

MA MODULE: INVESTIGATING RESEARCH

ASSIGNMENT: DISSERTATION RESEARCH PROPOSAL

SUBMITTED BY: GILLIAN COULTON

(EDUCATIONAL ASSESSMENT MA)

DATE OF SUBMISSION: 3rd SEPTEMBER 2010

WORD COUNT: 3766

Defining the research issue

England's statutory national assessment system, steadfast for many years in its reliance on standardised tests, is currently undergoing something of a sea change. In 2011/12 key stage 2 (KS2), the only key stage still retaining externally marked tests, switched to teacher assessment as the statutory outcome for writing. And this came just two years after the discontinuation of the KS2 science test. Suddenly then, primary teachers are faced with different demands and apparently greater freedom following the relinquishing, at least in part, of an arguably inflexible externally imposed assessment system that recognises nothing but performance on the day. This begs the question, what happens to teachers' practice when high-stakes tests are discontinued?

Since the inception of the national tests, they have sparked debate, dissent, inquiry and considerable teaching to the test (House of Commons Committee, 2008). They have justifiably grown into the label of 'high-stakes' in the sense that, for schools and teachers, the stakes are high given that test outcomes and the mass of data spawned in their wake constitute the overriding accountability indicator used to judge both school and, in many cases, teacher effectiveness. From both the professional and academic communities, there have been claims of restricted teacher autonomy and reduced opportunity for educative activity not solely purposive in raising test scores. Yet, since the partial dismantling of the national test system from 2008 onwards, there has been scant attention of how teachers in situ have responded to test cessation. Ofsted's appraisal suggests that resulting action is not uniform: in respect of science it reports positively on teachers' increased freedom resulting in beneficial changes to curriculum delivery but finds little impact, or certainly none characterised in the affirmative, in English (Ofsted, 2010, 2012).

Against this backdrop, the interest here is in exploring the effect, or absence of tangible effect, of test abolition on teachers' pedagogy and practice, which at this stage is defined as a component

or amalgam of teaching style (instructional or student centred, for example) and curriculum delivery at classroom level (content, activity type, emphasis, time allocation and so on). The study would seek to increase understanding of how the teachers most directly affected by the termination of the tests respond to this occurrence; it would search for underlying explanations through identification and exploration of factors influential in determining these responses. To this end, the focus will be exclusively on Y6 teachers as a clearly definable and identifiable group of professionals situated in the front line both during and following the period of KS2 national testing in writing and science; professionals who theoretically may be experiencing greater freedom to shape the direction and creativity of their teaching, perhaps for the first time.

Key research questions

- How do Y6 teachers view the discontinuation of the science and writing national tests at KS2, particularly in terms of the impact on their own situation?
- In what ways have Y6 teachers responded or intend to respond to the discontinuation of the science and writing national tests in terms of their practice?
- What factors have influenced teachers' decisions or capacity to either change aspects of their practice or maintain the status quo following the abolition of the tests?

Literature review

The proposed research rests on the premise that there exists a relationship between government mandated high-stakes testing and teachers' practice at classroom level. If evidence suggests that the former exerts a change in the latter then it would seem reasonable to investigate the response at the chalk-face when the precipitator of this identified change is removed. Theoretical and empirical work abounds on the effects of such testing in schools, both in the UK and beyond.

Based on empirical study in the United States, Au, (2007) argues that high-stakes tests impact on teachers' practice in that they leverage "content, formal and pedagogic" control (p263) over the curriculum by policy makers. In his qualitative metasynthesis, Au interrogates the effects of high-

stakes testing on curriculum (which he defines as incorporating pedagogy) through analysis of 49 qualitative studies all conducted in the United States, mostly between 2001 and 2006. The prolific nature of such inquiry is perhaps explained by the 2001 No Child Left Behind Act, which federally legislated for compulsory school based performance standards testing. The studies under scrutiny found contradictory evidence regarding the influence exerted by these tests. By far the most reported effects were a narrowing of content, a change in pedagogical emphasis to be more didactic and an increased tendency for knowledge “fragmentation manifested in the teaching of content in small, individuated and isolated test-sized pieces” (p262). In a small minority of cases, however, studies reported the opposite effects. These contrary findings though, are largely interpreted by Au as anomalies borne out of particular contextual factors including the atypical features of specific tests that accordingly required atypical responses in terms of preparatory teaching. Interestingly, social studies is identified as a possible ‘special case’. Although Au concedes that the slightly contradictory findings “complicate the understanding of the relationship between high-stakes testing and classroom practice” (p258), the fact that pedagogy, in whatever direction, was so widely affected is the crucial point from the perspective of the research issue under discussion. Also of note, is the implication that the extent and nature of test impact varies, at least somewhat, across contexts, including between particular subject disciplines.

