Working with the IMPaCT Taxonomy:

Encouraging Deep and Varied Questioning in the Mathematics Classroom

by

Jo Denton

Student Number: 1255558

Thesis presented as a partial requirement for the award of the Doctor of Education (EdD) of the University of Warwick

Affiliation: Centre for Education Studies

April 2017

Contents

List of Figures

List of Tables

List of Abbreviations

IMPaCT	Intended Mathematical Processes and Cognitive Thought
AfL	Assessment for Learning
SOLO	Structure of Observed Learning Outcomes
MATH	Mathematical Assessment Task Hierarchy
SPSS	Statistical Package for the Social Sciences
IRIS	A 360° remotely operated camera system
CPD	Continuing Professional Development

Acknowledgements

Foremost I would like to express my sincere gratitude to my supervisor Sue Johnston-Wilder for the guidance and feedback through the learning process of this thesis. Your support has been invaluable and despite the distance, I have learnt so much from you. Furthermore I would like to thank Michael Hammond for everything I learnt in my Advanced Research Methods (ARM) taught modules to get me where I am today. When I met you Michael my plan was simply to complete my Masters, however you, and my dear friend Emily from the ARM course, gave me the confidence in myself to embark on this doctorate.

I would like to thank the participants in my action research, particularly the four teachers who have willingly shared their precious time during the processes of interviewing, training and being filmed and observed. I would also like to thank my colleague and friend Zoe for taking her time to proofread the final draft of this thesis. Furthermore I would like to thank the technical support department at my school for supporting the participant teachers to record their lessons, particularly when our subscription to IRIS expired! I could not wish for more supportive and professional colleagues.

Finally I would like to thank my husband, Chris. You have supported me throughout this entire process, from convincing me to continue my studies following my success in my MSc, to transcribing interviews and testing the reliability of the IMPaCT Taxonomy. Thank you also for finding so many ways to entertain our young children Emma and Samuel to allow me the space and time to research and write.

Declaration

The work in this thesis was developed and conducted by the author between June 2013 and September 2016. I declare that, apart from work whose authors are explicitly acknowledged, this thesis and the materials contained in this thesis represent original work undertaken solely by the author. I confirm that this thesis has not been submitted for a degree at another university.

The following essays by the author form the basis for the pilot study in this thesis:

Denton, J. (2013a) Foundation Research Methods (FRM)

Part 1 - *Formative Questioning in Mathematics: An Open or Closed Case Study?*

Part 2 - *Formative Questioning in Mathematics: Probing Deeper than the Surface*

http://www2.warwick.ac.uk/fac/soc/ces/postgrads/pgr/eportfolios/edpmnu/assignments/frm/

The following essays are also referred to in this thesis:

Denton, J. (2013b) Researching Mathematical Learning (RML)

Part 1 - *Process-Object theories: A Mathematician's struggle with the ambiguity of language*

Part 2 - *Process-Object Theory: A Chicken and Egg Scenario?*

http://www2.warwick.ac.uk/fac/soc/ces/postgrads/pgr/eportfolios/edpmnu/assignments/rml/

Denton, J. (2015) Advanced Research Methods 3 (ARM3)

http://www2.warwick.ac.uk/fac/soc/ces/postgrads/pgr/eportfolios/edpmnu/assignments/arm/

Abstract

Despite a wealth of research into improving questioning in mathematics, recent research has identified the need for more effective questioning strategies which are accessible to mathematics teachers. This thesis investigates the types of questions which encourage mathematical thinking and participation, with the aim of deepening and varying mathematical thinking for learners through working with my own tool to develop questioning in mathematics, the *Intended Mathematical Processes and Cognitive Thought (IMPaCT) Taxonomy.*

Following a literature review of existing taxonomies which can be used for classifying questions, and framing the research around the establishment of sociomathematical norms in the classroom, this thesis develops a new taxonomy and presents the findings from the pilot study for this research followed by the empirical research to explore the effectiveness of the IMPaCT Taxonomy. Over the course of an academic year, 28 mathematics lessons, from four participant teachers and five classes of 14-16 year olds were observed, and the questioning in each lesson was analysed using the IMPaCT Taxonomy. After the first 15 of these observations, intervention was given to the participant teachers on using the IMPaCT Taxonomy before the remaining lesson observations were conducted. Proportions of question type and depth were analysed according to several factors: the participant teacher; the level of attainment of the class; the stage of the lesson; the Assessment for Learning techniques used; and the mathematical topic.

