthermore, as each questionnaire contained standardised content, the data derived from the responses for each question was easily comparable and thus more straightforward to

analyse, a clear benefit experienced when utilising self-completion questionnaires as a method for data collection.

However, one must also be very clear as to the limitations of this research instrument. Firstly, it is widely documented that using this form of data collection can result in a low response rate (Kiddler and Judd, 1986; Denscombe, 1998; O’Leary, 2004; Bryman, 2008). As Kiddler and Judd state “a low response rate can call into question *any* conclusions based on the data” (1986, p.223, emphasis in original). Clearly, if there are a high number of non-respondents, claims made upon the data obtained may be unfounded as perhaps the opinions of many of those who did not respond contradicts the findings of the research. To avoid this situation every effort was made in this study to facilitate the easy return of questionnaires, including enclosing stamped-addressed envelopes with the questionnaire so as not to impose any cost on the participants, and ensuring that the length of the form was not inappropriately long. Secondly, when completing the questionnaire, a participant may become confused by the wording of a question and consequently may not provide a response for this item. As the researcher is not present to clarify such misunderstandings, it is vital that each question must be carefully and unambiguously worded. Here, the researcher carefully worded all questions and piloted the research instrument to assess the effectiveness of the questionnaire prior to its formal administration, ensuring that no questions were unclear, so that the quality of data gained from the research instrument was not adversely affected.

3.5.2 Semi-Structured Interview as a Method of Research

Interviewing, though deemed unsuitable as the main method of data collection, was effectively utilised to provide specialised information, supporting the data derived from the self-completion questionnaires. This method of data collection can be defined as “a two-person conversation initiated by the interviewer for the specific purpose of obtaining research-relevant information, and focused by him on content specified by the research objectives” (Cannel and Kahn, 1968, as quoted in Cohen and Manion, 1994, p.271). In this case, the aim of the interview was to obtain specialised information from a Department of Education Inspector that would support the data obtained from the questionnaires, thus adding another perspective to inform and enhance the research.

It was decided that a semi-structured approach to the interview would be employed, whereby “the interviewer has a series of questions that are in the general form of an interview schedule but is able to vary the sequence of the questions” (Bryman, 2008, p.196). This method was deemed most suitable as when eliciting a person’s opinion they may offer a view that corresponds to a later question on the interview schedule and so the researcher has the flexibility here to deviate from the order that questions appear on the schedule. This approach to interviewing can often lead to the generation of more complex data than perhaps a more rigid approach to the interview process, a point supported by Cohen and Manion (1994).

A semi-structured approach also allowed the researcher to formulate questions on the schedule directly relating to each research question, thus ensuring that the interview addressed areas requiring attention. In this manner, a semi-structured approach was far more purposeful than an unstructured approach, as often it can be hard to focus the interviewee on the research objectives, and thus unrelated data can dilute the relevant data obtained. In stark contrast, a highly-structured interview was also considered unsuitable, as the interviewer does not have the freedom to deviate from a set list of questions. In this case, the researcher wanted to be able to probe responses if they felt that the interviewee had not addressed the questions properly, or if they felt that the interviewee may have more to contribute on the topic of discussion.

Upon close inspection of methodological literature on this topic, it became apparent that the type of semi-structured interview employed in this research was considered an ‘Expert Interview’ (Meuser and Nagel, 1991), whereby “the interviewee is of less interest as a person than in his or her own capacity of being an expert for a certain field or activity. He or she is integrated into the study not as a single case but as representing a group (of specific experts)” (Flick, 1998, p.92). Here, the specialised data gained from the interview with one primary school Inspector is seen as representative of that which would be obtained were a number of Inspectors from the primary school sector to be interviewed, as they each have direct experience in interpreting and utilising the scores obtained from standardised tests as per Department of Education guidelines. In this regard, semi-structured interviewing was considered an appropriate method of data collection as it allowed experts in the field to provide information that contextualised the data collected from the questionnaires,

and it also enhanced the validity of the research in ensuring that all perspectives were taken into account.

Many of the disadvantages associated with this form of data collection were not applicable in this case. Firstly, owing to time constraints experienced by the participant, the semi-structured expert interview had to be conducted over the telephone, thus costs involved in physically travelling to meet with the participant did not apply. Furthermore, subconscious body-language which can often influence a participant's response was not present as the interviewer and participant were geographically separated.

However, certain limitations were experienced and addressed. Issues relating to reliability were considered, in that "the impact of the interviewer and of the context means that consistency and objectivity are hard to achieve" (Denscombe, 1998, p.137). In this case, as the interview is based upon specialised knowledge, the content of which is regulated by the Department of Education, the researcher was satisfied that the reliability of the interview was not compromised as the knowledge disseminated to all Inspectors is of a uniform nature, thus were the interview to be repeated with another Inspector the resulting data would be very similar.

3.6 The Sampling Process

As issues relating to standardised assessment in the primary school context were being examined by the research questions, teachers in these settings constituted the main focus of this study. In this manner, the sampling technique was considered 'purposive', a technique Mason defines as "selecting groups or categories to study on the basis of their relevance to your research questions" (2002, p.124). Thus, based on their relevance to the study, it was decided that primary school teachers in Galway City and County would be considered for inclusion in the research. Random sampling, whereby "each individual has an equal probability of being selected from the population" (Creswell, 2003, p.164) would not have been practical in this case, as perhaps those selected would have no experience utilising standardised assessment, and thus would be of no benefit to this study.

Bearing in mind the sheer number of schools and teachers included in the sample frame, a further selection procedure had to be introduced that would aim to target a strategic, yet representative sample, aimed at producing the greatest number of survey responses.

The Department of Education and Science primary school database 2009/10, relating to Galway City and County, was used as the initial sampling frame. In order to simplify procedures governing the multi-stage participant selection process, the researcher used himself, as a fellow educator, as a resource. Here it was decided to target schools in which the researcher had worked or in which educational personnel were known to the researcher as this would promote a favourable response rate. However, not all schools in the sampling frame were chosen in this manner as the researcher wanted to ensure that a realistic combination of school types were included in the study, again ensuring representativeness in so far as was practically possible. Thus upon completion the sample included six rural schools, five urban schools and six suburban schools. As there were significantly fewer urban schools documented in the database, compared to both the rural and suburban schools, this was reflected in the sample selected. However, two urban schools and one rural school opted out of this study due to their prior commitments to other ongoing studies in their sites. An additional suburban school was thus contacted and agreed to partake in the research. This resulted in an overall sample of fifteen schools, which entailed the cooperation of thirty participants.

As a central theme in the research focuses on accountability, it was felt that it would be most appropriate to target the two teachers from each school selected, as outlined in the Department of Education guidelines (Ireland. Department of Education and Science, 2006), who are legally obliged to send home the childrens' standardised test results. These two teachers were chosen as it is they who are most likely to experience this by-product of standardised testing, a belief further validated by the literature on this topic (INTO, 1997, 2010; Kohn, 2000; Zitlow, 2001; NCCA, 2005; Vyrostek, 2009; Mac Ruairc, 2009). This can vary from school to school as in some cases this responsibility befalls the first and fourth class teachers, in others the second and fifth class teachers. Therefore, in order to know which teachers to

include, this point was clarified when initially inviting schools to participate in the research.

When devising a suitable sampling strategy with reference to the interview, a much more simplified method was conceived. Here, working from the Department of Education and Science Inspectorate Database 2009/10, an Inspector from the Galway district was chosen at random and invited to take part in the study. This random selection procedure was deemed appropriate as it “eliminates researcher bias” (O’Leary, 2004, p.106-7), enhancing validity, and as all of those in the sampling frame were experts in the field of standardised assessment, any participant selected could have contributed as effectively as any other present on the database. Also, given that the database is so restricted in size, it was felt that the privileged information elicited from one Inspector in this regard was a sufficient representation of that which could be gained from others, thus enlisting further participants to interview was deemed unnecessary.

Given the limited size of the sample population involved in this study, the researcher is aware that the findings of this dissertation cannot be generalised beyond their means and applied to the whole primary school teaching population. However, the views and opinions of those involved in the study, together with the related conclusions drawn from careful analysis of the data collected, are nonetheless of interest and may validate the need for further research in this area.

3.7 Ethical Issues

Before the research could begin, ethical issues, such as negotiating access and ensuring confidentiality and anonymity, had to be addressed. To ensure such ethical matters were duly taken into consideration, a number of supporting documents were devised. Firstly, a ‘Letter of Invitation’ was sent to the Principal of each school selected to officially secure access to both participants in each site. This letter (included in appendix three) outlined what the research entailed and specified the research objectives so that each Principal could give their informed consent for

teachers in their school to take part in the study. Each Principal was then contacted a week after receiving the letter to confirm access had been granted.

Having gained access to participants in the above manner, ethical considerations again came to the fore when constructing supporting documentation that would be sent to those selected in each site. Ethically, it was vital to ensure that all the teachers selected to participate fully understood the nature of the study and the associated research objectives. To this end, 'Letters of Invitation' were also sent to each participating teacher, using the outline advocated by Cohen, Manion and Morrison, who state that the function of this document is to give "as much information about the aims, nature and procedures of the research as is appropriate" (2007, p.57). Thus, this letter (included in appendix four) clearly stated the aims of the research and outlined the role expected of each participant.

From an ethical perspective, it was deemed necessary to construct a 'Form of Consent' to be signed by each participant to indicate that they had read and understood the expectations and procedures involved in the study. This document also outlined the numerous benefits and any known side-effects (of which there were none) associated with participation in this study. The consent form (included in appendix five) thus embodied "an implicit contractual relationship between the researcher and the researched and will serve as a foundation on which subsequent ethical considerations can be structured" (Cohen, Manion and Morrison, 2007, p.53).

Confidentiality and anonymity were both key features in the research, predominantly in relation to the interview. Here, the Inspector's right to privacy was respected when treating all data obtained in the interview confidentially, a feature clearly outlined on the consent form. As each questionnaire was anonymous in nature (as specified in the 'Letter of Invitation') assurances pertaining to confidentiality were not required. However, in respecting each individual's right to privacy, all participants were dutifully assured that all the data obtained from the questionnaires and/or interview would be treated confidentially and that no participant would be in any way identifiable in the final report.

3.8 The Research Instruments: From Inception to Analysis

3.8.1 The Self-Administered Questionnaire

3.8.1.1. Designing the Self-Administered Questionnaire (Main Research Method)

As Oppenheim notes “Questionnaires do not emerge fully-fledged; they have to be created or adapted, fashioned and developed to maturity” (1992, p.47). In the construction of such a complex document, one encounters a number of issues requiring attention in order to ensure the effective collection of data relevant to the research questions.

Firstly, validity is a key concern when using questionnaires, whereby it is vital to ensure that each question contained in the research instrument contributes constructively towards answering the research questions at the centre of the study. As Punch notes, “the general validity question for survey questionnaire data is: Do the responses which I have, and which I will score, really measure the variables which I think they measure?” (2003, p.36). Therefore, qualitative questions were carefully constructed to elicit valuable information about each of the research questions and to facilitate the investigation of reoccurring themes (the variables) for later analysis. Knight further supports this practice stating that “Good questionnaires, like other instruments, rest on a good working theory of what ought to be explored and why it might be significant” (2002, p.93). Therefore, each question on the pilot questionnaire (included in Appendix six) was constructed specifically with validity in mind, as outlined in Table 3.1 below.

Table 3.1: Relevance of Each Question on the Pilot Questionnaire

Question on the Pilot Questionnaire	Relevance to Research Questions
Question 1	Identifying which standardised assessments were used in order to see if one set required more preparation than another. <i>(Relevant to research questions probing preparation and ‘teaching to the test’)</i>
Question 2	Identifying years experience in order to ascertain if this has any bearing on accountability experienced. <i>(Relevant to research exploring pressure on teachers)</i>
Question 3	Ascertaining how effective teachers think standardised assessment is for their classroom practice. <i>(Included to ascertain if a further research question regarding the benefits of standardised testing should be included, as denoted by level of satisfaction– consequently it</i>

Table 3.1, Continued

	<i>was decided to include this inherent feature as a separate research question after the piloting process)</i>
Question 4	Outlining pressures experienced when using standardised assessment. <i>(Relevant to research question examining pressure/ accountability)</i>
Question 5	Outlining how much time is spent preparing students for the test. <i>(Relevant to research questions exploring preparation and 'teaching to the test')</i>
Question 6	Indicating what activities are undertaken to prepare students for standardised tests. <i>(Relevant to research questions exploring preparation and 'teaching to the test')</i>
Question 7	Indicating if the participant is male or female. Does this impact their response in any of the questions being addressed above? <i>(Relevant to all research questions)</i>

Secondly, upon construction of the pilot questionnaire, it was vital to ensure that researcher bias did not impact upon the content, compromising the reliability of the research. This feature is of the upmost importance as, due to researcher bias, leading questions may be present whereby the participant answers in a manner that fulfils the “perceived expectation of the researcher” (Denscombe, 1998, p.116). Thus, each question was revised in light of this possibility, ensuring that all questions were posed objectively, thus enhancing the reliability of the data collected.

Thirdly, syntax became a vital feature requiring close attention. On this topic, Denscombe notes that “the wording of the questions is one of the most difficult features of the questionnaire design. It is also one of the most important to get right” (1998, p.98). Initially when constructing the pilot questionnaire, it was decided that the wording of some of the questions was too complex. This issue was dutifully addressed, as if a participant is not clear about what they are being asked, this may render their responses irrelevant to the study. Knight advocates that “Questions should be as short as possible, clear and in plain English” (2002, p.93). Therefore, all questions were constantly revised and simplified. Interestingly, the length of the questionnaire also comes to the fore here as when simplifying the wording of questions the researcher was also aiming to reduce to length of the overall questionnaire to promote a favourable response rate, an objective that can be difficult to achieve when employing a questionnaire approach as the length of some questionnaires can be quite off-putting for participants, a limitation previously noted.

On this matter Denscombe states that “it is worth remembering that there is, perhaps, no more effective deterrent to answering a questionnaire than its sheer size” (1998, p.96). Thus, it was decided that the overall questionnaire would not exceed one A4 page, as it was felt this would encourage participants to complete it more readily than if they were faced with a longer document entailing pages of questions.

Lastly, Knight notes that “the order in which questions are arranged can have an effect on the answers” (2002, p.94). Extensive revision took place whereby the position of each question in the layout was considered and validated, a procedure advocated by leading theorists in questionnaire design (Kiddler and Judd, 1986; Denscombe, 1998; Knight, 2002). Personal questions, such as those denoting gender, may arouse undue suspicion and thus subsequent answers may be more conservative and carefully addressed as a result. In order to avoid such unfavourable practices this question was placed at the end of the pilot questionnaire where it was less likely to affect the quality of previous answers. Oppenheim further validates this practice stating that “unless there are very good reasons to do otherwise, personal data questions should always come near the end of a questionnaire” (1992, p.109).

It was also decided to include a piece at the end of the questionnaire inviting participants to partake in an interview based on the content presented in the questionnaire. This was done in order to allow the researcher the opportunity to collect more qualitative data on the topic of standardised testing to supplement that obtained in the questionnaire, if required.

3.8.1.2 Piloting the Questionnaire

The pilot sample was constructed again using a purposive sampling technique and consisted of three female and two male primary school teachers. Only two of the participants were known to the researcher, the other three participants derived from neighbouring schools to that of the researcher. Having gained access, each teacher was sent the questionnaire along with a letter of invitation and consent form. Ethical considerations, such as those outlined in these accompanying documents, were of key importance throughout the piloting process and certainly contributed to the one-hundred per cent response rate achieved in this regard.

Each questionnaire was returned in a separate envelope to the consent form, as specified in the accompanying documentation, ensuring that the questionnaires were of an anonymous nature. Upon receipt of the questionnaires, analysis took place to assess the effectiveness of the pilot instruments.

3.8.1.3 Enhancing the Initial Research Instrument

Overall the pilot questionnaire was deemed successful in providing data that effectively addressed the research questions. However, there were a number of features that required alteration to enhance the quality of data gained from this research instrument before its final administration.

It was felt that the ‘sensitive’ question (Denscombe, 1998) regarding gender (question seven) was unnecessary as this aspect did not impact on the nature of responses obtained in the piloting phase. Also, taking into account ethical considerations, this question could possibly have served to identify those who completed the questionnaires as there were significantly less male teachers than female in the sample surveyed. For these reasons, this question was omitted from the final questionnaire instrument.

Upon analysis of the data obtained from the pilot questionnaire, it was also found that question two, relating to ‘years experience’, was redundant in providing any valuable information as attitudes and opinions were not affected by this aspect in any significant manner. Thus, in the interest of efficiency, this question was also omitted from the final research instrument.