In the UK, the national tests initially instigated for seven, eleven and fourteen year olds, have similarly attracted a great deal of study and comment. Between 1993 and 1995, during the period when these assessments gained momentum moving from voluntary pilot to statutory status, Brown et al (1996) report widespread, rapid changes at classroom level. Using quantitative and qualitative data from the National Assessment in Primary Schools project (NAPS) they found school level pedagogic changes in the form of more subject based and whole class teaching as well as increased use of ability setting. In the qualitative case studies, all the individual Y6 teachers described engaging in test preparation activity even for the pilot study and in subsequent years made curriculum adaptations in maths, English and science by, for example, amending schemes of work and incorporating test aligned activity formats. Whilst this research is limited to a

fact finding mission rather than generating theory and argument, the authors do note the potential for a narrowing of the curriculum in undesirable ways, given the amount of test preparation activities they discovered even at this early stage of policy implementation.

Teacher views of the tests are of course also highly relevant here given the interest in exploring how teachers respond to their discontinuation. According to White et al (2002), in a study examining perceptions of the influence of high-stakes tests on their practice, beginning teachers “tended to perceive that the test program affected their instructional beliefs less than it affected their instruction: that their beliefs were more resilient” (p58). Comments from the participants also showed that some teachers held a determination to adapt their instruction only as required in order, presumably, to attend to the demands of tests that were not wholly compatible with their pedagogic beliefs. The authors, however, drawing upon the work of Smith 1991, caution that regardless of this intention, “teachers may lose the capacity to use particular methods if they are not permitted to do so,” (p58). Similarly, although contextualising the proposed research interest in the broader frame of policy impact on teachers’ lives, McNess et al (2010), concluded that there was a “disjunction between policy and preferred practice” (p255) and that the predominant performance culture in primary schools left teachers with reduced autonomy and choice.

In light of what the empirical evidence suggests to be the extensive nature of test preparation activity and its associated constraints on pedagogic autonomy and practice, it makes sense to further advance understanding of this dynamic by examining what happens in reverse. Situated more closely in the area of interest of the proposed research, is a study by Collins et al (2010) that explores teachers’ opinions and experiences of science teaching at KS2 in relation to the influence of national tests; the research employs telephone questionnaires and focus groups and adopts a comparative approach drawing on Y6 teachers (in addition to heads and science co-ordinators) in Wales following test abolition, and their counterparts in England where, at the time, statutory science testing was still in force. Whilst the reported evidence for England adds weight to existing empirical work, the findings relating to Wales are more intriguing. Participants greeted test

abolition favourably, perceiving direct pedagogical benefits through new flexibility to determine lesson content. Teachers, from both nations, articulated the view that KS2 science national test scores failed to provide accurate pupil attainment information. Even so, the authors still conclude that test abolition “has not immediately led to radical changes in teachers’ practice,” (p284).

Although a greater emphasis on practical work was a perceived consequential adjustment in teaching, participants also pointed to the on-going use of non-statutory test materials continuing to form an integral part of teaching and learning. There was even an apparent need for comprehensive revision of the KS2 science curriculum in Y6; in fact, participants offered the opinion that summative teacher judgements required “a revision schedule that closely resembled that reported by Y6 teachers in England” (p282).