The results show that although all these factors affect the questioning, the attainment of the class and the individual teacher have the biggest impact on the questioning employed. The empirical analysis shows that both the teachers' understanding of how their questioning impacts on learners and the teachers' variety and depth of questioning can be increased through working with the IMPaCT Taxonomy, although there are variations between teachers to the extent of the impact of the IMPaCT Taxonomy.

Introduction

Background Context and Research Aims

Questioning is at the heart of what teachers do in the classroom (Hollingsworth, 1982), and strategies for supporting questioning have been developed in educational research for decades (see, for example, Bloom et al., 1956; Anderson & Sosniak, 1994). There exists a wealth of training and support for teachers on questioning (see DfES, 2004a; 2004b; Black et al., 2003). Yet according to Ofsted (2012), in the mathematics classroom questioning is not being used effectively to check and develop learners' understanding.

This thesis aims to investigate both the *question types* and the *questioning techniques* which encourage mathematical thinking and participation, with the aim of deepening and varying mathematical thinking for learners through working with a newly developed *Intended Mathematical Processes and Cognitive Thought (IMPaCT) Taxonomy*; in doing so it builds on research started in a previous micro-research study (Denton, 2013a) for the award of Master of Science.

Definition of Terms

Through both my Masters and Doctoral research, I have found that the terms methodology, methods and research design are often used interchangeably in educational research (Mackenzie & Knipe, 2006). Each of these terms are often interpreted slightly, if not drastically, different, with methodology being particularly difficult to define. Through my own research on methodology, I have concluded that methodology has two principal definitions: the theory linking epistemology to methods in terms of deductive and inductive reasoning (Creswell, 2003; Opie, 2004; Robson, 2002); and the overall approach to a study, for example case study, action research or ethnography

(Sikes in Opie, 2004). To avoid ambiguity in this thesis, *methodology* refers to the theory of generating and verifying knowledge (Punch, 2009) through inductive and deductive reasoning. The term *research strategies* refers to the overall approach to the research (Denzin & Lincoln, 2005), in this case, action research. *Methods* refers to the data collection techniques used in the research, for example interviews or questionnaires. *Research design* refers to the chapter in this thesis dedicated to describing the methods, research strategies and methodology employed in this research (Denton, 2015).

With respect to questioning, *question type* refers to the mathematical thinking intended for learners as a result of the teacher posing a question. *Questioning technique* refers to the strategies that teachers put in place for learners to think about and respond to questions. The *Intended Mathematical Processes and Cognitive Thought (IMPaCT) Taxonomy* attempts to classify questioning in terms of the type of mathematical thinking required by the learner; the word *intended* here is important, as we cannot guarantee that a learner will think through a problem in the way the teacher planned. All these terms are explored in detail later in the thesis.

Research Questions

Pilot Research Hypotheses

The previous research (Denton, 2013a) forms the pilot study for this action research where the following hypotheses were investigated:

1. A larger proportion of questions requiring a 'surface approach' are used in mathematics lessons than those requiring deeper thinking.
2. Using formative questioning techniques supports a wider range of intended mathematical thinking, compared to questions posed without these techniques.