When considering the layout of the pilot instrument, it was decided that one A4 page was simply not long enough for participants to record their opinions as many had written on the back of the questionnaire to supplement responses given. To this end, the format of the questions was carefully revised. For questions four, five and six it was deemed more appropriate to employ an alternate format, whereby the first part of each question simply involved ‘ticking the box’ to indicate a positive or negative response to the topic. The second part then invited participants to develop upon the initial stance taken in each case. Importantly, the questions were more dispersed on the final instrument, and sufficient room was left after each question to add in extra detail if needed, addressing the problem experienced in the piloting process. In this

regard, the revised layout promoted the documentation of more detailed responses. Also, the questionnaire, now comprising of one A4 double-sided page, was still an acceptable length so as not to be off-putting to participants.

Furthermore, it was decided to include another research question (question three) at this point to ascertain participants' views relating to the benefits of standardised assessment. The original focus of the research had evolved since the piloting process, as is often the case when investigating social phenomena, and it was decided that a more balanced approach to the topic was to be undertaken. Thus, in addition to the collection of data relating to possible negative consequences of utilising standardised assessment, data supporting positive experiences was also explicitly required in the final questionnaire, the outline of which is documented in Table 3.2 on the following page. Data relating to this theme was informally collected throughout the research process up to this point, yet was formalised with the inclusion of this research question hereafter.

It was also decided to leave in the piece at the end of the questionnaire inviting participants to partake in an interview based on standardised testing. Though not utilised during the piloting stage, this feature was again included on the finalised questionnaire in case more information was required after the formal administration of this research instrument, or if perhaps the researcher needed to clarify the content of any questionnaire.

The above account outlining the construction of the questionnaire from its inception to its final embodiment (included in appendix seven) clearly validates the view that questionnaires are “a very subtle and tricky type of research instrument, far more complex and difficult to design and use than they are ordinarily assumed to be”. (Kidston, 1985, p.150).

Table 3.2: Outline of Final Questionnaire Research Instrument

Question	Question Content	Question Format	Linking to Research Question exploring:
1	Identifying which tests were utilised	Tick the box, linked to question 4.	Accountability
2	Judging how effective standardised tests are in indicating achievement	Scaling of opinion, linked to question 3.	Benefits of Standardised Assessment
3	Outlining the benefits of using standardised assessments	Open-ended	Benefits of Standardised Assessment
4a	Identifying if teachers experience pressure when using this form of assessment	Tick the box, Yes/No	Accountability
4b	Elaborating on the answer given in 4a	Open-ended	Accountability
5a	Identifying how much time is spent preparing students for the tests	Open-ended	‘Teaching to the Test’
5b	Identifying which activities are undertaken when preparing students for the test, if preparation occurs	Open-ended	‘Teaching to the Test’ and Classroom Practice
6a	Ascertaining if standardised assessment impacts upon classroom practice.	Tick the box, Yes/No	Classroom Practice/ Benefits of Standardised Assessment
6b	Elaborating on the view expressed in 6a	Open-ended	Classroom Practice/ Benefits of Standardised Assessment

3.8.1.4 Administering the Qualitative Questionnaire

Upon gaining access to participants in the fashion outlined previously, each school was contacted to elicit which two teachers in that site sent home the standardised testing scores formally each year to parents. Each of these teachers were then sent a package during the last week in February containing the letter of invitation, the consent form and the self-administration questionnaire. They were also provided with two envelopes so that the completed questionnaire and consent form could be sent back separately, thus ensuring the participant’s anonymity.

An eighty-three per cent response rate was achieved, whereby twenty-five of the thirty questionnaires were completed and returned. Each questionnaire was then transcribed, facilitating later analysis. This high response rate can be partially attributed to the purposive selection of participants. As Babbie notes, “If a high response rate is achieved, there is less chance of significant non-response bias than with a low rate” (2010, p.272). Therefore, the high response rate achieved can be viewed as an important attribute supporting validity in the research.

3.8.1.5 Analysing the Data from the Questionnaire

Each questionnaire was assigned a serial code and carefully transcribed. Serial codes such as ‘Q1, Q2, Q3, Q4’ were employed signifying the order in which each questionnaire was received, thus allowing the researcher to match each original questionnaire to its transcribed version. Once all questionnaires were transcribed, the researcher then explored the data obtained in a thematic fashion. Themes identified within the data were assigned unique ‘codenames’, allowing the researcher to flag similar data in other questionnaires and to differentiate between the various themes and sub-themes evident in the overall data set (All codenames utilised in this process are included in appendix eight). As Babbie notes “the aim of data analysis is the discovery of patterns among the data, patterns that point to theoretical understanding of the social world. The coding and relating of concepts is key to this process” (2010, p.400).

Once all the data obtained from the questionnaires was thematically coded in the manner outlined above, a number of Microsoft Word documents existed that embodied each category and its inherent sub-categories within the data, a sample of which can be seen in appendix nine. The data in each category was then compared to that in other categories, in order to identify if any interconnections existed, a practice advocated by Babbie (2010) above. This analytical method, defined as the Grounded theory model (Glaser and Strauss, 1967), more comprehensively outlined in the next chapter, allowed the researcher to make connections between themes in the data and explore the influence each category exerted in such relationships. This activity thus led to the generation of theory relating to standardised testing, the results of which are documented in the following chapter.

Furthermore, following analysis on the wealth of qualitative data obtained from the questionnaires, it was decided that interviewing teachers further on the topic, a possibility suggested on the questionnaire, would not be required.

3.8.2 The Interview Schedule

3.8.2.1 Designing the Interview Schedule

The process involved in designing an effective interview schedule was much less complex than that outlined above when designing the questionnaire. As the interview schedule was designed after the questionnaire, the researcher was already aware of many of the difficulties that can be encountered when designing an objective research instrument, particularly those involving validity and reliability, which aided the researcher in this task.

As the researcher would be interviewing an Inspector, time would undoubtedly be a constraining factor as these educational personnel are few in number and are assigned large areas to manage. Therefore, the interview schedule had to make efficient use of the time available for interviewing. In this regard, the four specific research questions were studied and corresponding questions were carefully constructed to address each area, as documented in Table 3.3 on the following page. In this manner, the researcher was “translating the research objectives into questions that will make up the main body of the schedule” (Cohen, Manion and Morrison, 2007, p.356). A further question was included that aimed to elicit the Inspector’s ‘privileged knowledge’ on the topic of standardised assessment (question two). Such information, inaccessible to the mainstream teachers who completed the questionnaires, was of vital importance in obtaining all relevant perspectives. The format of the questions on the interview schedule was open-ended, which would allow the researcher to probe responses and develop concepts, if necessary. Closed-ended questioning would have been unsuitable when collecting experience-rich data, as by definition closed-ended questions demand that “answers fit into categories that have been established in advance by the researcher” (Denscombe, 1998, p.101). However, when engaging in a qualitative approach to research it is impossible to pre-assign categories, as the researcher cannot predict the nature or content of the responses that the participant will provide. Thus, open-ended questioning was

adopted, whereby the participant was allowed the freedom to express their views, later categorised according to the data gained from the interview.

Table 3.3: Questions contained on the Pilot Interview Schedule

Question Number	Wording of Question	Linking to Research Question exploring:
Question 1	What are the benefits of standardised testing?	Benefits of Standardised Assessments
Question 2	For what purposes are the results of standardised assessments used by the Inspectors?	Accountability (Specialised Knowledge)
Question 3	Do you feel that standardised assessment can create pressure for teachers?	Accountability
Question 4	In your opinion, does standardised assessment impact classroom practice?	Classroom Practice
Question 5	Do you think there is a 'teaching to the Test' culture present in Ireland?	Teaching to the Test

The number of questions to be included on the interview schedule presented an obvious issue. By focusing the content of the interview schedule on areas directly related to the research questions, the researcher was further ensuring that the interview does not “get lost in topics that are of no relevance and permits the expert to extemporize his or her issue and view on matters” (Meuser and Nagel, 1991, p.448). Here, the interview focused on what was most relevant while also allowing ample opportunity for the Inspector to talk about their area of expertise.

It was decided to discuss the positive attributes of standardised assessment prior to exploring the more questionable features. On this topic, research has proven that “it is important to ease into main questions and themes. If you start off with ‘a sensitive’ question or one that might be considered threatening, you may find yourself facing an up-hill battle” (O’Leary, 2004, p.168). Thus, by initially asking more general questions, a positive and trusting rapport was developed prior to engaging in a more sensitive line of questioning.

Validity and Reliability were again features that were of the upmost concern. The wording of all questions was objective, as previously evidenced in Table 3.3, and thus ensured that interviewer bias did not affect the reliability of data gained from the

interview. It was noted that validity could have been affected by the researcher's own background and interest in the area. Here, Blaxter, Hughes and Tight note "the very choice of the subject or objects of your research will influence your findings" (1996, p.76). By investigating standardised testing, a form of assessment the researcher has experience utilising, personal beliefs led the researcher to include participants in the study, such as the Inspector, that would produce beneficial data for the research, and thus the question of validity comes to the fore. However, it is my belief that as an experienced practitioner in this form of assessment I will be able to identify and appropriately address any discrepancies in the data thus, in effect, enhancing the validity of the research.

3.8.2.2 Piloting the Interview Schedule

Ideally, it was envisaged that the interview schedule would have been piloted on an Inspector, as it was constructed with this aim in mind. However, constraining factors were experienced in this regard. As there are a very limited amount of Inspectors working in the area, the researcher did not feel it would be wise to exhaust the possibility of procuring one such Inspector for formal interviewing purposes until every effort was made to develop the interview schedule, thus maximising the efficient use of time permitted for the interview. Therefore, it was decided to pilot the interview schedule on an experienced teacher, known to the researcher, who has an extensive background in utilising standardised assessment and would therefore be professionally competent in answering the questions posed on the interview schedule, with the obvious exception of the question probing the Department of Education's use of standardised test scores.

Having been granted access by the participant to conduct the pilot interview, a suitable time and venue was then agreed upon by both parties. The interview, recorded using a small hand-held digital device, took place in the interviewee's home and lasted approximately nine minutes long. Throughout the interview the researcher was very conscious to act in an objective manner so as not to influence the interviewee's responses. Upon completion of the interview the data recorded on the digital recording device was transcribed, allowing for post-interview analysis.

3.8.2.3 Assessing the Effectiveness of the Pilot Interview Schedule

Upon examination of the data obtained, it was decided that the interview schedule (included in appendix ten) was very effective in providing data directly relevant to the research questions. Therefore, it was decided that this research instrument did not require any adjustments and so the pilot interview schedule was thus deemed suitable to collect data in a formal capacity.

3.8.2.4 Conducting the Interview

An Inspector was selected at random from the sampling frame and contacted by telephone to see if they would partake in the study. Having gained access in this manner, it was planned to conduct a face-to-face interview. Yet, due to time constraints experienced by the Inspector, a telephone interview was deemed more appropriate for use in this case. The participant was then contacted to organise a time that would be suitable to conduct the interview.

Prior to the interview the participant was also provided with the interview schedule. This practice has many advantages. By providing the participant with the interview schedule prior to the interview, they have time to reflect on their answers and thus will provide more in-depth data (Denscombe, 1998) Also, they are less likely to digress from the topics included as they are aware of the interviewer's areas of interest, thus the interview progresses more efficiently. This feature is further highlighted by O'Leary, when stating "If you have a limited amount of time [...] you will want to make sure you are keeping your interviewee on track and moving at a decent pace (2004, p.168). Also, when presented with the questions prior to interviewing, less probing can be required as participants are more likely to offer in-depth opinions as they have been allowed time to reflect on the topic and do not feel pressurised or intimidated into answering immediately. This feature is especially advantageous when researching controversial or sensitive topics, such as standardised testing, as probing can sometimes be perceived as invasive and thus the participant may be less likely to offer their true opinion.

The interview was successfully conducted, via telephone, on 3 February 2010 at 5pm. The interviewee participated in the interview from their home. All areas of focus on the interview schedule were addressed by the participant and thus probing

responses was not required. The researcher acknowledges that providing the participant with the interview schedule prior to the interview most likely resulted in this advantageous aspect. The completed interview, lasting approximately seven minutes, was recorded using the digital device previously employed in the piloting stage.

3.8.2.5 Analysing the Data from the Interview

Once completed, the interview was fully transcribed. The transcription was then sent to the interviewee for verification, and to see if they wanted to change or omit any of the information previously provided. In this case the interviewee did not make any such changes and so analysis of the interview could proceed immediately. In this regard the researcher employed the approach advocated by Denscombe (1998), the layout of which is outlined in appendix eleven. Here, having transcribed the interview using Microsoft Word, each line is assigned a number “so that parts of the data can be identified and located precisely and quickly” (Denscombe, 1998, p.130). Two columns are then added to the left hand margin of the page to aid in the analytical process.

The first column allowed the researcher to document codes. Here, the researcher identified snippets of data in the Inspector’s responses that matched the criteria of categories previously used when analysing data from the questionnaires. Codenames were then documented in this column to allow the researcher to promptly locate material specific to a certain theme when comparing data obtained from the research instruments.

The second column was included in order to document any notes the researcher wished to make relating to data obtained from the interview, such as those documenting the relationship between data obtained from the interview with that obtained from the questionnaires. The results of such analysis are documented and discussed in the next chapter of this dissertation.

Chapter 4: Data Findings, Analysis and Discussion

4.1 Introduction

Having collected data using the methodology outlined in the previous chapter, the focus of the research then progressed to analysing this data and interpreting the findings in order to better understand the current position of standardised testing in the primary school classroom, as indicated by those participating in this study. The main themes, as well as the numerous sub-themes, identified and explored in this chapter directly link to specific research questions and thus a strategic approach to data analysis was adopted in order to effectively address and answer each research question, a process outlined in Table 4.1 below.

Table 4.1: Outlining the Manner in which Each Research Question was Addressed

<i>Research Question</i>	<i>Associated Theme Identified in the Data Collected and Analysed:</i>	<i>Sub-themes Identified, aimed at answering the Research Questions:</i>
<i>Research Question 1:</i> What are the benefits of standardised testing?	‘Benefits of Standardised Testing’	<ul style="list-style-type: none">• Indicates Achievement.• Facilitates National Comparisons• Diagnostic/ Screening Value.• Provides Parents with Evidence of Achievement.• Other Miscellaneous Benefits.
<i>Research Question 2:</i> How does accountability affect teacher’s attitudes towards standardised testing?	‘Accountability/ Sources of Pressure’	<ul style="list-style-type: none">• Parents• Reflection of Own Teaching.• Other Teachers.• Curriculum Overload.• Educational Management.

Table 4.1, Continued

<u><i>Research Question 3:</i></u> In what ways are children prepared for standardised testing?	‘Preparation’	<ul style="list-style-type: none"> • Procedural Revision. • No Revision. • General Revision. • Directed Revision.
<u><i>Research Question 4:</i></u> What are the impacts on classroom practice as a result of the adoption of standardised testing in our schools?	‘Impact on Classroom Practice’	<ul style="list-style-type: none"> • No Impact. • Impact on Pre-testing Activities. • Impact on Post-testing Activities.

It was also decided to adopt a strategic approach in the layout of this chapter. In this regard, each theme is examined separately, allowing the researcher to explore inherent sub-categories more comprehensively. Data provided by participants in the self-administration questionnaire is supplemented with the views expressed by the Inspector, and further contextualised with literature deriving from both a national and international perspective. Thus, applying this approach, theory based on standardised testing was carefully formulated within the confines of this study and is dutifully discussed in this chapter.

Throughout this chapter when referring to data obtained from participants in the study a logical approach is adopted, as devised by the researcher. The letter ‘R’ is used inside the brackets when referring to a response provided by a participant on the survey instrument, together with the corresponding participant number. Thus, (R4) refers to the response obtained from the participant who completed the fourth questionnaire returned from the thirty sent out. When referring to content from the Inspector’s interview, the word ‘Interview’ appears inside the brackets, together with the corresponding line numbers from the interview from which the data was derived. In this regard (Interview, lines 4-5) indicates that the data being discussed was obtained from lines four and five of the transcription of the Inspector’s interview.

4.2 Grounded Theory Approach to Data Analysis

Many of the techniques associated with the ‘Grounded Theory’ approach to data analysis were employed in this study. Developed by Glaser and Strauss (1967), Babbie defines Grounded theory as “an inductive approach to research [...] in which theories are generated solely from an examination of data rather than being derived deductively” (2010, p.396). Here, the researcher’s professional observations formed the initial interest in the area of standardised assessment, in which subsequent research aimed to derive objective data whereby common themes and patterns could be identified, ultimately facilitating the development of informed theories. Interestingly, Grounded Theory is not a theory in itself, rather its inherent features lead to the generation of theory through careful and objective analysis of the data collected.

Open Coding, a central process in Grounded Theory, was effectively used to organise the data obtained in this study. Charmaz clarifies that “codes [...] serve as shorthand devices to *label, separate, compile and organise data*” (1983, p.186, emphasis in original). When adopting a qualitative approach, as is the case in this study, data that has been collected is broken down into a series of statements and subsequently sorted into categories, each assigned a distinctive codename to differentiate between the topics being discussed. For example, in this study codenames such as ‘General Information’, ‘Benefits’, ‘Pressure’, ‘Preparation’, and ‘Impact on Classroom Practice’ were assigned to indicate what the data in each category was referring to. Importantly, categories such as these “depict the problems, issues, concerns and matters that are important to those being studied” (Strauss and Corbin, 1998, p.14).