So, the theoretical and empirical work evidences a strong relationship between high-stakes testing and consequential change in teachers’ pedagogy and practice. However, although the question of what happens when this testing stops has been afforded less attention, there is evidence to suggest that the effect on practice may not be as predictable, even when teachers’ articulate positive attitudes to test termination. Of course, the processes are very different: introducing a high-stakes test may create an imperative to act but it should not be assumed that removing the test creates an equally compelling imperative to counteract, particularly when the tests and associated practice have become such a longstanding, embedded feature of schooling. The importance of the proposed research, therefore, comes in exploring the relatively unexplored space of test cessation, complementing the larger scale research by Collins et al (2010) in examining perspectives from teachers in England now that test abolition has spread across the Welsh border, and in contributing theory that may be a useful dialogic tool in relevant settings.

Methodological approach

In determining a methodological approach, a clear sense of the intended purpose of this inquiry is vital. The aim of the proposed research is to enhance understanding of teacher response to a specific occurrence. Given its small scale nature, the inquiry will be far more limited in scope than

the reviewed empirical work, for example by involving fewer participants, but it will seek to go beyond description of the occurrence under scrutiny into the realms of generating new knowledge relevant to the setting. As there will be no attempt to generalise the findings to a broader population, and having adopted the position that depth rather than breadth is crucial in interrogating the data, a purely qualitative approach will be employed. By formulating theory, the intention is to provide further insight into Y6 teachers' post-test experiences, in particular, by helping to make sense of their subsequent pedagogic decisions and actions in relation to test discontinuation.

It is intended, therefore, to adopt an approach based on elements of grounded theory, with its emphasis on understanding human action and illuminating process centred around a topic, (Cresswell 2007). In addition to its suitability for the topic, grounded theory also offers guidance on the research process in terms of clearly delineated procedures for data analysis and theory development, (Lal et al, 2012). At this stage there is also intent to draw on practices found in narrative inquiry, though clearly as the research unfolds, the scope and ideas formulated in planning may be refined in the light of emerging issues further down the line. Still, the thinking is that incorporating an element of narrative may be a useful device for capturing what is a temporally sequenced process as teachers move from teaching within a test orientated regime, through abolition, and then beyond, even including the possibility of planned future responses given the relatively little time that has elapsed since these assessment changes. Although the primary intention will still be to analyse and develop theory, eliciting and presenting some of the data in a narrative form may help to illuminate aspects of this theory, as well as, perhaps, specific contextual factors or participant responses of particular interest. Bailey and Jackson (2003), have articulated their search for complementary methods to a grounded theory approach after discovering that within the bounds of their research, the necessary fragmentation of data resulted in the unique perspectives of participants becoming somewhat lost. By considering a combination of these methods then, the hope is that the study will result in theory development that helps to explain teacher response, illustrated more vividly by supporting narrative episodes that tell key

stories across the changes in statutory assessment in the words of the participants. Both approaches are entirely compatible with the proposed data collection method and, most importantly, with the necessity of ensuring that theory is fully grounded in the data.

Research design

Sampling

Though clearly there will be no attempt to draw on a representative sample given the qualitative nature of this study, there is a need to consider sampling strategies in selecting individuals “because they can purposefully inform an understanding of the research problem and central phenomenon,” (Cresswell 2007 p125). In the first instance then, selection would depend on the fulfilment of stipulations integral to the study, namely experience of teaching in Y6 both during and after the period that testing in science and writing was statutory. In effect, that will necessitate participants having been in situ for four years or longer in order for them to have sufficient first-hand knowledge of practice relevant to both tests. The sample will also be constrained by issues of access; the method of data collection will necessitate geographical proximity between the researcher and participants. Ethical considerations are another crucial factor to be attended to from the outset, although as there is no intention to include observation of the teachers in their place of work, approval and permissions are likely to be limited to the individual participants themselves, viewed for the purposes of this study as autonomous practitioners.