Refined Research Questions for Thesis:

In the main body of this research, I investigate these initial research questions in greater depth and develop an action research study to investigate whether working with the IMPaCT Taxonomy affects the range of questioning types used by teachers and whether it supports teachers to understand the effect of their questioning on learners' mathematical thinking. The aim of this research therefore is to investigate the following questions:

1. **What factors affect the type and depth of questioning used by mathematics teachers?**
2. **Does working with the IMPaCT Taxonomy affect the type and depth of questioning used in mathematics lessons?**
3. **Does the IMPaCT Taxonomy affect mathematics teachers' understanding of how their questioning impacts on their learners' mathematical thinking?**

Overview of Study

Chapter 1 presents and discusses a review of both the relevant literature of conceptual learning theories to frame this research as well as literature on questioning and formative assessment, particularly that literature relating directly to mathematics. The literature review aims to present and analyse the research that exists to date. The conceptual element of the review will look at how social norms and sociomathematical norms describe the construction of knowledge through questioning; the empirical element of the review will examine prevailing research in the field of questioning and formative assessment to identify what is already known in this area of research and to also identify elements of the field which require further empirical investigation. This chapter reviews existing learning objective taxonomies and their limitations, and builds on the literature reviewed to develop an early version of the IMPaCT Taxonomy (see Appendix 1). In particular, Chapter 1 discusses taxonomies which are considered suitable for identifying question types and levels of complexity in terms of probing mathematical understanding. Formative questioning is discussed with a review of Black's and Wiliam's (1998) theoretical framework on formative assessment in order to analyse techniques which could support the asking of a broader range of question types in mathematics.

Chapter 2 describes in brief the pilot study carried out for this research as part of the Foundation Research Methods module for a Master of Science degree at the University of Warwick (Denton, 2013a). This chapter outlines the research methods employed and the subsequent findings from the pilot study. It should be noted that the pilot study and main study were conducted at different schools, so the outcomes of the pilot study could be used to refine the research questions, but did not establish a starting point for the main study as the contexts for the two schools were very different.

The refined IMPaCT Taxonomy is introduced detail in Chapter 3, and this chapter describes the changes made to the original version of the taxonomy based on the findings from the literature review and pilot study. Chapter 3 shows how a combination of the most relevant taxonomies was used to classify questions for the IMPaCT Taxonomy and how the visual representation of a Venn diagram (see Figure 3.1) was developed for teachers to work with the IMPaCT Taxonomy.

Chapter 4 describes the action research strategy to investigate the research questions for the main study and considers the methodology and data collection methods employed in the research design. A combined methods approach of both qualitative and quantitative methods is adopted, as recommended by many researchers, such as Hartas (2010) and Robson (2002). These methods include semi-structured teacher interviews and classroom observations. This mixed methods approach was selected to address the diverse nature of the data and to attempt to triangulate any findings. This chapter reflects on the reliability and validity of the chosen methods, in order to consider the limitations of the research, and outlines strategies to increase the reliability and validity of the study. Sampling and ethical implications are also addressed in this chapter.

The empirical research for this dissertation was conducted by means of action research at my current school, a mixed gender aged 13-18 upper school on the south coast of England, working with teachers to develop their questioning in the mathematics classroom. The school had recently become an elective academy. The school is non-selective and the proportion of learners who are from minority ethnic backgrounds or speak English as an additional language at the school is below average. This study does not make any attempt to generalise beyond this school however; the context of the school simply provides the reader with the potential to compare the findings and conclusions from this research with similar schools.

The findings from the initial observations in the first action research cycle, including the lesson observations and teacher interviews, are presented in Chapter 5. These are compared to the earlier findings from the pilot study, before presenting, in Chapter 6, the intervention element of the action research for working with the participant teachers on the IMPaCT Taxonomy. Chapter 6 describes the training that took place for the teachers on how their questioning impacts learners' understanding of mathematics and how establishing certain social and sociomathematical norms in the classroom can support the process.

The findings from the second cycle of the action research are presented in Chapter 7. Firstly, the findings from the interim monitoring observations, which took place in February 2016, are presented and the implications for further training are described. Secondly, the combined post-intervention results are presented. These results are compared to the earlier findings and are discussed in depth in Chapter 8. In this chapter, the findings are also compared and contrasted with the previous research in this field, making explicit links back to the literature review.