Axial Coding, an additional feature of the Grounded Theory approach, was also widely utilised throughout the course of this study and can be defined as “the appreciation of concepts in terms of their dynamic interrelationships. These should form the basis for the construction of theory” (Goulding, 2002, p.78). Here, data that was broken down during open coding, as outlined above, was then reassembled in order to formulate meaningful categories that facilitated the researcher in identifying relationships between categories. Strauss and Corbin explain that the data is

“reassembled through statements about the nature of relationships among the various categories and their subcategories” (1998, p.103). In this process it is common that a category may contain numerous sub-categories that directly relate to the main category “giving it further clarification and specification” (Strauss and Corbin, 1998, p.101), a feature that can be seen when examining all major categories in this study. Interestingly, throughout this entire multi-stage process, the properties that define each category are in constant flux, altering their dimensions to include further related data, or indeed restricting their scope if it is felt that a group of data could be more appropriately re-coded as a separate category in its own right. This process allows the researcher to ascertain how strongly pieces of data within a category connect, and thus the extent to which they are connected to other categories can be more critically assessed.

Constant comparison constitutes a fundamental process associated with Grounded Theory, as advocated by Babbie (2010), and was also used extensively throughout the entire analysis stage in this study. Here, the researcher “explicitly compares each incident in the data with other incidents appearing to belong to the same category, exploring similarities and differences” (Spiggle, 1994, p.493-4), as well as comparing such ‘incidents’ with those documented in related literature to make sense of and contextualise their presence. Applying this feature to this study, the sub-categories identified in section 4.5 relating to the pressures felt by teachers when exploring the theme of ‘Accountability’ can be seen to directly influence data contained in other categories, whereby the pressure experienced by teachers can lead to the adoption of ‘result boosting’ practices, which in turn can impact negatively upon general classroom practice, a connection further explored in later sections. Relationships identified from such analysis were then dutifully explored in literature based on standardised assessment in order to identify the prevalence of these associations, both nationally and internationally.

Therefore, it is when we investigate the interconnections between and within such thematic categories that theory relating to standardised testing is meaningfully formed, a journey embodied in the following sections of this chapter.

4.3 Standardised Testing at Primary Level

This section aims to explore the types of Standardised tests currently in use within the primary education sector, as well as examining teachers' attitudes towards the effectiveness of standardised testing. Such information serves to inform and supplement subsequent findings relating to the main themes identified in the data.

4.3.1 Standardised Test Instruments Selected

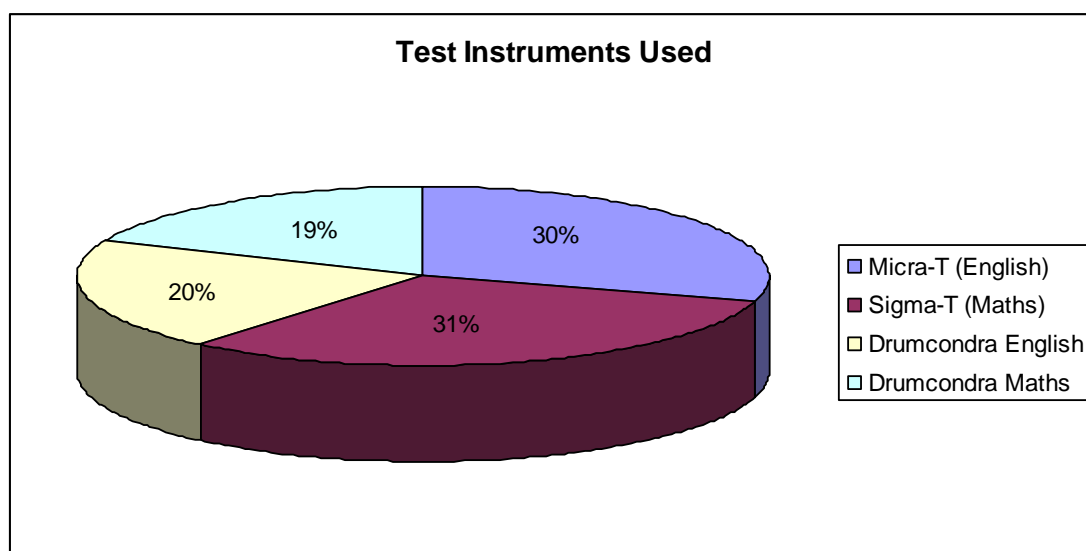


Figure 4 1: Test Instruments Used

Figure 4.1 above outlines the test instruments currently in use in the primary schools of those surveyed. The Sigma-T Maths and Micra-T English series (Wall and Burke, 2004, 2007) produced by CJ Fallon Publishers and the Drumcondra Tests produced by the ERC (2006, 2007) are currently the two main standardised assessment series' available. From the above information, one can clearly ascertain that the Sigma-T Maths assessments was more highly favoured than the Drumcondra Maths assessments, with twelve per cent more of the sample population opting to use the Sigma-T test instrument. In like manner, the Micra-T English assessment was used by ten per cent more of the sample population than was the case with the Drumcondra English assessment. Therefore, one can confidently state that the Sigma-T/ Micra-T series (chosen by 61% of those surveyed) were the leading standardised assessment instruments used within the context of this study, a tendency more comprehensively addressed when discussing conclusions based on this data in the next chapter.

4.3.2 Perceived Effectiveness of Standardised Testing

It was decided to include a question on the questionnaire probing teachers' attitudes in relation to how effective they felt standardised assessments were in indicating a child's level of achievement so as to contextualise subsequent data obtained relating to the main themes identified in this study. The results of this enquiry are clearly visible in Figure 4.2 below.

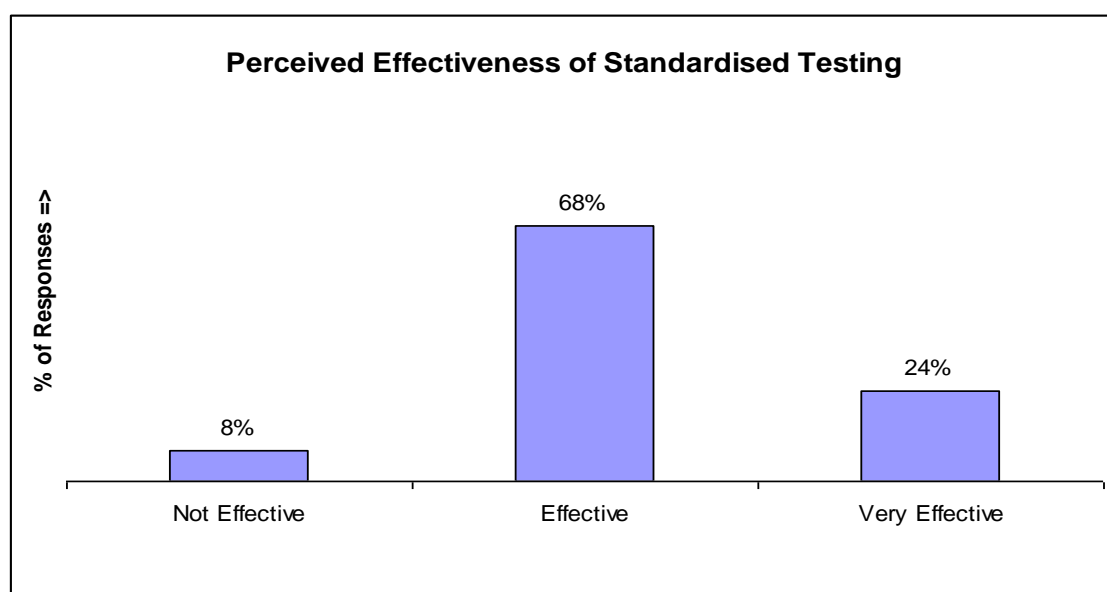


Figure 4.2: Perceived Effectiveness of Standardised Testing

From the above representation it is clearly discernable that the majority of those surveyed (68%) felt that standardised testing was 'effective' in indicating a child's level of achievement. Those who responded in this vein are confirming that the test instruments are successful yet have limitations, as indicated by the participants' reluctance to define the tests as 'very effective'. This finding would support the stance that teachers are utilising these tests in context. In this regard, the NCCA state that "it is important to use a range of assessment information when making decisions about a child's progress and achievement" (2007, p.61), rather than merely relying on a single standardised test score. Paralleling this belief, one participant notes that "this type of test is of use only in some aspects of assessing a child" (R21), while another participant further states that standardised assessments are "only worthwhile when used in conjunction with teacher observation and teacher designed tests/tasks" (R2). Such comments highlight teachers' professional practice in this area, whereby

in recognising the tests limitations they are utilising the assessments in the manner envisaged by their creators (Wall and Burke, 2004, 2007; ERC, 2006, 2007).

However, a minority of those surveyed (8%) felt that the tests were ‘not effective’, directly contrasting with a significant number of participants (24%) defining standardised assessments as ‘very effective’. Though at opposing ends of the spectrum, both perspectives are undoubtedly of interest as when one is considering ‘overall effectiveness’ it can be ascertained that the vast majority of participants (92%) rated the tests either ‘effective’ or ‘very effective’. This indicates that generally those utilising standardised testing appreciate the positive contribution that this form of assessment can make in the classroom.

Furthermore, no significant correlation between perceived effectiveness and the test instrument chosen was identifiable in the data collected, therefore one cannot concisely state that one set of test instruments are necessarily more effective than another. It was however noted that ‘not effective’ responses were solely confined to the Sigma-T/ Micra-T series of assessments, yet within the confines of this study these constituted a mere eight per cent of the responses obtained and were furthermore in stark contrast to the twelve per cent of participants who defined the same series of assessments as ‘very effective’. Therefore, no significance can be drawn from this finding as on average the Sigma-T/ Micra-T series was considered ‘effective’ by those utilising this series in their primary school classrooms.

4.4 Theme 1: Benefits Associated with Standardised Testing

Figure 4.3 below depicts the sub-categories identified within the data obtained, providing a clear indication of the various benefits those surveyed have associated with this form of assessment, thus directly answering the research question posed on this topic.

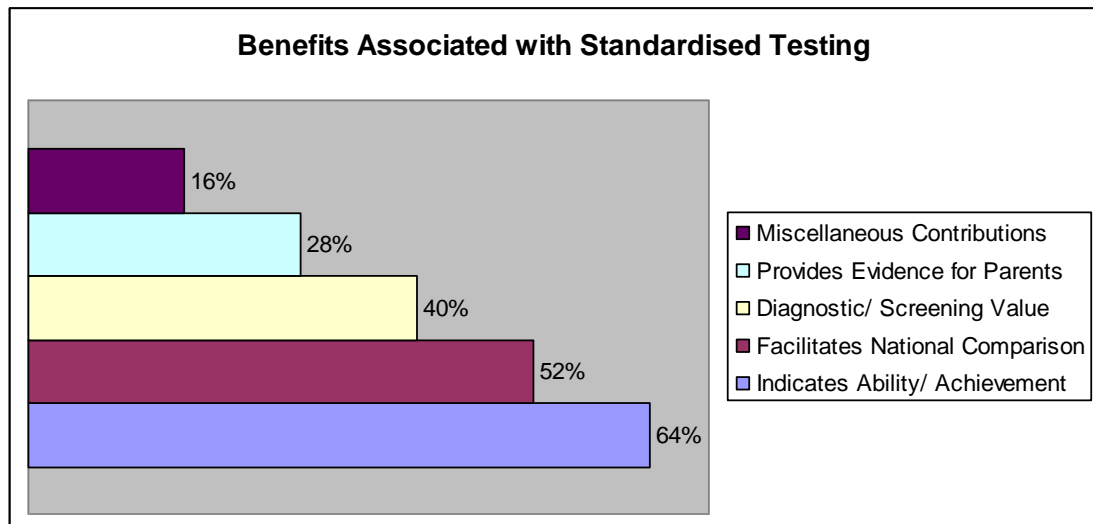


Figure 4.3: Benefits Associated with Standardised Testing

4.4.1 Indicating Achievement (64%)

From careful analysis of Figure 4.3 above, one can clearly ascertain that the leading benefit associated with standardised assessment is its inherent ability to indicate achievement. A significant sixty-four per cent of participants explicitly identified this benefit in their response on the questionnaire, supporting the test creators' belief that these assessments "can enable teachers to obtain a clear and comprehensive account" (Wall and Burke, 2007, p.1) in relation to their pupils achievement. Therefore, teachers' attitudes in this area, embodied in responses such as "Gives the child's level of achievement on a national scale" (R22) and "Good reflection of the pupils' achievement" (R25), are in line with those envisaged by the test creators (Wall and Burke, 2004, 2007; ERC, 2006, 2007). Data obtained from the expert interview further supports this finding, whereby the Inspector explicitly promotes the view that standardised testing "does help you find out how your children are getting on" (Interview, line 14).

When considering the key role assigned to standardised assessment in primary school, one participant interestingly notes that "Our education system in Ireland is based on how well a pupil performs on paper and on a particular day so standardised assessments prepare pupils for what lies ahead" (R3). Here, this participant states that the test's role in indicating achievement furthermore introduces pupils to the format of formal testing that they will experience throughout their schooling career. However, O'Leary (2006) contradicts this view, questioning if perhaps the test's

ability to outline achievement, though clearly beneficial when used in conjunction with other sources of information, is being over-relied upon in such cases to provide a comprehensive evaluation of pupil ability rather than achievement.

4.4.2 National Comparisons (52%)

Over half (52%) of those surveyed agreed that a major benefit experienced when utilising standardised assessment is that it allows for national comparisons of student achievement to be undertaken. This is further evidenced in some of the qualitative responses obtained, with one participant noting that standardised testing “does give an indication of how a child is performing in relation to the rest of their peers nationally” (R13), while another participant clearly states that it “profiles students against their peers nationally” (R7). Such data suggests that teachers are aware that standardised testing affords them the opportunity to place their class’s learning on a national scale, while also allowing them to ascertain where further consolidation of concepts may be necessary in order to ensure that their class is on par with their peers nationwide. This perspective is further supported by the Inspector, who defines this feature as the “the biggest benefit of standardised testing” (Interview, lines 12-13). Thus, agreement is evident in relation to this beneficial feature among educational practitioners, and is further reinforced by literature in this area stating that standardised testing can be used to compare children “to other children throughout the country at the same class level or age level” (NCCA, 2006, p.1), thus differentiating standardised testing from other forms of assessment. Furthermore, the test creators themselves explicitly publicise this widely-known feature, stating;

The chief advantage of nationally standardised norm-referenced tests, such as the Sigma-T, is that they almost uniquely allow teachers to compare any individual pupil’s mathematical performance with standards of mathematical attainment nationally.

(Wall and Burke, 2007, p.1)

Thus it is apparent that this beneficial aspect clearly promotes the continued use of standardised assessment as no other form of testing currently available affords educational personnel the opportunity to track a child’s achievement nationally to the same extent as the standardised testing model, as noted by Wall and Burke (2007) above.

4.4.3 Diagnostic/ Screening (40%)

Forty per cent of participants noted the tests ‘Diagnostic/ Screening Value’ when identifying the benefits associated with standardised testing. Here participants state that standardised tests can be “very helpful in confirming children who need extra help” (R15), as well as advocating that the tests “pinpoints specific areas where children have difficulty, allows for plans to be put in place for children that need extra support” (R17). Here, one can determine that teachers can use the assessments to identify areas of the curriculum that children may have problems with, as well as using the results to help identify/ confirm where children may need added support, a practice further promoted in literature relating to this topic (NCCA, 2005, 2007). Interestingly, when commenting on this matter, the Inspector states that

Generally speaking they are used as screening results for learning support. Now, I have always been keen to point out to teachers that that is not the purpose of standardised testing, and while that might be one of the main ways in which they are used, it is only one of the ways that they should be used”

(Interview, lines 48-52)

Here, we can identify conflict in the data between ‘professional’ practice and ‘common’ practice. Professional practice would advocate that no single standardised test result should be taken in isolation when considering a child for inclusion in a learning support programme, as outlined by Wall and Burke (2007) below, yet, as the Inspector highlights, this questionable practice does occur and is furthermore viewed as a benefit, as opposed to a harmful by-product. In support of the Inspector’s concerns, two participant responses from the ten included in this category explicitly commented upon learning support allocation, with one participant stating “We use it to decide on learning support priorities the following year” (R10). In such instances the standardised assessment results are being used out of context, as clearly outlined by the test creators who dutifully highlight that “for a comprehensive appraisal of a child’s achievement level, the results ought to be combined with other strands of evidence” (Wall and Burke, 2007, p.18). However, upon close inspection of the questionnaire data, such over-reliance does not appear to be commonplace in the sample surveyed in this study, with eighty per cent of responses in this category not containing any reference to learning support at all.