Consensus regarding appropriate sample size in a grounded theory study seems to be within the range of 10 to 60, (Cresswell 2007, Lal et al 2012) There are, however, certainly examples of smaller samples within the methodological literature; for example, Kunkwenzu and Reddy (2002), who recruited six subjects for their study of beginning teacher socialisation. Grounded theory frequently draws on theoretical sampling which allows for adjustment and further selection of the sample as categories emerge in data analysis, enabling the researcher to pursue lines of inquiry coming to the fore or to test emerging theory, (Cresswell, 2007). Propitious though this may be in

ensuring saturation of categories, it clearly makes prediction of necessary sample size difficult, presenting practical difficulties and perhaps insurmountable challenges for small scale and first time research. Therefore, the initial intention is to include six teachers in the sample, proceeding not simply on pragmatic grounds but also with the possibly naïvely optimistic stance that theoretical saturation “depends in part on the quality of data obtained as opposed to the quantity of individuals recruited,” (Lal et al, 2012 p10). Within these limits then, the aim will also be to seek maximum variation, loosely interpreted to take account of one or two potentially influential factors discernible before commencing data collection with an individual interviewee; for instance, some variation in relevant subject specialisms and length of Y6 teaching experience might well be pertinent. Use of a snowball method may be of assistance in increasing the manageability of recruiting potential participants both in terms of identification and encouragement to commit (Dowling and Brown, 2010). A fusion of recommendation generated from participants themselves coinciding with interests coming from the data would, of course, form the ideal sampling route.

Data Collection

Semi structured, face –to-face interviews are the planned method of data collection, a mode frequently utilised in both grounded theory and narrative inquiry. Digital recording will facilitate a fully transcribed record. Questions, which will largely be formulated to be open in nature, may start with some preliminary fact finding requests framed deliberately more tightly to function as ice breakers and provide some valuable upfront contextual information. For example, asking participants who, in say a parallel two form entry school, are quite likely to plan collaboratively with a colleague(s), to describe relevant team structures within their school setting that have a direct influence on what happens in their classroom. The intention is to pre-plan sufficient structure to be sure that key topics are covered whilst maintaining the necessary flexibility to “explore the world from the perspective of the interviewee and to construct an understanding of how the interviewee makes sense of their experiences,” (Dowling and Brown p78). In keeping with grounded theory, as the interview stage progresses, emerging themes in the data may also have a bearing on the

interview direction determining, for example, when to use probes to clarify or pursue potentially illuminating aspects of data in more depth. The use of narrative methods may also be a useful device in giving fluency to the interview. Although gathering teachers' perspectives will, no doubt, be a lot 'messier' than establishing a linear chronology of practice, encouraging participants to recount their teaching 'stories' in relation to the existence and termination of the tests may provide a familiar and non-threatening structure through which to access the data.

Further consideration of both the relation between interviewer and interviewee and what Dowling and Brown (2010) term "the interactional context" (p79), raises a number of points. Although coming from different work places means formal hierarchical relationships will be avoided, the actors here will have a shared experience of teaching in Y6 throughout the period of interest, a dynamic which could lead to the interview functioning almost like a focus group with the risk of the interviewee construing a fellow practitioner as professionally intimidating, thereby inhibiting full and frank responses, or conversely, as an audience to 'impress' with responses that may not be based on reality. With this in mind, deliberate efforts will be made to establish a relationship that will, as far as possible, provide reassurance and engage participants. For example, assurances will be made about anonymity (name changing) and the entirely non-judgemental nature of the enquiry. In the interests of putting interviewees at ease, the choice of location will ultimately be theirs, though if they select their own school, there will be an ethical need to seek agreement from the headteacher. Finally, in terms of perspectives, although the interviewer will have views of the issues in question formulated from personal experience, the requirement here will be to use this knowledge productively to help make sense of the interviewees perspective and not to impose or assume any personal interpretations or opinions.

Data analysis

Within this proposed inquiry it is perhaps data analysis that will draw most heavily on the distinguishing methods that mark out grounded theory. Though arguably a daunting process to execute for the fledgling researcher, the method does offer a tangible scaffold upon which to frame a course of action. Mindful though of this complexity and the time commitments involved in Strauss and Corbin's (cited in Cresswell 2007) rigorous formulation of procedures, a somewhat condensed, less ambitious adaptation will be pursued in this project. Beginning with open coding, interview transcripts will be analysed to separate out key components (words, phrases, statements) in the data, using tags or labels drawn largely from the utterances of the participants during interview. In effect this first analytical layer will ensure that the eventually emerging theory is grounded on a foundation entirely constructed from the participant's responses; the empirical setting will 'be heard', (Dowling and Brown 2010). Overlapping the open coding process, inductive analysis will be taking place of the fragmented, but now indexed, information. This process will effectively entail reassembling the data, moving towards identified conceptual categories that gradually become more inclusive and defined as commonalities and relationships between them are discovered. In terms of the narrative element to the research design, participant's accounts may be storied or re-storied around the chronology of test abolition. Where using episodes of narrative will help to elaborate on categories or aspects of the final theory, they will be integrated into the report. Without, of course, introducing any pre-existing hypothesis, it is always possible that there will be temporal elements integrated as forming part of the theory.