Finally, in Chapter 9, conclusions are drawn from the empirical findings for working with the IMPaCT Taxonomy and the limitations of this research is considered.

1. Literature Review

Teachers' questioning is "not always well judged or productive for learning" (DfES, 2004a, p.4). Furthermore, research highlights the need to use "open, higher-level questions to develop pupils' higher-order thinking skills" (ibid, p.18). But what constitutes higher-level questions and higher-order thinking and how can these be established in the mathematics classroom?

This chapter outlines the theories of social norms and sociomathematical norms, particularly in relation to establishing teachers' use of higher-order questioning in the mathematics classroom, and discusses how Yackel's and Cobb's (1996) *emergent perspective*, which combines social interactionist theory and psychological constructivist theory (ibid), will be used to frame the research. This chapter also outlines the research to date in developing the classification of thinking skills in the classroom, through an analysis of existing taxonomies available to support teachers' questioning, with particular consideration of how these taxonomies can be applied to developing learners' conceptual mathematical thinking, and looks at previous research which has analysed the proportions of question types in the mathematics classroom. In addition, the chapter reviews literature on formative assessment techniques which optimise learners' mathematical thinking when used in conjunction with teachers' questioning (Black & Wiliam, 1998; Black et al., 2006; Hodgen & Wiliam, 2006).

The Emergent Perspective

Cobb and Yackel (1996) consider learners' mathematical activity to be "social through and through in that it develops as they participate in classroom mathematical practices" (p.180) and set about to "coordinate analyses of classroom processes that are conducted in psychological and in social terms" (ibid). The *emergent perspective* is a theoretical framework which combines both social and psychological perspectives in the classroom (Cobb & Yackel, 1996). By social perspective, Cobb and Yackel (1996) are referring to Bauersfeld et al.'s (1988) interactionist perspective and how the classroom functions collaboratively in the development of social norms, which could allow learners to explain and justify their thinking. The psychological perspective refers to the constructivist theory about how learners contribute to the development of collective processes through their individual actions (Cobb & Yackel, 1996).

From an emergent perspective, "mathematical learning is both a process of active construction [...] and a process of acculturation into the mathematical practices of wider society" (Yackel & Cobb, 1996, p.460). Table 1.1 clarifies this interpretive framework as played out in the classroom context.

Social Perspective	Psychological Perspective
Classroom social norms	Beliefs about own role, others' roles, and the general nature of mathematical activity in school
Sociomathematical norms	Mathematical beliefs and values
Classroom mathematical practices	Mathematical conceptions

Table 1.1. An Interpretive Framework for Analysing Individual and Collective Activity at the Classroom Level (Cobb & Yackel, 1995, p.177).

The emergent perspective considers positive social norms to be established by the teacher in the classroom which are "characterised by explanation, justification, and argumentation" (Yackel & Cobb, 1996, p.460). These characteristics are not specific to mathematics lessons, as learners should be expected to justify their own thinking and challenge the thinking of others across the curriculum, not just in mathematics (ibid). Yackel and Cobb (1996) believe that to develop learners' mathematical thinking, norms which are unique to the learning of mathematics need to be established, which they refer to as *sociomathematical* norms. Sociomathematical norms are established as a result of classroom discussion between the teacher and learners (Hershkowitz & Schwarz, 1999). These sociomathematical norms include developing a learner's understanding of what constitutes an *acceptable mathematical explanation and justification*, as well as developing an understanding of *mathematical difference*, *mathematical sophistication*, *mathematical efficiency* and *mathematical elegance* (Yackel & Cobb, 1996). A positive sociomathematical norm is intended to "set an expectation in the classroom that encourages strong mathematical activity in the form of justification" (Gerson & Bateman, 2011, p. 115).