4.4.4 Providing Evidence for Parents (28%)

Over a quarter (28%) of participants highlighted the standardised tests' role in providing evidence for parents in terms of their child's yearly achievement. Comments such as "Can be beneficial to support or provide evidence to parents" (R2) and "They are a good source of evidence to provide the parents with" (R20) denote teachers' belief that the tests are useful in providing objective 'evidence' to parents regarding their child's achievement. This view strongly correlates with that outlined in the literature relating to this beneficial feature. Here, the NCCA explicitly promote the communication of standardised test results "to report to parents on their children's achievement and progress" (2007, p.61), a procedure formalised in the Education Act (Government of Ireland, 1998). This view is also supported by those who constructed the tests, whereby the test results are seen as measures that can "prove especially helpful in communicating with parents" (Wall and Burke, 2007, p.1) when discussing their child's progress, along with other contributing sources of information.

However, the findings obtained in this study also suggest that one must be careful in their application of this beneficial feature in practice. Comments such as "parents seem to accept a standardised result more readily than just 'teacher' opinion" (R6) highlights how teacher professionalism can be undermined when such results are over-relied upon, a belief further supported by O'Leary (2006) whereby he argues that other less formal methods of assessment are also integral components in assessing achievement levels in the classroom.

4.4.5 Other Noteworthy Benefits Identified

Additional singular comments were also noted by participants when outlining the benefits associated with standardised testing. Such comments included the use of standardised testing "for revision purposes and preparation for entrance exams" (R8), as well as those stating that such tests are "free from teacher bias/ favour" (R1). Though not relevant to previously defined sub-categories, such contributions were informative in considering all perspectives relating to the benefits experienced when using this form of assessment.

As noted by R1 above, these tests are “standardised norm-referenced tests” (Wall and Burke, 2007, p.1), ensuring that teachers cannot in any way influence the content of the test or the manner in which they are scored, and so the tests are an objective measure of achievement, aimed at supplementing the accumulated portfolio of information relating to a child’s progress that teachers compile throughout the year.

Participant R8 above alludes to the tests use in revising and preparing for entrance exams. Though the format of the test instruments may entail a general overview of the curriculum, it is not the purpose of these assessments to be used in a revisionary capacity, especially prior to administration which is generally the time of year when most secondary school’s entrance exams would take place. This dubious feature is clearly denounced by the test creators who vehemently state “‘teaching to the test’ [...] will render the results entirely invalid” (Wall and Burke, 2007, p.5). Although the findings derived from this study contradict views expressed in the literature here, though only relating to one participant, such instances are important to identify and analyse so that procedures governing ‘best practice’ can be ascertained.

4.5 Theme 2: Accountability

This section aims to answer the research question relating to accountability when ascertaining if teachers feel pressure when utilising standardised testing in their site of practice, and if so, from what sources this pressure derives.

4.5.1 The Presence of Pressure

Data obtained from the self-completion questionnaires indicates that over twice as many teachers (68%) experience pressure when using standardised testing than those who do not (32%). Thus, within the context of this study, the findings imply that the majority of teachers do experience pressure when using this form of assessment, a view also held by many researchers on this topic (Turner and Clift, 1988; Neill and Medina, 1989; Bartlett, 2000; Abrams, Pedulla and Madaus, 2003; Hursh, 2005; NCCA, 2005; Mac Ruairc, 2009; Vyrostek, 2009; INTO, 2010). However, this is a view that the Inspector did not share. When discussing this topic the Inspector stated “I have never felt in any school I have been in that it has been a pressure for teachers,

no” (Interview, lines 89-90). Although it is evident that there are conflicting perspectives here in terms of the majority of teacher responses obtained and that of the Inspector, when one probes this conflict it becomes apparent that this variance may simply be due to a lack of communication by both parties relating to standardised testing. In this regard, it was noted that both parties seem to be constructing attitudes towards this form of assessment somewhat autonomously, without clear knowledge of the other party’s practice, a point reiterated by Mac Ruairc (2009)

The various sources of pressure relating to standardised assessment identified by participants in this study are illustrated in Figure 4.4 below, and are further explored in this section.

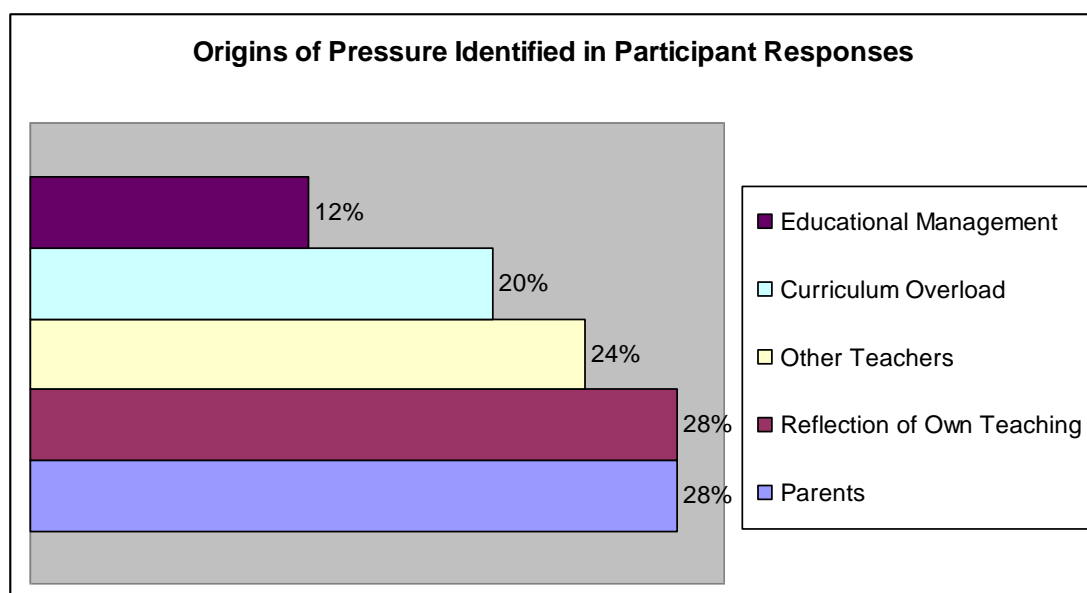


Figure 4.4: Origins of Pressure Associated with Standardised Assessment

4.5.2 Parental Pressure (28%)

Results indicate that twenty-eight per cent of the questionnaire respondents felt that ‘Parents’ constituted one of the leading pressures in relation to standardised assessment. Here, the data obtained suggests that teachers felt they were held responsible by parents if their child did not perform well on the standardised test or if their score decreased from the previous year, due to lack of parental understanding surrounding the process of standardised assessment. In this regard, one participant notes “Parents use it as a way of judging teachers” (R24), a view further supported

by Posner, who states that the competence of a teacher “is being measured by students’ scores on standardized achievement tests” (2004, p. 749). Here, Posner is highlighting the parental tendency to equate a teacher’s professional efficacy in terms of how well their child performs on these tests. The findings of this study further support Posner’s view, with one participant noting “If their child’s score is lower than the previous year they can see this as a reflection of the teacher” (R13). In such cases parents are identified as an external source of pressure when undertaking this form of assessment.

Literature emanating from the NCCA supports this finding, advocating that there needs to be “a public information campaign explaining standardised testing in accessible terms for parents” (2005, p.11) so that misinterpretation and over-reliance on the results of this form of assessment are minimised, thus decreasing pressure experienced by teachers in this area. As noted by one participant, “some parents may not be aware that standardised testing is only a fraction of the assessment procedure. They read/ take standardised test results as the only indication of progress” (R22). Work has certainly been initiated in this area, for example the availability of information for parents on the NCCA’s website, yet the comment outlined above would certainly suggest more awareness on this topic needs to be raised.

An interesting point to note here is that the sub-category ‘Parents’ appears in both the ‘Sources of Pressure’ category and the ‘Benefits of Standardised Testing’ category. Here, both positive and negative connotations are assigned to the same sub-theme, presenting a conflict within the data. However, having worked extensively with the data obtained in this study, it is this researcher’s belief that this conflict can be viewed merely as a matter of context. As the data obtained suggests, when used appropriately, standardised assessments can constitute a very useful source of supplementary evidence to present to parents regarding their child’s progress, as outlined in section 4.4. Yet, when used out of context, parents lack of understanding can create pressure for teachers who recognise that “if their child’s score drops from the previous year they can see this as the fault of the teacher so some teachers feel pressurised to have high scores” (R19), a concept explored above. Interestingly, the interview with the Inspector did not provide any data relating to this sub-category.

4.5.3 Reflection of Own Teaching (28%)

Of equal concern amongst the respondents surveyed in this study was the belief that the test results were viewed as an indication of their own teaching ability.

Importantly, this sub-category can be viewed as producing both internal and external pressure, as explored below.

When considering internal accountability, one participant notes that “if a pupil/pupils drop in score then you feel very responsible and exposed” (R3). Here the data would indicate that a drop in standardised test score has a negative impact upon a teacher’s self-image and self-worth. Another like-minded participant notes “teachers feel it is a reflection on their teaching. Of course it pressurises teachers to get good results” (R8). Here again the data suggests that some teachers perceive their pupils’ scores as being intrinsically linked to their professional ability, whether internally by that teacher themselves or externally by others. Recent literature would concur with this finding, stating that “some teachers were of the view that the results of standardised tests bore a reflection of themselves” (INTO, 2010, p.114). Thus, evidently, teachers can place significant pressure on themselves to achieve good results in their class, thereby believing that this will reflect favourably on their professional image. Consequently, this can result in the adoption of educationally questionable practices, such as ‘teaching to the test’, in order to raise pupils’ scores, a linkage in the findings explored later in this chapter.

External pressure in this area can also be identified in the data collected in comments such as “If a high achiever performs poorly on the day of the test for whatever reason, some may feel it was a reflection of the teacher” (R16). Here, the data suggests that external sources can view a standardised test score as a reflection of the teacher’s ability. This feature can be viewed as an unintended by-product of the ‘medium stakes’ testing climate present in Ireland, as discussed in chapter two, whereby stakeholders other than the teacher and pupil view the standardised test results and make judgements upon them. This belief is further supported in literature exploring the link between accountability and the medium-stakes testing situation currently evident in Ireland (NCCA, 2005). One such stakeholder in this situation is the Inspector who uses the test results to monitor standards in their assigned jurisdiction and to look for trends in the data obtained. In supporting the findings

obtained on this matter, the Inspector interviewed commented that low pupil performance “could highlight problems with the teacher” (Interview, line 56). Here, the data analysed confirms that a teacher’s ability can be viewed by those monitoring standards as a contributing factor to class performance, and so the somewhat ambiguous knowledge of this practice can create pressure for teachers, whereby they want to ensure that their class performs favourably, thus reflecting upon their teaching ability positively (Mac Ruairc, 2009). Ex-Minister for Education, Noel Dempsey, further reinforces this link when stating that “standardised test results play an essential role in the evaluation of the work of schools” (2004). Thus, the findings obtained in this area clearly parallel the those outlined in the literature, indicating that teachers can experience external pressure to maintain/ raise standardised test scores in order to ‘prove’ their ability as a teacher to those from the Department of Education who can view their pupils’ standardised assessment scores as a contributing factor towards whole-school evaluation.

4.5.4 Peer-Teacher Pressure (24%)

Interestingly, twenty-four per cent of those surveyed indicated that ‘peer-teacher pressure’ also constituted both an internal and external source of accountability. Here, the data indicates that the scores achieved by a teacher’s class are often directly compared with those obtained the previous year, and may also be viewed by their teacher the following year. Therefore, some teachers experience accountability in terms of ‘measuring up’ to the scores obtained from previous years, and are acutely aware that the results their class obtain in the standardised tests may be viewed by the next teacher of that class as an indirect indication of their professional ability. In support of this finding, one participant confirms that external pressure does exist “to ensure that a class that has a high standard maintains it from one year to the next and from one teacher to the next” (R20). Here, we can identify concern being expressed in the data relating to the accountability experienced by teachers when trying to maintain high standardised test scores, a worrisome situation referred to in a discussion document published by the INTO (2010).

Furthermore, another participant adds that “a teacher can also feel pressure from other teachers in the school because you don’t want to have low results when other teachers have higher results because they have revised with the children for the test”

(R13). In a similar vein of reasoning to the last participant quoted, this participant identifies ‘peer-teacher pressure’ as an internal source of pressure, as when placing an over-importance on achieving high test scores in their class to match those obtained by other teachers in the school, increased accountability is internally experienced. In this manner this sub-theme can be seen to directly relate to the ‘Reflection of Own Teaching’ sub-theme previously explored, as other’s perceptions can create internal pressure in terms of a teacher’s awareness that their class’s standardised test scores can constitute a reflection of their teaching by those misusing this form of assessment. Also, of particular note in participant R13’s comment above is the identification of ‘revision’ as a means to enhance the scores obtained. Such activities, more comprehensively explored in section 4.6, are explicitly opposed by the test creators, as previously outlined, and are undertaken with a view to ‘boosting’ the scores of children in the class to decrease the pressure experienced by the class teacher in this regard, a finding further supported by Abrams, Pedulla and Madaus who state that:

Researchers have cautioned that placing a premium on student test performance can reduce instruction to test preparation, thus limiting the range of educational experiences to which students are exposed and minimising the skill that teachers bring to their craft.

(2003, p.20)

Here, Abrams, Pedulla and Madaus (2003) acknowledge that test preparation can be undertaken by teachers in a response to increased accountability being experienced when teachers are engaging in ‘competitive’ testing, as can be observed in participants’ comments above. Interestingly, data relating to this sub-category is lacking from the Inspector’s perspective, which would suggest that the Inspector is unaware of this source of pressure.

4.5.5 Curriculum Overload (20%)

‘Curriculum Overload’ constituted a significant sub-category within the data, with twenty per cent of participants referring to this feature in their responses. Comments denoting “pressure on teachers to get all the work that will be coming up in the tests covered by May” (R25) and those indicating the perceived “need to cover the entire curriculum before testing” (R10, emphasis in original) typify responses that constitute this sub-category. Here findings based on the data suggests that teachers

feel they must teach as much of the curriculum as possible before administering the standardised assessments, a situation that creates pressure as when teaching an expansive curriculum this proves no easy task.

The curriculum itself states that it is “characterised by its breadth and balance. [...] and is directed towards the development of the full potential of every child” (Government of Ireland, 1999, p.10). However, as indicated above, when aspiring to reflect this ‘breadth and balance’ in their instruction, it would seem that some of the teachers surveyed felt that they were already being over-stretched in terms of the time available to comprehensively engage with such a curriculum. This finding broadly agrees with literature in this area warning that “Curriculum backwash may occur: that is, test content may have an impact on teacher practice” (Gipps, 1988, p.31). Here the literature suggests that in some cases non-tested curricular areas may be omitted due to time constraints experienced by teachers under pressure, and in doing so “the taught curriculum contracts to conform to the assessed curriculum” (INTO, 1997, p.19), as a result of this source of pressure.

4.5.6 Educational Management (12%)

In this final sub-category twelve per cent of teachers felt that educational management, namely Principals and/or Inspectors, constituted a source of pressure for teachers when administering standardised assessments, as it was felt that both parties view the results with high expectations relating to performance, not necessarily correlating with the more realistic measures of achievement obtained.

In the case of the Principal, some of those surveyed felt that s/he wants “to see good results, to show that the school has a high achievement level” (R8). Here, the data clearly suggests that some teachers feel directly accountable to the Principal, as it is expected that their class will uphold the high level of achievement that reflects favourably on the school. Logically, such expectations can result in the formation of pressure, particularly for teachers who are legally obliged to communicate the results home to parents, as a Principal’s focus is more likely to be drawn to their scores rather than others in the school as they are aware that parents can view such results as a measure of ‘school effectiveness’. International findings discussed at the Consultative Conference on Education in 2008 further support this finding, stating

that “a teacher’s professional reputation, prestige or even promotional prospects or the enrolment levels and future viability of a school may be effected by the results of assessment” (INTO, 2010, p.84). Taking this international perspective into consideration, it is evident that not only teachers may experience accountability, as Principals too may well feel accountable to the public at large, as well as the Department of Education Inspectorate, to maintain high standards previously achieved in the school. This pressure can then transfer to the teachers working in the school when endeavouring to fulfil such high expectations.

Importantly, data relating to Inspectors was also included in this sub-category. Here, findings derived from careful analysis indicate that some of those surveyed felt that “Inspectors [...] also look at results and I think sometimes they see a classes’ overall performance as a reflection of the teacher’s ability” (R13). This participant highlights the belief that the scores obtained by one’s class can act as a measure of ‘teacher effectiveness’. Clearly, one can see a direct link emerging here between the data within this sub-category and that contained in the ‘reflection of own teaching’ sub-category. In both cases, teachers can feel directly accountable for their class’s performance, yet here the Inspector is specifically cited as the source of pressure. This particular source is also highlighted by Mac Ruairc who stating that “the use that is made of test results [...] within the Department of Education and Science has yet to be fully determined” (2009, p.48). Here, Mac Ruairc alludes to the fact that the test scores are collected and utilised by personnel from the Inspectorate in a manner undisclosed to the teaching community. Ambiguity surrounding this situation can thus create pressure for those who wish to ensure that the test scores achieved reflect favourably on their professional practice.