Where compromises in methods may well figure is in the amount of what Cresswell (2007) depicts as 'zig-zagging' between data analysis and data collection, creating a cycle of one feeding into the other and providing constant momentum in the analysis process. As mentioned, although theory based sampling decisions about who to speak to next would be the ideal, they cannot be guaranteed in the confines of this study. Nonetheless, the two processes will run concurrently and interrogation of data will help to inform subsequent interviews. Another aspect to navigate is the process of constant comparison integral to analysis which entails an intense back and forth scrutiny of data codes, categories and emerging theory, continually cross checking them against

each other until saturation occurs and theory is coherent. Again, manageability will almost certainly limit this process and the move from open codes to categories may occur more speedily than prescribed in the Strauss and Corbin perspective delineated in Cresswell (2007). It will be important though to try and avoid potentially consequent difficulties with the formation of theory, for example, by settling for a 'thin' theory that may be underdeveloped or incapable of explaining data that may not fit. Also challenging, might be the task of keeping theory development firmly in the foreground as opposed to lapsing into rich description, which is probably easily done in first time research.

Contribution to professional practice

As a longstanding Y6 teacher responsible for assessment and curriculum development, further insight into how teachers in different settings have adapted their practice in light of the changes to the statutory testing system would be of the utmost relevance. The proposed study provides the opportunity to discover different approaches to curriculum delivery in a shifting assessment system. Of greater importance though, is its potential to enhance understanding of factors that can either facilitate or prevent this from happening, thereby serving as a platform from which to develop effective means of supporting individual teachers in overcoming barriers prohibiting change.

References

Au W, 2007 High-stakes Testing and Curricular Control: A Qualitative Metasynthesis, *Educational Researcher*, 36:5, 258-267

Bailey D.M and Jackson JM (2003), 'Qualitative Data Analysis: Challenges and Dilemmas Related to Theory and Method', *American Journal of Occupational Therapy*. 57, 57-85

Brown M, Taggart B, McCallum B and Gipps C(1996):'The impact of key stage 2 tests, Education 3-13' International Journal of Primary Elementary and Early Years Education, 24:3, 3-7

Collins S, Reiss M and Stobart G, (2010) 'What happens when high-stakes testing stops? Teachers' perceptions of the impact of compulsory national testing in science of 11-year olds in England and its abolition in Wales, Assessment in Education: Principles, Policy and Practice',17:3, 273-286

Cresswell J (2007), 'Qualitative Inquiry & Research Design: Choosing Among Five Approaches', London Sage Publications London

Dowling P and Brown A, (2010) 'Doing Research/Reading Research. Re-interrogating Education' London, Routledge

House of commons Children, Schools and Families committee, Testing and Assessment Third report of Session 2007-08 Volume 1

Kunkwenzu E and Reddy C (2008) 'Using grounded theory to understand teacher socialisation: A research experience' Education as Change, 12:1,133-149

Lal S and Suto M, (2012) 'Examining the Potential of Combining the Methods of Grounded Theory and Narrative Inquiry: A Comparative Analysis' The Qualitative Report Vol 17, Article 41, 1-22

McNess E, Broadfoot P and Osborn M (2003): 'Is the Effective Compromising the Affective?', British Educational Research Journal

The Office for Standards in Education, Children's Services and Skills (Ofsted) 2012 Moving English Forward. www.ofsted.gov.uk/resources/110118 Crown Copyright

White S, Sturtevant E and Dunlap K, (2002): Preservice and beginning teachers' perceptions of the influence of high-stakes tests on their literacy-related instructional beliefs and decisions.' Reading Research and instruction, 42:2, 39-62