According to Hershkowitz & Schwarz (1999), sociomathematical norms are "social constructs specific to mathematics that individuals negotiate in discussions to develop their personal understandings" (p.150) and can often be viewed as observable classroom actions (Levinson et al., 2009). However Levinson et al. (2009) debate whether sociomathematical norms are in fact more complex to interpret than this, as a classroom community consisting of a wide range of individuals is a "complex environment" (ibid, p.171).

Sociomathematical Norms and Mathematical Autonomy

> Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. (Article 26 of the Universal Declaration of Human Rights, cited in Piaget. 1973, p.41)

Piaget (1973) interprets this statement to mean that the main goal of education is learners achieving intellectual and moral autonomy. Cobb and Yackel (1996) define intellectual autonomy as an "awareness of and willingness to draw on their own intellectual capabilities when making mathematical decisions and judgements" (p.9), however Holster (2006) argues that learners' intellectual autonomy is often taken for granted in many classrooms. From their empirical research, Yackel and Cobb (1996) explain that for learners to establish mathematical autonomy, teachers have to ensure that the sociomathematical norm of acceptable explanations and justifications involves "described actions on mathematical objects rather than procedural instructions" (p.461). Therefore just explaining *what* they did was insufficient, of great importance was the *how* and, even more importantly, justifying *why*.

The teacher plays an important role in developing this autonomy (Holster, 2006) by providing opportunities for learners to explain and justify their ideas, which are key aspects in learners developing reasoning skills in mathematics (Whitenack & Yackel, 2002). However when some learners reason their own ideas, other learners in the classroom are also given the opportunity to reflect on those ideas and evaluate them (ibid). Whitenack and Yackel (2002) describe how all learners can benefit from classroom discourse involving explaining and justifying thinking. The learner who is explaining or justifying has the opportunity to revisit their mathematical thinking, giving the opportunity to build upon the argument, with the hope of developing a stronger understanding of the mathematics. Or perhaps in the act of explaining, the learner will find a new way of approaching the task, thereby developing new ideas or understanding. According to Whitenack and Yackel (2002), sharing mathematical thinking in this way can develop the mathematical thinking of all the participant learners. This relates well to Vygotsky's (1978) social constructivism and the Zone of Proximal Development (ZPD), where learning is constructed through social interaction. However Vygotsky states that this is maximised through "adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p.33) with the aim that the less knowledgeable partner, for a particular topic, makes progress in their understanding. However, as described above, there is scope for all parties to benefit from mathematical discourse.

Yackel and Cobb (1996) recognise there is a place for *explaining* in the classroom to clarify learners' thinking, especially if it might not be clear to other learners. Whitenack and Yackel (2002) state that explaining is in fact essential for learners to develop mathematical argument. However, in an inquiry-based classroom, that is, one which focuses on learning mathematics through problem-solving, learners should distinguish between explanations which simply describe the process or procedure undertaken as opposed to explanations which "describe actions on experientially real mathematical objects" (ibid, p.467). Yackel and Cobb (1996) found that the highest order explanations allowed for most reflection if given sufficient time.

The sociomathematical norm of *mathematical difference,* where learners explore the conceptual differences between approaches to a given mathematical problem, can support higher-order thinking (Yackel & Cobb, 1996). In establishing the sociomathematical norm of mathematical difference in Yackel's and Cobb's (1996) research, it was common for the participant teachers to ask if learners had solved problems in a different way. In these classrooms, what was considered as different was not measured against set criteria, rather what constituted mathematical difference came about as a result of the classroom discussions.

Yackel and Cobb (1996) note that "additional learning opportunities arise when children attempt to make sense of explanations given by others, to compare others' solutions to their own, and to make judgements about similarities and differences" (p.466). In addition, a greater focus on explaining other learners' answers as opposed to their own could reinforce the justifying *why* instead of the explaining *what,* since the learner cannot simply cite a procedure as it might not have been the procedure used by their peer (Martino & Maher, 1999). The classroom norms of allowing learners to make sense of other learners' explanations, being given opportunities to agree or disagree with other learners, and exploring alternative solutions to problems (Cobb & Yackel, 1996), hence need to be established for this justification to occur. However, Yackel and Cobb (1996) found that in order for teachers to maximise these learning opportunities, they needed to listen to the learners' explanations carefully themselves and to respond appropriately.