Interestingly, data obtained from careful analysis of the Inspector’s interview on this topic validates such concerns. The Inspector notes that the standardised testing scores of a school are “always looked at” (Interview, line 24), in the proper context, and “if a trend seems to go down in one particular year it could highlight problems with the teacher for example” (Interview, lines 55-56). Here, the Inspector confirms that “standardised tests contribute to the evaluation of schools by the Inspectorate” (NCCA, 2005, p.4), further adding that low performance by children in the test may be attributed to low ability of a teacher. The findings in this area thus suggest that

this association results in the formation of pressure for those utilising this form of assessment. When discussing this feature with teachers Turner and Clift (1988, p.160) also note that many feel “there’s a growing feeling amongst our members of staff of ‘Big Brother’” as the results obtained by one’s class may be directly or indirectly used to appraise their professional competence. Therefore, the findings obtained in this area are in line with the literature, indicating that teachers, having a somewhat vague knowledge of this aspect of standardised testing, can experience significant pressure in this regard.

4.6 Theme 3: Preparation

This section explores the methods of preparation employed by teachers when getting ready for the standardised assessments, as well as examining how much time is spent engaged in such activities. Furthermore, this section links with the theme of accountability previously explored, whereby the various methods of preparation undertaken are adopted as a direct result of pressure to boost/maintain high standardised testing results, a stance further supported by Cizek (2001).

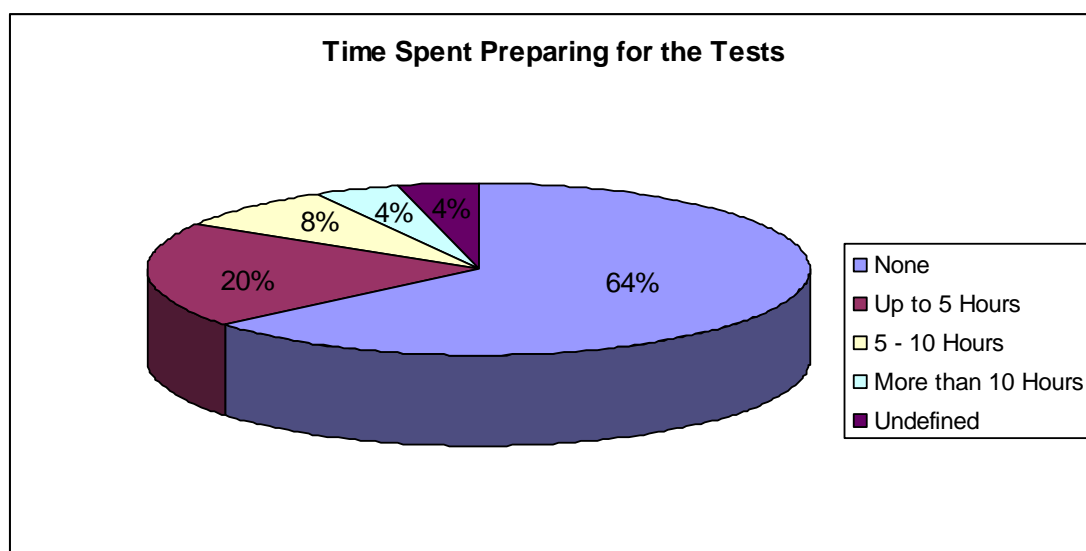


Figure 4.5: Time Spent Preparing for Tests

4.6.1 Indicating Time Spent Preparing for the Tests:

As can be seen in Figure 4.5 above, sixty-four per cent of those surveyed indicated that no preparation for the tests was undertaken. Vehement comments such as “We

don't teach to the test!" (R9) and "I refuse to 'teach to the test'" (R2) were noted, indicating that participants were astutely aware of the negative connotations associated with this activity, whereby "'teaching to the test' in order to boost test scores will render the results entirely invalid" (Wall and Burke, 2007, p.5). Here, we can identify that teachers' professional conduct in this area parallels that envisaged by the test creators, suggesting that standardised assessments are being utilised in a manner befitting their purpose.

However, over one-third (36%) of participants did prepare their pupils for the standardised tests. Here, twenty per cent of teachers estimated that they spend up to five hours involved in preparation, with a further twelve per cent indicating that preparation activities undertaken exceeded this sub-category, to varying extents. Therefore, it is vital to explore what this time is spent doing, together with the information provided by those who indicated that 'no time' was spent preparing yet go on to provide subjective data regarding how they 'ethically' prepare students for the test. This anomaly in the data may suggest that some of those who indicated that they do not spend time preparing for the test equated the word 'preparation' on the questionnaire with 'teaching to the test'. However, as is explored in Figure 4.6 below, the data collected clearly indicates that several forms of test preparation, both ethical and otherwise, exist.

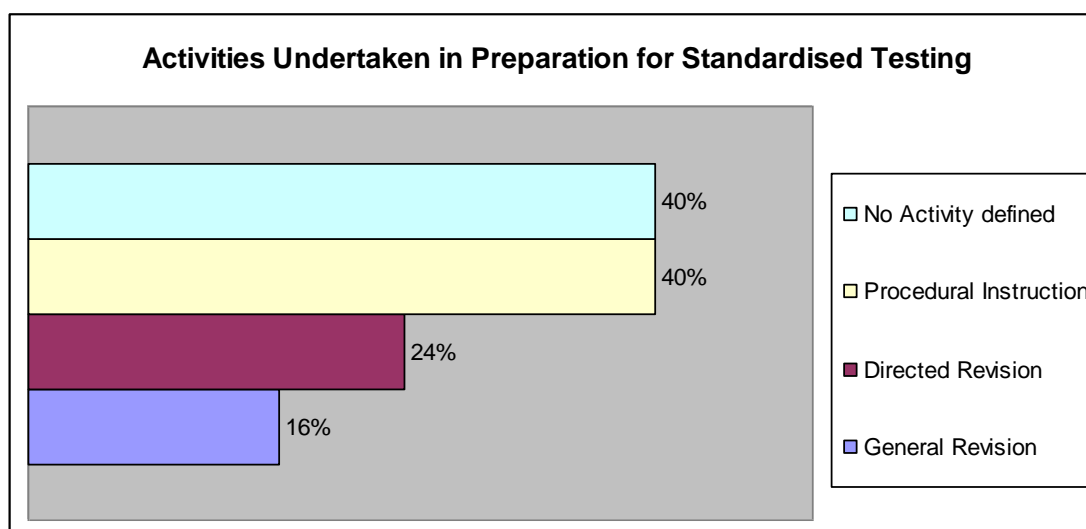


Figure 4.6: Activities Undertaken in Preparation for the Test

Firstly, the data depicted in Figure 4.6 clearly shows that a significant number of participants did not define any activities undertaken as they had previously stated that no time was spent preparing for the tests, a logical correlation presented in the data here. However, those that did identify preparatory activities provided rich qualitative data predominantly in three specific areas, each explored below. In some cases, two or more sub-categories were identified in the same response obtained.

4.6.2 Procedural Instruction (40%)

‘Procedural Instruction’ constituted a major category within the data relating to the theme of ‘Preparation’. Here, the findings indicate that forty per cent of those surveyed described how children were prepared ethically in relation to the structure of the test and the rules governing testing-taking in general, the efforts of which do not have any bearing on the results obtained. This is a practice further advocated by Haladyna, Bobbit-Nolen and Haas (1991). When outlining such methods, participants note “I explain the details as per test instructions and that’s it” (R15), “Throughout the year we practice the skills needed for tests, i.e. skip a question, look back, guess, double check” (R3) and “I find it helpful to practice cloze tests because this is the format of the tests and it’s not in their books really” (R20). Upon careful analysis of the above comments one can see that teachers are equipping their students with the skills they need to complete the standardised assessments, but are in no way addressing the content matter of the test. Therefore, the results obtained will not be artificially inflated as the students are not being explicitly taught areas of the curriculum that feature on the test instruments, as is the case when ‘teaching to the test’. Thus, such ethical preparation serves only to boost a child’s confidence and lower anxiety levels without positively affecting their score. Furthermore, the testing manuals themselves advocate that “it is important not to arouse undue anxiety in pupils” (ERC, 2007, p.4), thus indicating that teachers who engage in procedural instruction are utilising standardised assessments appropriately, ensuring that anxiety levels are lowered before test administration. Haladyna, Bobbit-Nolen and Haas comment upon such “training in testwiseness skills” (1991, p.5), further supporting the implementation of this practice as it serves only to develop a child’s general test-taking skills without affecting their result.

4.6.3 Directed Revision (24%)

‘Directed Revision’ constitutes the next most prominent sub-category within the data, with twenty-four per cent of the participants referring directly to this activity in their responses on the questionnaire. Responses typically associated with this method of revision include “I revise key areas in maths that come up on the test” (R13), as well as “Giving examples similar to the test questions, revising different aspects of the maths curriculum that may come up on the test” (R25) and undertaking “a revision of the language and terminology in the test” (R8). The findings here clearly indicate that nearly a quarter of the teachers surveyed in this study are undertaking revision activities that directly enhance student performance in standardised tests, a detrimental practice frequently highlighted in literature in this area. (Fontes *et al*, 1980; INTO, 1997, 2010; Cizek, 2001; Cimbricz, 2002; Hursh, 2005; NCCA, 2005, 2007; Mac Ruairc, 2009). Of note in this regard is the admission by some of those surveyed that “I do not teach to the test” (R13 quoted above for example). However, these teachers go on to describe ‘acceptable’ activities employed that ironically do serve to partially ‘teach to the test’, artificially enhancing childrens’ scores in any case. In this regard, the data suggests that ‘acceptable’ practice is an ambiguous aspect that can, in some instances, create a situation whereby “the test becomes the curriculum” (Neill and Medina, 1989, p.694).

A finding of particular note in this subcategory is the admission by some of those surveyed that they have experienced ‘teaching to the test’ in their site of practice, though not engaging in this activity themselves. In this regard, participant R15 comments “It’s important that teachers do not ‘teach to the test’. I have experienced this in the past where this happened and results were distorted”. Furthermore, participant R22 highlights a teacher in their site of practice who engages in this activity, thus indicating their aversion to this practice. In such instances, the data collected confirms the presence of this negative by-product of standardised testing within the sample surveyed.

When discussing this theme Cizek comments that there is a growing trend of “educators themselves attempting to subvert accountability systems by artificially inflating student test scores” (2001, p.1). Here, Cizek notes that as a result of increased accountability, as discussed in section 4.5, teachers can deliberately engage

in educationally questionable practices to enhance their test results, thus reflecting favourably on their teaching ability. Here, the extent to which a teacher experiences pressure can directly result in the adoption of result-boosting activities, such as ‘teaching to the test’ which, again, highlights the link in this study between this category and that of ‘Accountability’.

Interestingly, when discussing this topic the Inspector notes that “You know that teachers are actually ‘teaching to the test’. I have no doubt some individuals are doing it perhaps with a view of puffing up their own results [...] but I have no evidence of it” (Interview, lines 112-114). Here, findings based on the data objectively obtained in this study confirms the Inspector’s belief whereby the responses given on questionnaires explored above clearly outline activities undertaken by some teachers that serve to enhance a child’s performance on standardised assessments beyond that which they would have been capable of achieving unaided.

4.6.4 General Revision (16%)

The final sub-category constructed from the data pertaining to the theme of ‘Preparation’ is that entitled ‘General Revision’. Here, the data collected indicates that general exercises in revision were carried out by sixteen per cent of the teachers surveyed in the time leading up to the administration of standardised tests, a practice strongly opposed by Wall and Burke (2004, 2007) and the ERC (2006, 2007). While the findings indicate that no material on the actual test instruments was directly examined, as is the case with ‘teaching to the test’, general areas of the curriculum were nonetheless addressed, some of which may appear on the test. Thus, literature in this area situates ‘General Revision’ at the lower end on a continuum of preparatory activities that can be undertaken, “done without regard to a particular test” (Bond, 2004, p.1), yet such activities can still distort the results obtained on standardised tests, providing an inflated indication of student achievement.

During analysis of the data it was noted that teachers were prompt in clarifying that “revision is not done specifically for the test but a recap is done on curricular areas which have been covered” (R17). Here the data indicates that teachers recognise the dangers of directly ‘teaching to the test’ in relation to result distortion, yet do not

deem general revision inappropriate in this regard, even-though an indirect advantage is being given to those who have ‘revised’ and the curriculum becomes increasingly based on that which is tested, a stance further supported by Turner and Clift (1988).

When discussing this controversial sub-theme, the Inspector also notes that “you can never be certain that the test has been administered correctly” (Interview, lines 28-29), thus further supporting the findings of this study whereby a number of ethical and unethical activities have been outlined when ascertaining the methods by which children are prepared prior to the tests. Interestingly, the findings of Cimbricz’s insightful study concur with those noted here. In her research based on standardised assessment, forty per cent of teachers did not recognise similar revision activities as “a narrowing of the curriculum, a sacrifice to instruction or a change to their instructional methods overall” (2002, p.10). Here we can see a direct link between the findings of this study and those obtained by Cimbricz, whereby some teachers do not view altering their professional practice and focusing on general revision as a contributing factor that may result in artificially boosting standardised test results and can, in effect, limit the curriculum to a series of consolidatory exercises prior to test administration. Neill and Medina further support this finding, summarising that “as teaching becomes ‘coaching for the test’ in too many schools, real learning and real thinking are crowded out” (1989, p.694).

4.7 Theme 4: Classroom Practice

This section examines if teachers feel the adoption of standardised assessment in their classroom influences their professional practice, and if so, in what manner.

4.7.1 Ascertaining if Classroom Practice is Affected

Initially, fifty per cent of participants clearly indicated that standardised testing does influence their professional practice, with a further fifty per cent stating that it did not. However, careful analysis of the open ended question that followed this initial indication, the results of which are depicted in Figure 4.7, found that in actual fact sixty per cent of those surveyed provided information relating to pre-testing and/or post-testing impacts, in comparison to forty-eight per cent of participants denoting that no impact was experienced. This variance in the data suggests that teachers do

not experience impacts relating to standardised testing in a uniform manner in their classrooms.

When exploring this topic, Abrams, Pedulla and Madaus strongly advocate that “Regardless of one’s position on this issue, it is impossible to deny that statewide testing policies influence classroom practice and student learning” (2003, p.18) which may indicate that some of those who stated no impact was experienced in this regard may have been mistaken. Yet, within the confines of this study, one must be careful not to place interpretations upon the data beyond its means.

However, regardless of one’s position on the matter, along with the Inspector, the majority of participants, even those who indicated that no impact on classroom practice was experienced, provided rich qualitative data when justifying their stance on this topic. The results of this analysis are documented in Figure 4.7 below.

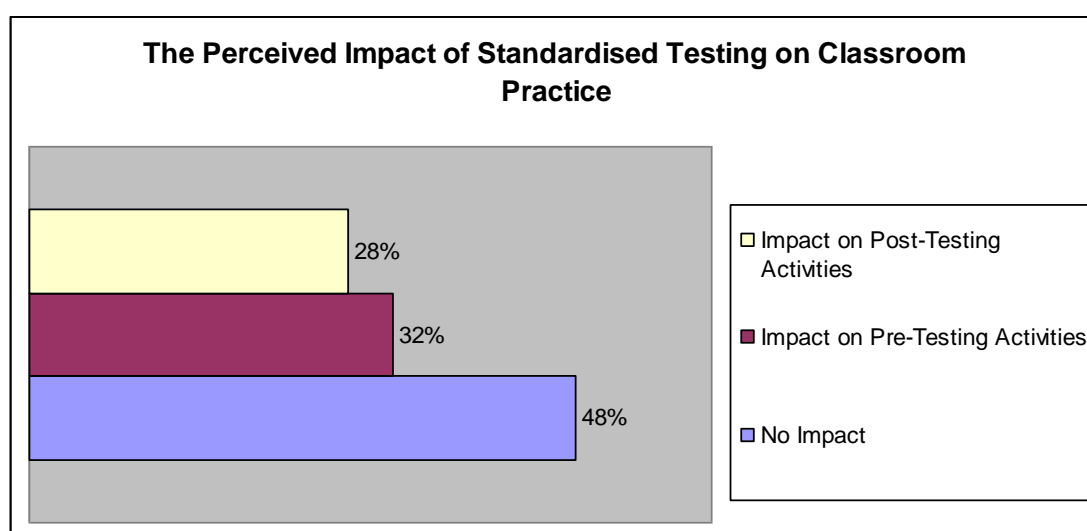


Figure 4.7: The Perceived Impact of Standardised Testing on Classroom Practice

4.7.2 No Impact on Classroom Practice (48%):

It is clear that the leading sub-category defined on this topic is that of ‘No Impact’, with just under half the participants (48%) providing responses relating to this sub-theme within the data. Here participants are vehement in their views that standardised testing should not impact on their classroom practice, suggesting that the lack of impact in this area is more the result of deliberate effort as opposed to indirect outcome. When discussing this feature one participant notes that “it in no way influences my classroom practice” (R12), while another strongly states “it

should not influence the teaching in the class” (R9). Logically, here the data suggests that classroom practice neither benefits nor suffers as a result of standardised testing, as no impact is defined. However, such views logically imply that when no impact on classroom practice is experienced the concept of testing becomes redundant, as it serves no significant purpose (Vyrostek, 2009).

This perspective is also shared by the Inspector who notes that “it is a way of finding out how children are getting on and [...] should inform the teaching more” (Interview, lines 80-81). Here, the Inspector supports Vyrosek’s stance, stating that the results of standardised assessments should influence teaching more as it is a valuable tool, together with other forms of assessment, that can help a teacher to identify areas that may require attention and devise a programme suitable to address such matters. Literature constructed by the NCCA further validates the Inspector’s belief, stating that teachers can use standardised testing to “identify children with learning difficulties so that appropriate supports can be put in place” (2007, p.61) as well as “to identify children with exceptionally high scores so that appropriate learning experiences can be provided for them” (2007, p.61). Here, we can see the positive influence that standardised assessment can have on classroom practice in providing a supplementary and objective perspective of achievement that can serve a teacher to further support their students and their learning. Thus, one must question if perhaps those who stated that no impact was experienced are utilising the assessments to their best effect.

However, within this sub-category, insightful observations validating participants’ negative responses were also provided. One participant notes that “not all children perform well on these tests due to a variety of factors” (R18), while another comments “I’ve learnt that teaching is not all about standardised tests and they are just a small fraction of what one can do in order to assess children and their learning” (R23). Findings based the data provided here highlights the cautious manner in which teachers approach the assessments, considering these assessments as a separate educational tool not affecting their classroom practice. In this regard, when identifying the possibility that children can underperform on these tests due to a variety of reasons, as well as the appreciation of the dangers associated with solely relying on a single test score, as noted above, those surveyed are indicating that they

are being careful not to let such negative by-products of standardised testing pervade their classroom practice. Such comments also confirm that teachers' appraisal of the negative by-products of standardised assessment are generally in line with current literature exploring this theme (Wall and Burke, 2004, 2007; NCCA, 2005, 2007; ERC, 2006, 2007). However, though admirable in their efforts to negate the effects of such negative by-products, these teachers must also ensure that their practice is benefitting from the numerous benefits previously explored, otherwise undertaking standardised testing in such cases is of little or no use to a teacher's professional practice.

4.7.3 Impacts Identified on Pre-Testing Activities (32%)

Thirty-two per cent of participants provided responses relating to this sub-category, (comprising of only one more participant than in the Post-Testing sub-category below), creating somewhat of a balance in the data provided by participants from both the Pre- and Post-testing perspective. Findings derived from the data provided in the pre-testing context indicate that revision seems to be commonplace amongst those who responded in this sub-category. One participant notes "Time is spent revising before the test" (R19), while another comments that "I spent some time doing general revision, would never 'teach to the test' though" (R20). Here one can identify that general revision activities are undertaken prior to test administration, as specified in sub-section 4.6.4. As noted by R20 above, this activity does not directly constitute 'teaching to the test' as previously defined by Wall and Burke (2007), insofar as curricular areas included on the test instrument are not being selectively taught. However, such activities can still serve to boost test scores and thus distort the results of standardised testing, a point posed by those who have devised the tests (Wall and Burke, 2004, 2007; ERC, 2006, 2007). As noted by the Inspector, one can never be sure that the tests have been administered correctly (Interview, lines 28-29), and in instances where teachers are revising general concepts before administration, as outlined in the findings here, it would seem such concerns are validated. Furthermore, when engaging in these revisionary exercises, literature would suggest that "as a result, a great deal of time is devoted to preparing for assessment, and overall educational quality suffers" (INTO, 2010, p.26). Logically, here the INTO are advocating that revision-style activities take up a considerable amount of time

leading up to the assessments which could have been spent developing more holistic skills, as opposed to concentrating solely on academic material.

A further finding derived from this sub-category highlights the belief held by some participants that time should be spent preparing the children to cope with the format of standardised assessments. In this regard, one participant notes “Prior to the test I work on cloze procedures to familiarise the children with this procedure” (R5), while another states “I would do extra comprehension work before testing” (R10). Here the data suggests that ethically sound activities are also undertaken prior to testing so that children will not become anxious when faced with exercise formats that they have not encountered before, a method of ethical preparation previously outlined in sub-section 4.6.2 that has clear links with the findings obtained here. As specified, such activities do not give the child an unfair advantage, as they are not consolidating prior learning of concepts included on the test. Nor are the children being directly taught items on the actual standardised test instrument. Rather, they are practising the skills they will need to understand what they are being asked to do and how to address questions on the test appropriately. Haladyna, Bobbit-Nolen and Haas support this practice and further define such ethical activities, stating that “training in testwiseness skills includes familiarizing students with the formats of answer sheets and test items and general strategies for optimum performance” (1991, p.5). Here, the literature supports the findings obtained in this study, clearly specifying the benefit of such pre-testing activities for children undertaking standardised assessments.

4.7.4 Impacts Identified on Post-Testing Activities (28%)

This sub-category, directly contrasting with that outlined above, documents the *post*-testing impacts standardised testing can exert on classroom practice. Here, findings obtained from the careful analysis of data indicate that twenty-eight per cent of those surveyed altered their professional practice based on the results obtained by pupils in their class. Such impacts were highlighted in comments such as “If a group shows a large discrepancy in computation skills one may spend added time on this area of the curriculum” (R11), “the scores can also identify if a child may need more help in a certain area in Maths/English” (R19) and “I would go back and revise any area that proved difficult for the children in the test” (R15). Here the data clearly implies that

post-testing consolidation of the concepts children found challenging in the test is undertaken, indicating that the tests are being used in a constructive manner, informing teaching and enhancing the quality of learning in a classroom. Thus, the findings of this study support O'Leary's belief that "assessment is integral to good teaching and learning" (2006, p.15).

Moreover, the data collected signifies that if a child's score is quite low in the standardised test, and is consistent with their previous class-based tests, this may indicate that the child may be a candidate for extra learning support. This post-testing impact is also identified by the Inspector, who states that "they do help you to find out how your children are getting on number one, but also of course the whole area of screening them for extra support" (Interview, lines. 13-15). After testing, the Inspector envisages that a teacher would be able to gauge how their children are 'getting on' and the score obtained could also provide a valuable objective perspective to supplement the teacher's own observations, together with an accumulated pupil assessment portfolio, in judging if a child would benefit from extra support being provided for them. Circular 0138/2006 further supports this finding when stating "There is widespread acceptance of the value of standardised testing as one of a range of modes of assessment that help teachers make more informed decisions" (Ireland. Department of Education and Science, 2006, p.1), particularly when considering candidates suitable for learning support. Here, the literature devised by those charged with implementing standardised assessment in primary schools explicitly advocates that the scores of such assessments should have an impact on classroom practice, supplementing other 'modes' of assessment, and aiding the teacher in ascertaining if a child may require added support. Clearly then, the data collected in this study correlates with proposed 'best practice' guidelines (Wall and Burke, 2004, 2007; ERC, 2006, 2007; NCCA, 2007), suggesting that teachers experiencing such impacts are utilising standardised testing in a professional manner, allowing it to inform their classroom practice positively in supporting the learning needs of those in their care.

4.8 Conclusion

Upon careful consideration of the findings presented in this chapter it is apparent that teachers' attitudes towards standardised testing generally concur with those presented in literature based on this form of assessment. However, instances of conflict, though few in number, were nonetheless noted and addressed.

Furthermore, considering the analysis and discussion documented in this chapter it is clear that the research questions posed from the outset, and subsequently expanded, have been appropriately addressed and answered when exploring the various sub-themes identified in the data obtained, as outlined in Table 4.1. Conclusions resulting from this process are dutifully explored in the next chapter.

Chapter 5: Conclusions and Recommendations

This chapter presents key conclusions based on the findings and discussion explored in the previous chapter. Furthermore, recommendations in this area are also included in this chapter, as well as highlighting areas that would benefit from future research.

5.1 Conclusions

5.1.1 Standardised Testing in Ireland

Firstly, the most obvious conclusion evident in this regard is that there is a severe lack of research undertaken to date regarding standardised testing in the Irish context. The limited existing body of Irish literature predominantly derives from the NCCA (2005, 2007), INTO (1997, 2010) or the test creators themselves (Wall and Burke, 2004, 2007; ERC, 2006, 2007). Such publications, mostly offer advice on standardised testing and allude to the problems experienced in other countries when using this form of assessment. Yet, a very limited number of directed studies (Kelly, 2007; Mac Ruairc, 2009) have been undertaken to explore and document standardised testing in the Irish situation relative to those undertaken from an international perspective.

Secondly, when focusing upon the data collected in this study, it became evident that the majority of teachers (68%) utilising standardised assessments found the tests ‘effective’, with a further twenty-four per cent considering the tests ‘very effective’. Thus, it is apparent that the vast majority (92%) of teachers generally viewed the assessments in a positive light.

Thirdly, when examining which test instruments were selected, a tendency amongst those surveyed towards choosing the Sigma-T/ Micra-T test series was noted. This tendency may possibly be attributed to the fact that the Sigma-T/ Micra-T pupil booklets are shorter in length than the Drumcondra pupil booklets and contain colour images, a feature lacking in the Drumcondra series. Thus it is possible that the Sigma-T/ Micra-T series may be considered more ‘child friendly’. This conclusion is

reiterated by those marketing the standardised testing instruments when stating “Distinguished by their attractive layout and high colour content, the tests are visually appealing to pupils, a factor that is important in securing pupil interest” (C.J. Fallon Publishers, 2010). Also, as colour is used extensively to animate the test instruments, both the Sigma-T and Micra-T pupil booklets contain a short colour blindness test, an advantageous feature not included in the Drumcondra test instruments.

5.1.2 Theme 1: The Benefits Associated with Standardised Testing

Conclusive data obtained in this study indicated that the main benefits associated with this form of assessment include the tests’ ability to measure achievement, facilitate national comparisons, diagnose/ screen pupils and to provide objective evidence to parents regarding their child’s achievement.

Furthermore, it became evident that those surveyed were generally aware of such beneficial features yet adopted a cautious approach when utilising standardised assessment. Those who highlighted the tests contribution in the area of diagnosing/ screening, for example, were clear that care must be taken when interpreting and utilising the pupils’ scores, a view best practice guidelines would support (Wall and Burke, 2004, 2007; NCCA, 2007; ERC, 2006, 2007; INTO, 2010). In this regard, the INTO note that, “standardised tests should not be accorded undue importance in making such decisions for a comprehensive appraisal of a child’s level” (2010, p.83). Therefore, this cautious approach, though not affecting the aforementioned beneficial features, ensures that the tests are being used in an appropriate manner, and thus one must conclude that generally teachers’ professional practice in this area is highly commendable.

However, less prominent conclusions in this area also indicate that instances do exist whereby the tests’ function is being distorted when used as a revision instrument for secondary school entrance exams, or simply as a general revision aid, due to the inherent breadth in test content. Though in the severe minority, teachers engaged in this practice are ultimately distorting subsequent test results and undermining the potential contribution of standardised assessment in their classrooms.

5.1.3 Theme 2: Accountability

Conclusions derived from the data in this area confirmed that the majority of those surveyed (68%) experienced pressure when utilising standardised testing, a belief that concurs with studies undertaken on this topic (Gipps, 1988; Turner and Clift, 1988; Neill and Medina, 1989; Halaydna, Bobbit-Nolen and Haas, 1991; Bartlett, 2000; Abrams, Pedulla and Madaus, 2003; Posner, 2004; NCCA, 2005; Mac Ruairc, 2009; Vyrostek, 2009, INTO, 2010). Logically, as a direct result of experiencing such pressure, the majority of teachers' attitudes towards standardised testing, specifically pertaining to accountability, can be generally viewed as 'negative' in light of this finding.

Careful analysis of the data in this area specified that the origins of this pressure can include; parents, reflection of own teaching, peer-teacher pressure, curriculum overload and educational management, each resulting in the formation of the informed conclusions outlined below.

Firstly, it became evident that 'Parents' constituted a significant source of accountability, as the scores obtained by a teacher's class can be seen as a reflection of their professional ability due to parents' lack of understanding of the standardised testing process. This conclusion is also highlighted by Posner (2004) and further serves to explain why over a quarter (28%) of those surveyed felt that parents constituted a major source of pressure, generating negative attitudes amongst teachers in this area. Interestingly, teachers' concerns relating to this parental theme were seen to directly link with those expressed when examining the sub-theme 'Reflection of own Teaching', and were equally as pertinent with the same percentage of participants explicitly identifying this particular source of accountability. Here the data suggests that teachers can view standardised testing in a negative fashion owing to the tendency by some to internally link the scores obtained by their class to their own teaching ability. This conclusion was also discussed at the Consultative Conference on Education 2008 (INTO, 2010), and can be seen to relate to the above parental theme, as not only the teacher themselves may make this misinformed association, parents too may formulate a misplaced link between the teacher and the test scores obtained. This conclusion thus clearly highlights the

manner in which negative attitudes towards standardised testing may be socially and/or self-constructed by teachers.

The identification of peer-teacher pressure, as indicated by twenty-four per cent of the participants, was a surprising finding as one would expect that those who are utilising standardised testing would be aware of the complexities inherent in the administration and interpretation of the test results obtained from this form of assessment. However, conclusions in this regard suggest that some teachers are not utilising the tests in an appropriate manner, opting instead to use the standardised test results as a measure of teacher efficacy, thus generating negative attitudes towards standardised assessment amongst the teaching community in a school in this regard.

Curriculum overload was cited by a fifth of those surveyed as an origin of accountability, as these teachers felt there was not enough time permitted to comprehensively cover all required curricular areas before administering the standardised assessments. In this regard, one participant noted that the school year is thus shortened to May (R25). Therefore, the pressure experienced by teachers is intensified when trying to cover a curriculum “characterised by its breadth and balance” (Government of Ireland, 1999, p.10) in the time approaching standardised testing administration, a conclusion further highlighted by the INTO (2010).

When analysing data relating to ‘Educational Management’ the responses given on the questionnaire explicitly implicated both Principals and Inspectors in this regard. Conclusions in this area advocate that owing to a school’s ‘good reputation’ for having high standardised test scores, pressure was indirectly experienced owing to the Principal’s high expectations in this area. Also, teachers felt that Inspectors can look at the standardised testing scores when assessing the effectiveness of a school and formulate judgements of a teacher based on these results. Research undertaken in this area would certainly support these conclusions (Gipps, 1988; Turner and Clift, 1988; NCCA, 2005; Mac Ruairc, 2009; INTO, 2010), adding that “teachers pressured to ensure their students pass the criterion-referenced test spend huge amounts of time drilling minimal competency skills” (Turner and Clift, 1988, p.314) to the detriment of other more meaningful forms of learning.

A surprising conclusion based on this theme was the lack of awareness by the Inspector that teachers experience pressure when undertaking standardised testing in their classrooms. When asked about this issue, the Inspector felt that no pressure was experienced by teachers in this regard (Interview, lines 89-90). Here, the Inspector's conflicting perspective may point to the fact that teachers may simply not have voiced their opinions on this matter, and so those from the Inspectorate remain unaware of this burden of pressure upon teachers. It may also indicate that teachers are not provided with a suitable forum to air such concerns, and so issues, such as those pertaining to accountability, are permitted to go unaddressed. However, it must be noted that some attempts have been made in this regard. Here, NCCA employee Ann Looney vehemently states "We have regularly made a strong plea for in-service" (INTO, 2010, p.110). However, as of yet, this aim has not been realised.

Interestingly, data analysis on this theme indicated that the negative attitudes teachers have constructed towards accountability, as explored above, are outweighed by an informed appreciation (amongst 92% of those surveyed) of the numerous benefits associated with this form of testing, resulting in a prevailingly positive attitude currently existing towards standardised testing in the primary school context.

5.1.4 Theme 3: Preparation

Firstly, upon initial analysis of the data on this theme, it became evident that the majority of teachers (64%) indicated that they did not spend any time preparing for the tests. Furthermore, some of those surveyed explicitly clarified that they did not engage in 'teaching to the test', a feature which points to the conclusion that teachers are aware of both the presence of this negative by-product of standardised testing, as well as the unconstructive manner in which it can affect the test's results. Therefore, the aforementioned teachers, dutifully trying to circumvent the effects of 'teaching to the test' in their site of practice, are striving to ensure that the standardised test instruments are being utilised in the manner intended by the test creators (Wall and Burke, 2004, 2007; ERC, 2006, 2007). Thus, conclusions derived from the data on this topic support the stance that the professionalism shown by such teachers in this regard is certainly commendable.