Although teachers often asked learners if they had a different solution, it was less common in Yackel's and Cobb's (1996) research in the US that questions posed by the teacher involved asking learners if they had a more efficient or a more sophisticated method, so different methods were not compared in terms of their effectiveness in a given situation. Teachers need to prompt learners to consider the advantages and disadvantages of different ways to solve a problem. Simply asking learners to give an alternative method does not constitute an understanding of mathematical difference without reference to the structure of the mathematics, however in comparing efficiency and sophistication, a deeper level of cognitive activity is required (ibid).

In framing their research on the emergent perspective, Kazemi and Stipek (2001) found the following sociomathematical norms encourage conceptual thinking:

> (a) an explanation consists of a mathematical argument, not simply a procedural description; (b) mathematical thinking involves understanding relations among multiple strategies; (c) errors provide opportunities to reconceptualise a problem, explore contradictions in solutions, and pursue alternative strategies; and (d) collaborative work involves individual accountability and reaching a consensus through mathematical argumentation (p.59).

The first sociomathematical norm parallels Yackel's and Cobb's (1996) stance that for explanations to elicit higher-order thinking, learners need to move away from describing procedures and progress to describing actions on mathematical objects. The second relates to Yackel's and Cobb's (1996) mathematical difference and, potentially, mathematical efficiency, sophistication and elegance depending how the teacher encourages learners to compare and contrast the multiple strategies. Kazemi's and Stipek's (2001) final two sociomathematical norms relate to teachers providing an environment where learners see the value of learning from mistakes and in ensuring all learners participate in mathematical thinking during small group work. These sociomathematical norms are in agreement with Hershkowitz and Schwarz (1999) who describe the norms that “meaningful activity is valued more than correct answers and that partners should reach consensus as they work on activities” (p.150). This notion is also supported by Whitenack and Yackel (2002) who found in their classroom-based research that mathematical activity is more effective when the emphasis is put on exploring mathematical ideas as opposed to just getting correct answers.

Through questioning, teachers can “stimulate students’ conceptual understanding of mathematics” (Kazemi & Stipek, 2001, p.60). However in a similar conclusion to Yackel and Cobb (1996), Kazemi and Stipek (2001) found that although teachers regularly ask questions requiring learners to describe their methods, being able to compare the mathematical concepts behind the various methods was more pedagogically demanding for the teachers. Kazemi and Stipek (2001) hence consider the former as a social norm and the latter as a sociomathematical norm. Kazemi and Stipek (2001) found that small differences in pedagogy can increase learners’ opportunities to think conceptually about their mathematics. Similarly, Kazemi and Stipek (2001) found that teachers found it easy to establish the social norm that making mistakes is acceptable in the course of learning, but establishing the sociomathematical norm of how those mistakes are used to engage learners with the mathematics was more problematic.

Whitenack and Yackel (2002) list questions that learners may start to ask themselves as they go about problem-solving in mathematics:

> Why might I use one approach over another? What information might I use to help me solve this problem? Can I solve the problem in more than one way? Are some approaches ‘easier’ or more efficient? (Whitenack & Yackel, 2002, p.526)

Whitenack and Yackel (2002) believe that allowing learners to reason independently will help them contribute to class discussions, however these questions need to be first posed by the teacher to model that reasoning, in order for learners to develop the required language to develop their reasoning with others, as using the correct language in reasoning can often be problematic for learners (Holster, 2006).