However, further conclusions based on the data analysed indicated that over a third of those surveyed do prepare their students for the tests. Results here indicated that the activities outlined by participants could be placed in one of three general categories; Procedural Instruction, Directed Revision and General Revision. Thus, one can conclude that there are a number of methods in existence by which pupils are prepared for standardised testing, each explored below.

Firstly, it was found that a significant number of those who indicated that preparatory activities were undertaken provided data which was categorised into the 'Procedural Instruction' sub-category. Here, conclusions highlight that forty per cent of the teachers involved in this study prepare their students for these formal assessments in an ethical manner, being careful not to 'teach to the test'. Procedural instruction prior to testing is further advocated by Haladyna, Bobbit-Nolen and Haas (1991), as the development of "testwiseness skills" (p.5) serves to abate a pupil's anxiety when faced with an unfamiliar format, without positively affecting the score they achieve.

In stark contrast to the ethical manner in which preparation was undertaken above, conclusions based on the data provided by participants confirmed that nearly a quarter (24%) of those surveyed in this study engage in 'directed revision' prior to testing. When discussing this topic, the Inspector confirms the existence of 'teaching to the test' in the Irish context, but indicates that the extent of this problem has not been ascertained by the Inspectorate (Interview, lines 112-114). Here, the conclusive data collected in this study validates and confirms the Inspector's belief, and further indicates that some teachers are employing activities that constitute 'teaching to the test', but they do not view them in this regard. In such cases it is evident that the term 'teaching to the test' is not understood homogeneously and thus 'acceptable practice' in the area of preparation needs to be clarified to ensure that those utilising standardised testing in their classrooms are doing so in an appropriate manner.

In a similar vein to the sub-category outlined above, sixteen per cent of those surveyed indicated that they engage in 'General Revision' prior to testing. Here again, conclusions based on the data obtained imply that 'acceptable practice' is a subjective concept requiring clarification as 'General Revision', though admittedly not enhancing pupils' scores to the same extent as 'teaching to the test', still serves to

distort the results achieved as children will perform better on the test than they would have done unaided, albeit in a somewhat unintentional fashion. This conclusion is further supported by literature devised on this topic (Turner and Clift, 1988; Bond, 2004; Wall and Burke, 2004, 2007; ERC, 2006, 2007; INTO, 2010).

5.1.5 Theme 4: Impact on Classroom Practice

When analysing data relating to this theme it was found that half of those surveyed indicated that no impact upon classroom practice was felt, with a further half indicating that they did experience an impact. However, conclusions arising from the careful analysis of data obtained from the open-ended qualitative question that followed this closed-ended question found that this apparent balance of opinion in the data would be more accurately described as a forty-eight per cent (No Impact) to sixty per cent (Impact Experienced) breakdown in statistics. Conclusions based on this topic further advocate that where an influence was indicated, two distinct categories of impact were evident in the data; *Pre-testing* and *Post-testing* impacts, each explored below. Furthermore, these impacts encompass both ethical and unethical activities, a view also shared by Haladyna, Bobbit-Nolen and Haas (1991). Thus, it would seem that there are various methods, embodied in the data collected, by which classroom practice can be influenced as a direct result of undertaking standardised assessment.

Most notably, nearly half (48%) of those surveyed indicated that no impact upon their general instruction was experienced. Interestingly, the data collected would suggest that some of those who specified that there was no impact upon their instruction viewed standardised testing as a distinctly separate educational entity, one that is completely removed from their teaching and in no way should impact upon it. In such cases, one must conclude that these teachers are not fully allowing the numerous benefits associated with standardised testing to enhance their classroom practice. However, further conclusions on this topic also advocate that there are those in this category whose attitudes are influenced with ethical considerations in mind. Here, some of those who indicated that no impact on classroom practice was experienced implied that this was as a result of deliberate effort, as this form of assessment should not be viewed as any more important than other less prolific forms, such as teacher observation and teacher-designed tasks and tests. However,

here again, one must question if perhaps the tests' potential role in such instances is being constrained somewhat as the results should be used more to inform teaching within the classroom context, a view supported by O'Leary (2006). This conclusion was also noted by the Inspector when commenting "I think it should inform the teaching more than it has" (Interview, line 81).

Just under a third of those surveyed (32%) indicated that they felt standardised testing impacted upon their pre-testing instruction. Here, we can see direct links with the methods of preparation sub-themes, explored above, emerging in the data. Arising from close analysis of the data relating to 'pre-testing' impacts, it was evident that both ethical and unethical practices were adopted. Thus, conclusions in this regard highlight the reality that there are varying extents on the ethical continuum to which teachers alter their classroom practice in the time leading up to the commencement of standardised testing. Thus, one can conclude that pre-testing impacts are experienced by a significant number of those utilising this form of assessment, the moral degree of which depends solely on the teacher.

Conclusions derived from the data analysed relating to 'Post-Testing' impacts, as experienced by twenty-eight per cent of those surveyed, indicate that, generally, the impacts on classroom practice in this regard, as outlined in subsection 4.7.3, serve to enhance teaching and learning. Furthermore, one can clearly conclude that post-testing impacts are not as likely to result in the adoption of unethical practices, as can be the case with impacts in the 'Pre-Testing' context, as employing result-boosting exercises becomes redundant once the test is completed. One can also conclude that the positive impacts in this area, such as identifying areas of weakness and strength, are in line with those envisaged by the Inspector, and are much more favourable than some of the more questionable practices adopted in the 'Pre-Testing' context. Thus, overall conclusions formed on the topic of post-testing impacts indicate that those who experience such influences are more closely aligning their professional practice with that advocated in literature in this area, whereby;

Teachers may look at a set of class results to see whether any significant patterns or features are apparent. These patterns can provide the teacher with information to adapt his/her teaching methods, differentiation strategies, content of the learning experiences, and so on to meet the child's learning needs more effectively.

The positive impacts experienced in this regard also further support the previous conclusion that the assessments are not being used to their best effect by those who stated that ‘No Impact’ upon their instruction was felt, as they are restricting the potential benefits that standardised testing can bestow upon their teaching in the post-testing context.

5.2 Recommendations

Drawing from the main findings of this study, as well as those documented in related literature, a number of recommendations are being proposed in the area of standardised testing aiming to enhance current practice and increase the valuable contribution that this form of assessment can make in the primary school classroom.

5.2.1 Standardised Testing In Ireland

- As both the Sigma-T series (Wall and Burke, 2007) and The Drumcondra Test series (ERC, 2006, 2007) have not been standardised since 2006, and the Micra-T series (Wall and Burke, 2004) dating back as far as 2003 in this regard, it is therefore recommended that a revision of the norms against which children are being currently compared is dutifully undertaken. This activity is seen as “timely” (INTO, 2010) as bearing in mind that in the four years since the most recent test instruments were standardised, many innovative methodologies and technologies have revolutionised the classroom environment, such as the prevalent use of interactive whiteboards, as well as other pedagogical advancements. Thus, when one casts their mind back to the primary school-going population of 2006, and compares the knowledge and experiences of these children with those of today, one cannot deny that we are not currently comparing like with like. Thus an updating of the norms upon which the ultimate success of standardised testing depends is undoubtedly required.
- Throughout this study, the limited amount of research carried out on standardised assessment in the Irish context was noted. Though the NCCA

(2005, 2007), Mac Ruairc (2009) and, to a lesser extent, Kelly's dissertation research (2007) have certainly made some progress in this regard, more directed research needs to be undertaken in this area, a requirement that this dissertation has aimed to highlight and address. When one considers that standardised assessment is now compulsory, it is vital that large-scale research be undertaken in order to ensure that this form of assessment is being utilised correctly. This recommendation is further reiterated by the INTO, who note that "it is timely, therefore, to reconsider assessment policy and practice in primary schools in Ireland" (2010, p.19). Having learned from the pitfalls of other countries, such as those outlined in the United States and Great Britain, it is vital that our formal testing practices be closely monitored so that we do not enter into a 'high- stakes' testing situation, as explored in chapter two, in which the current pressures experienced by teachers would only be further exacerbated.

5.2.2 Theme 1: Benefits Associated with Standardised Testing

- The benefits associated with this form of assessment, as previously outlined, are undeniable. Yet, one must ensure that they are interpreting such benefits in context. Here, the measure of pupil achievement obtained from standardised testing is envisaged as a supplementary source of evidence which a teacher can use when considering a pupil's academic performance. Over-reliance on such results should be dutifully avoided as these "test results may be adversely affected by temporary states such as fatigue, anxiety, demoralisation, inattentiveness or boredom" (Wall and Burke, 2004, p.20). Also, when discussing the topic of learning support, research advocates that a single standardised test result should never be the sole determinant of whether a child will be considered for learning support or not (Wall and Burke, 2004, 2007; NCCA, 2005, 2007; ERC, 2006, 2007). In this regard, the NCCA justly state "It is important to use a range of assessment information when making decisions" (2007, p.61). Thus, in order to enjoy the benefits afforded by this form of assessment, it is recommended that teachers must also recognise the limitations of the results obtained from standardised tests, placing them within an appropriate context in their classroom.

5.2.3 Theme 2: Accountability

- As the majority (68%) of teachers in this study indicated that they felt pressure when utilising standardised assessment this is certainly an area that should be further addressed. Training should be provided for teachers by those charged with regulating assessment procedures in the Irish context, namely the NCCA. Initially when the NCCA released ‘Assessment in the Primary School’ in 2007, a number of once-off seminars on assessment were offered to schools, yet not all schools attended such events and those who have qualified since have received no such training, a neglectful situation that certainly requires immediate attention. This requirement is further highlighted by Kelly, whereby ninety-six per cent of those surveyed in her study felt that “in-service is urgently required” (Kelly, 2007 as cited in INTO, 2010, p.100). This in-service training should directly address the topic of accountability, providing teachers with the knowledge and skills to minimise pressure felt from this aspect of formal testing. The NCCA have certainly made recent attempts to provide such a service, as noted by Looney (INTO, 2010), but to date no professional development in this regard has been provided due to lack of funding, a situation clearly requiring immediate action.
- It is recommended that pre-service training be provided to all student teachers in the Colleges of Education regarding the correct usage of standardised assessment so that those who will be utilising this form of assessment will be aware of the negative by-products, such as pressurised accountability, that can become evident when standardised assessment is misused. In this manner an informed approach towards formal testing would be undertaken and sustained into the future.
- In the interest of reducing the amount of pressure experienced by teachers in relation to the Inspector’s usage of the standardised testing scores, it is recommended that transparency surrounding this process be established. This would serve to clarify the ambiguity currently surrounding this process, as noted by Mac Ruairc (2009), and thus the pressure that arises from teachers’ suspicions, quite possibly unfounded, on this topic would be abated.

- As previously highlighted in this study, the Inspector interviewed did not seem to be aware of many of the pressures experienced by teachers when undertaking standardised testing in their site of practice. This indicates that teachers are not voicing their concerns to the Inspectorate personnel in this regard, or perhaps they are not being afforded ample opportunities to do so. Here, it is recommended that a suitable forum should be established in which teachers can air such concerns, and also share their experiences in relation to standardised testing. This may merely comprise of a simple 'Discussion Forum' being set up on the Department of Education's website, www.education.ie, that could be monitored by the Inspectorate, thus helping to inform and enhance their own professional practice, and indeed, that of the primary school teaching community nationwide.
- In order to effectively reduce the pressure emanating from parents, it is recommended that whenever the scores are sent home to parents an accompanying leaflet should be provided explaining the process of standardised testing and how to place their child's score in the appropriate context. This leaflet could be easily devised by the NCCA, taking the current information available on their website for parents and merely summarising it to an appropriate length for inclusion in a clear, concise leaflet. Currently this information, though useful, is not readily accessible to parents who may not be aware of the NCCA, their website, or their work in this regard. Such a leaflet could thus address issues of misplaced accountability, as well as highlighting that the score derived from standardised testing is not a comprehensive indication of their child's annual academic performance. In this manner, parental misconceptions that can result in the formation of pressure for teachers would be effectively addressed and thus perhaps this source of pressure could be greatly reduced.
- Curriculum overload was cited as a significant source of pressure in this study. Here, it is recommended that in schools where teachers experience pressure in this regard, the time of year the children sit the tests could be changed to the Autumn option, as opposed to the more widely-adopted Spring/ Summer alternative. As the children would be sitting the tests at the

beginning of the school year, this would eliminate any pressure to have the entire curriculum covered before the end of the academic year. Also, as the teacher would only have received the class that September, both the teacher and the parents would be far less likely to view the resulting test scores as a reflection of the teacher's ability. Wilson further supports this recommendation, adding that "despite substantial progress in test design, standard administrations of norm- or criterion- referenced tests given at the end of a year or unit provide little guidance for classroom instruction" (2009, p.68). Here, Wilson further outlines that administering the assessments at the beginning of the year would more positively inform instruction than waiting until the end of the year, when perhaps the time for intervention may have passed. This is certainly an issue that should be discussed at individual school level, as "the optimal time for administering the test depends largely on the purpose for which it is being used" (Wall and Burke, 2007, p.2). Therefore, it is recommended that each school must decide if it is the assessment *of* learning or assessment *for* learning (NCCA, 2007) they wish to prioritise to meet their school's particular needs.

5.2.4 Theme 3: Preparation

- What was perhaps most evident from the analysis of data relating to this theme was the varied approaches undertaken by teachers when preparing their pupils for the tests. In light of this finding, it is recommended that a uniform approach in this regard be devised by the NCCA and subsequently communicated to teachers either via a Department of Education circular or in in-service training seminars. Of the upmost importance here is the need to establish which practices are ethical and which are not, as there are a range of activities currently being undertaken, such as general revision and 'teaching to the test' (directed revision), that serve to enhance the scores obtained in the tests, therefore ultimately rendering the results invalid, as previously explored. Thus, guidelines stipulating ethically sound methods of preparation need to be clarified and disseminated to teachers so the value of standardised assessment in our primary school classrooms is not being undermined.

- In keeping with the theme of preparation, it is recommended that for teachers preparing to undertake standardised testing in their classrooms in the coming year(s), a summer course (online or otherwise) should be devised which would aim to educate teachers about the practices and pitfalls associated with this form of assessment. Perhaps if this course provided an extra EPV¹ day to those who have completed it this would increase the uptake, thus ensuring that a greater number of teachers are being educated on the correct usage of this form of assessment than would otherwise be the case. This recommendation would also ideally constitute an opportunity for the NCCA, who recognise the need for training in this area, to contribute towards this situation in a productive and concrete manner.

5.2.5 Theme 4: Impact on Classroom Practice

- In keeping with advice offered by the Inspector, the key recommendation of this study in this area is that the results of standardised testing should be used more than is currently the case, to impact positively upon teaching and learning in the classroom. Many of those surveyed indicated that no impact on their instruction/ classroom practice was experienced, and in such cases one must determine that the true function of standardised testing is not being realised, as a key aim of the tests is to “enhance teacher professionalism by enabling teachers to make more informed decisions regarding how best to instruct pupils” (Wall and Burke, 2007, p.1). Here, the test creators, in agreement with the recommendations of this study, promote the generation of post-testing impacts upon instruction. Here, for example, teachers should use the standardised test results to identify areas of the curriculum where children are succeeding and/or having difficulty and put programmes in place to support their learning in this area. Such post-testing impacts clearly serve to enhance the quality of education and should be more widely experienced by those using this form of assessment.

¹ EPV days is an acronym for ‘Extra Personal Vacation’ days. Upon completion of a Department of Education approved summer course, a teacher receives 3 EPV days that they may take at their discretion throughout the school year in compensation for the time they have spent developing their professional skills in their chosen summer course.

- Also, arising from the careful analysis of the data obtained in this study, it is recommended that pre-testing impacts, such as those previously explored in this chapter, should be limited to procedural activities, as other harmful practices, such as ‘teaching to the test’, can pervade classroom practice in this respect and serve to undermine the quality of education afforded to the children in such instances.

5.3 Further Research

While this study does provide a much needed insight into the impacts of standardised assessment in Irish primary school classrooms, it also provides a base upon which further research could be conducted. Hence, based on the findings and conclusions of this study, areas that would benefit from further attention include:

- Conducting a more expansive research project in relation to teachers’ attitudes towards standardised testing. Though out of the scope of this study, it would prove beneficial to enlarge the sample population to encompass a larger variety of geographical areas and school sizes, thus providing a more extensive data base from which valid generalisations could be more appropriately formulated.
- Undertaking a more in-depth study relating to the role of parents in the standardised testing process. Such a project could explore parents’ beliefs surrounding this form of assessment, identify misconceptions that can result in the formulation of accountability and aim to synthesise appropriate methods by which parents could be educated about the features and practices associated with this form of assessment.
- Exploring initiatives that could be established to support teachers when using standardised testing, given the current lack of funding available. Such a study could examine the feasibility of setting up an appropriate internet discussion forum, whereby teachers can discuss the concerns and difficulties they experience when using standardised testing and get advice from those more experienced in this area.

- Conducting a study on the attitudes of newly-qualified teachers towards standardised testing, as they have been given no formal training in utilising this form of assessment. Such a study could examine how these teachers cope with initial test administration, correction and interpretation, as well as exploring how they develop their skills in this area. An exploration of the activities these teachers deem appropriate when preparing their students for the tests could also be undertaken, the findings of which could be easily compared with those outlined in this study.
- Devising a suitable programme that could be adopted to educate teachers about the standardised testing process. An aspect of this study may entail exploring if perhaps such a programme could be delivered effectively in a summer course (either online or otherwise), as previously discussed, or if such content could be summarised and presented in a leaflet/ circular that could subsequently be distributed to all teachers.
- Building upon the findings of Mac Ruairc's study (2009), future research could be conducted to further probe the bias present in the current tests' content. As standardised tests are the creation of those deriving predominantly from a middle-class background, it stands to reason that the situations and experiences included in the tests are based on the middle-class culture. In such instances Neill and Medina state "the purported objectivity of tests is often no more than the standardization of bias" (1989, p.692). Thus, research in this area could aim to identify cases of test item bias and provide a framework for updating the test instruments so that they are more accessible to those from other socio-economic groups.
- As foreign national/ newcomer pupils are often deemed unsuitable to undertake standardised tests, as advised in standardised testing manuals (Wall and Burke, 2004, 2007; ERC, 2006, 2007), research could be carried out to ascertain the most appropriate manner in which the achievement levels of such pupils could be measured. A key objective of this research may be to explore the feasibility of formulating a suitable standardised test instrument that would effectively achieve this aim.

- Updating and clarifying current Irish assessment policies in light of recent international trends, as advocated by the INTO (2010). When considering the work of Hall and Kavanagh (2002), the NCCA note “Assessment policy and practice here are characterised by conceptual uncertainty, by vagueness about goals and purposes and by lack of clarity about the place of assessment information in classrooms” (2005, p.13). Updating policies in this regard would certainly contribute towards ensuring that assessment procedures in Irish primary schools, in particular those relating to standardised testing, are regulated and monitored to ensure that a homogenous, informed approach is adopted by all schools in this regard, therefore ensuring that such practices serve to enhance the quality of education provided to the children within our education system.

References

- Abrams, L.M., Pedulla, J.J. and Madaus, G.F. (2003) 'Views from the classroom: teacher's opinions of statewide testing programs', *Theory into Practice*, 42 (1), pp. 18-29. JSTOR [Online]. Available at: <http://uk.jstor.org/> (Accessed: 10 October 2009).
- Anagnostopoulos, D. (2005) 'Testing, tests and classroom texts', *Journal of Curriculum Studies*, 37 (1), pp. 35-63.
- Babbie, E. (2010) *The practice of social research*. 12th edn. United States: Adsworth Cengage Learning.
- Bartlett, S. (2000) 'The development of teacher appraisal: a recent history', *British Journal of Educational Studies*, 48 (1), pp. 24-37.
- Bassey, M. (1995) *Creating education through research*. Newark: BERA.
- Bell, J. (2005) *Doing your own research project: a guide for first-time researchers in education and the social sciences*. Buckingham: Open University Press.
- BERA. (2003) *Good practice in educational research writing*. Southwell, Notts: BERA.
- Blaxter, L., Hughes, C. and Tight, M. (1996) *How to research*. Buckingham: Open University Press.
- Boland, J. (2008) *Curriculum and assessment: course handbook*. NUI Galway: Galway.
- Bond, L. (2004) *Teaching to the test: carnegie perspectives*. Unpublished Paper. The Carnegie Foundation for the Advancement of Teaching, California.
- Broekhoff, M. (1978) 'Criterion-referenced evaluation as an alternative', *The English Journal*, 67 (1), pp.32-37.
- Brown, D.F. (1992) 'Altering curricula through state-mandated testing: perceptions of teachers and principals', *American Educational Research Association Conference*. San Francisco April. San Francisco, CA: American Educational Research Association.
- Bryman, A. (2001) *Social research methods*. New York: Oxford University Press.
- Bryman, A. (2008) *Social research methods*. 3rd edn. New York: Oxford University Press.
- Cannel, C.F. and Kahn, R.L. (1968) 'Interviewing', in Lindzey, G. and Aronson, E. (eds.) *The handbook of social psychology Vol. 2: research methods*. New York: Addison-Wesley.

Charmaz, K. (1983) 'The grounded theory method: An explication and interpretation', in Emerson, R.M. (ed.) *Contemporary field research: a collection of readings*. Boston: Little Brown.

Cimbricz, S. (2002) 'State-mandated testing and teachers' beliefs and practices', *Education Policy Analysis Archives*, 10 (2), pp. 1-21.

Cizek, G.J. (2001) 'Cheating to the test', *Education Matters*, 1 (1), pp.1-8.

C. J. Fallon Publishers (2010) Micra-T. Available at:
http://www.cjfallon.ie/product_list.aspx?mode=1&id=63&l0=63 (Accessed: 10 June 2010).

Cohen, L. and Manion, L. (1994) *Research in education*. 3rd edn. London: Routledge.

Cohen, L., Manion, L. and Morrison K. (2007) *Research methods in education*. 6th edn. London: Routledge Falmer.

Corbett, H.D. and Wilson, B.L. (1991) 'Two state minimum competency testing programs and their effects on curriculum and instruction', in Stake, R.E. (ed.) *Advances in program evaluation: vol 1 effects of mandated assessment on teaching*. Greenwich, CT: JAI Press Ltd.

Creswell, J.W. (1998) *Qualitative learning and research design: choosing among five traditions*. London: Sage.

Creswell, J.W. (2003) *Research design: qualitative, quantitative and mixed-methods approaches*. 2nd edn. Thousand Oaks, CA: Sage.

Denscombe, M. (1998) *The good research guide: for small-scale social research projects*. Buckingham: Open University Press.

Denzin, N.K. and Lincon, Y.S. (1994) *Handbook of qualitative research*. London: Sage.

Educational Research Centre (ERC). (1976) *Drumcondra English Test, Level III, Form A*. Dublin: Educational Research Centre, St Patrick's College, Dublin.

Educational Research Centre (ERC). (2006) *Drumcondra Primary Mathematics Test- Revised, Levels 3-6: Administration and Technical Manual*. Dublin: Educational Research Centre, St. Patrick's College, Dublin.

Educational Research Centre (ERC). (2007) *Drumcondra primary reading test-revised: level 3-6 administration and technical manual*. Dublin: Educational Research Centre, St. Patrick's College, Dublin.

Flick, U. (1998) *An introduction to qualitative research*. Thousand Oaks, CA: Sage.

Fontes, P.J., Kellaghan, T., Madaus, G.F. and Airasian, P.W. (1980) 'Opinions of the Irish public on examinations', *The Irish Journal of Education*, 14 (2), pp. 53-74.

Gall, M.D., Borg, W.R. and Gall J.P. (1996) *Educational research: an introduction*. 6th edn. New York: Longman.

Gipps, C. (1988) 'The debate over standards and the uses of testing', *British Journal of Education Studies*, 36 (1), pp. 21-36.

Gipps, C.V. (1994) *Beyond testing: towards a theory of educational assessment*. London: Routledge Falmer.

Glaser, B.G. and Strauss, A.L. (1967) *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine.

Goulding, C. (2002) *Grounded theory: a practical guide for management, business and market researchers*. Thousand Oaks, CA: Sage.

Government of Ireland (1998) *Education act*. Dublin: The Stationery Office.

Government of Ireland. (1999) *The primary school curriculum: introduction*. Dublin: The Stationery Office.

Great Britain. Department of Education and Science (1983) *Teaching quality*. London: Her Majesty's Stationery Office.

Great Britain. Department of Education and Science (1985) *Better schools*. London: Her Majesty's Stationery Office.

Great Britain. Department for Education and Science (1991) *School teacher appraisal act*. London: Department for Education and Science.

Great Britain. Department for Education and Employment (DfEE) (1997) *Excellence in schools*. London: Department for Education and Employment.

Great Britain. Department for Education and Employment (DfEE) (1998) *Teachers meeting the challenge of change*. London: Department for Education and Employment.

Green, K.E. (1992) 'Differing opinions on testing between pre-service and in-service teachers', *The Journal of Educational Research*, 86 (1), pp. 37-42.

Haladyna, T.M., Bobbit-Nolen, S. and Haas, N.S. (1991) 'Raising standardised achievement test scores and the origins of test score pollution', *American Educational Research Association*, 20 (5), pp. 2-7. JSTOR [Online]. Available at: <http://uk.jstor.org/> (Accessed: 12 October 2009).

Hall, K. and Kavanagh, V. (2002) 'Primary assessment in the republic of Ireland: Conflict and consensus', *Educational Management and Administration*, 30 (3), pp.216-274.

Hursh, D. (2005) 'The growth of high-stakes testing in the USA: accountability, markets and the decline in educational quality', *British Educational Research Journal*, 31 (5), pp. 605-622.

- Hutton, P.F. (1990) *Survey research for managers: how to use surveys in management decision making*. 2nd edn. Basingstroke: Macmillan.
- INTO (1997) *Teaching and assessment: issues in assessment*. Dublin: Irish National Teachers Organisation.
- INTO (2010) *Assessment in the primary school: discussion document proceedings of the consultative conference on Education 2008*. Dublin: Irish National Teachers Organisation.
- Ireland. Department of Education and Science (1992) *Education for a changing world: green paper on education*. Dublin: The Stationery Office.
- Ireland. Department of Education and Science (1995) *Charting our educational future: white paper on education*. Dublin: The Stationery Office.
- Ireland. Department of Education and Science (2006) *Supporting assessment in primary schools*. [Online] Available at: <http://www.into.ie/roi/Informationforteachers/descirculars/descirculars2006.html>. (Accessed: 20 October 2009).
- Kellaghan, T. and Fontes, P. (1989) 'Relative bias in teacher judgements and standardized tests in the identification of literacy problems', *The Irish Journal of Education*, 23 (2), pp. 40-50.
- Kelly, I. (2007) *The assessment process- current practice and perspectives in Irish primary schools: improving teacher motivation to assess*. Unpublished dissertation. Mary Immaculate College, Limerick.
- Kiddler, L.H. and Judd, C.M. (1986) *Research methods in social relations*. London: CBS Publishing.
- Kidston, J.S. (1985) 'The validity of questionnaire responses', *The Library Quarterly*, 55 (2), pp.133-150.
- Knight, P.T. (2002) *Small-scale research*. London: Sage.
- Kohn, A. (2000) *The case against standardised testing: raising the scores, ruining the schools*. Portsmouth: Heinemann.
- Kraemer, D.J. (2005) 'Fighting forward: why studying standardized tests with our students is important', *The English Journal*, 94 (4), pp. 88-92.
- Lam, T.C.M. and Bordignon, C. (2001) 'An examination of English teachers opinions to the Ontario grade 9 reading and writing test', *Interchange*, 32 (2), pp. 131-145.
- Mac Kenzie, N. and Knipe, S. (2006) 'Research dilemmas: paradigms, methods and methodology', *Issues in Educational Research*, 16, no page numbers.

- Mac Ruairc, G. (2009) 'Dip, dip, sky blue, who's it? NOT YOU': children's experiences of standardised testing: a socio-cultural analysis', *Irish Educational Studies*, 28 (1), pp. 47-66.
- Madaus, M., Airasian, P. and Kellaghan, T. (1971) 'The effects of standardised testing', *Irish Journal of Education*, 5 (2), pp. 70-85. JSTOR [Online]. Available at: <http://uk.jstor.org/> (Accessed: 12 October 2009).
- Madaus, G.F. and Kellaghan, T. (1992) 'Curriculum evaluation and assessment', in Jackson, P.W. (ed.) *Handbook of research on curriculum*. New York: Macmillan.
- Mason, J. (2002) *Qualitative research*. London: Sage.
- Mc Neil, L. (2000) *Contradictions of reform*. New York: Routledge.
- Mertens, D.M. (2005) *Research methods in education and psychology: integrating diversity with quantitative and qualitative approaches*. 2nd edn. Thousand Oaks, CA: Sage.
- Meuser, M. and Nagle, U. (1991) 'Expert interviews: often, little tests mindfully: a contribution to qualitative methodology' in Garz, D. and Kraimer, K. (eds.) *Qualitative empirical social research*, pp. 441-468. Opladen: West German publishing house. (Translation)
- Miell, D., Phoenix, A. and Thomas, K. (eds.) (2002) *Mapping psychology 1*. Milton Keynes: The Open University.
- Miles, M. and Huberman, A. (1994) *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Minister for Education and Science. (2004) 'Minister for Education and Science, Noel Dempsey T.D. to make standardised testing of pupils a requirement' [Press Release]. 16 July. Available at: <http://www.education.ie/robots/view.jsp> (Accessed: 30/ 01/ 2010).
- National Council for Curriculum and Assessment (NCCA). (2005) *Supporting assessment in schools: standardised testing in compulsory schooling*. Dublin: NCCA.
- National Council for Curriculum and Assessment (NCCA). (2006) *Your child and standardised testing*. Dublin: NCCA.
- National Council for Curriculum and Assessment (NCCA). (2007) *Assessment in the primary school curriculum: guidelines for schools*. Dublin: NCCA.
- Neill, D.M. and Medina, N.J. (1989) 'Standardised testing: harmful to educational health', *Phi Delta Kappan*, 70 (9), pp. 688-697. JSTOR [Online]. Available at: <http://uk.jstor.org/> (Accessed: 11 October 2009).
- O'Leary, Z. (2004) *The essential guide to doing research*. London: Sage.

- O'Leary, M. (2006) 'Towards a balanced assessment system for Irish primary and secondary schools', *Oideas*, 52, pp. 7-25.
- Oppenheim, A.N. (1992) *Questionnaire design, interviewing and attitude measurement*. 2nd edn. London: Printer Publishers Ltd.
- Posner, D. (2004) 'What's wrong with teaching to the test?', *Phi Delta Kappan*, 85 (10), pp.749-751.
- Punch, K.F. (2003) *Survey research: the basics*. London: Sage.
- Robson, C. (1993) *Real world research*. Oxford: Blackwell Publishers Ltd.
- Salmon-Cox, L. (1981) 'Teachers and standardised achievement tests: what's really happening?', *The Phi Delta Kappa International*, 62 (9), pp. 631-634. JSTOR [Online]. Available at: <http://uk.jstor.org/> (Accessed: 11 October 2009).
- Santrock, J.W. (2001a) *Child development*. London: McGraw-Hill.
- Santrock, J.W. (2001b) *Educational psychology*. London: McGraw-Hill.
- Sherman, R. and Webb, R. (eds.) (1988) 'Qualitative research in education: forms and methods', in Blaxter, L., Hughes, C. and Tight, M. (1996) *How to research*. 3rd edn. Buckingham: Open University Press.
- Spiggle, S. (1994) 'Analysis and interpretation of qualitative data in consumer research', *Journal of Consumer Research*, 23 (3), pp. 491-503.
- Spolsky, B. (1981) 'Some ethical questions about language and testing' in Klein-Braley, C. and Stevenson, D.K. (eds.) *Practice and problems in language and testing*. Frankfurt: Peter Lang.
- Strauss, A. and Corbin, J. (1998) *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Turner, G. and Clift, P. (1988) *Studies in teacher appraisal*. East Sussex: Falmer Press.
- Tymms, P. (2004) 'Are standards rising in English primary schools?', *British Educational Research Journal*, 30 (4), pp. 477-494.
- USA. Department of Education (2002). *No child left behind Act (NCLB)*. Washington, DC: US Department of Education.
- USA. Department of Education (2002) *No child left behind: parents guide*. Washington: Office of Public Affairs.
- USA. National Commission on Excellence in Education (1983) *A nation at risk: a report to the nation and to the secretary on Education*. Washington, DC: US Department of Education.

Vyrostek, S. (2009) 'Accountability the individual way', *Educational Horizons*, 87 (2), pp. 128-34

Wall, E. and Burke, K. (1987) *Micra-T- test manual*. Dublin: CJ Fallon

Wall, E. and Burke, K. (1991) *Sigma T- test manual*. Dublin: CJ Fallon

Wall, E. and Burke, K. (2004) *Micra-T level 3: test manual*. Dublin: CJ Fallon

Wall, E. and Burke, K. (2007) *Sigma-T level 4: test manual*. Dublin: CJ Fallon.

Wilson, M. (2009) 'Assessment from the ground up', *Phi Delta Kappan*, 91 (1), pp.68-71.

Zitlow, C. (2001) 'The case against standardised testing: raising the scores, ruining the schools by Alfie Kohn: a review', *The English Journal*, 91 (1), pp. 112-114.