Yackel and Cobb (1996) found that sociomathematical norms are constrained by the classroom participants, including the teacher. If a teacher only asks questions which require lower-order thinking, then learners giving a superficial answer becomes a classroom norm. If, however, the teacher probes the learners' understanding further, then justification can be elicited and a deeper reasoning becomes the norm. It is therefore the teacher's responsibility to share with learners "what counts as an acceptable mathematical explanation and justification" (ibid, p.461) for it to ultimately become a sociomathematical norm

Classification of Questioning

Mason (2000) states three pedagogical purposes for classifying questions: focusing attention of learners, testing, that is, monitoring learners' understanding, and enquiry. Fraivillig et al. (1999) also propose a three tiered structure (see Figure 1.1) to examine the purpose behind the question: Eliciting Children's Solution Methods, Supporting Children's Conceptual Understanding, and Extending Children's Mathematical Thinking (ibid), which they call the ACT Framework (Advancing Children's Thinking in Mathematics). 'Eliciting Children's Solution Methods' concerns learners sharing their thinking in mathematical discourse and includes the teacher eliciting different ideas from the class. 'Supporting Children's Conceptual Understanding' is where the teacher encourages learners to make links with prior knowledge and to explain the ideas of others. 'Extending Children's Mathematical Thinking' is where mathematical reflection is encouraged and more sophisticated and efficient methods are considered, which parallels

Yackel's and Cobb's (1996) sociomathematical norms of mathematical sophistication and mathematical efficiency.

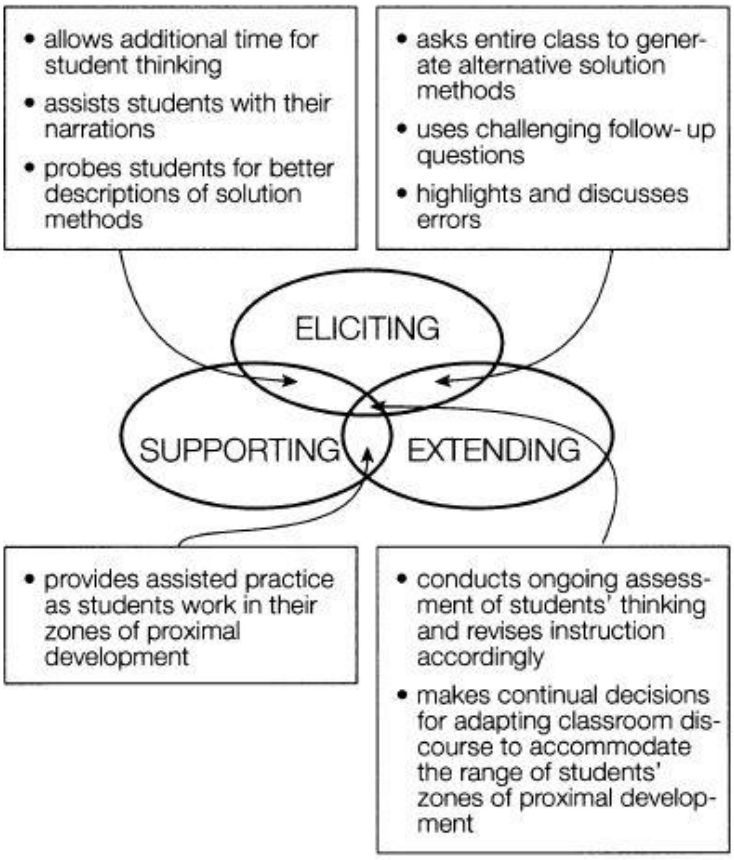


Figure 1.1. Advancing Children's Thinking (ACT) framework (Fraivillig et al., 1999, p. 154)

Morgan and Saxton (2006) also classify questioning in three ways: probing what the learner already knows; building a context for shared understanding between teacher and learners; and challenging learners to think both critically and creatively. This resonates with Mason (2000) and seems to support both the classifications in the ACT framework and Yackel's and Cobb's (1996) emergent perspective, however the classifications are very broad, particularly the second category which could contain a large array of types of mathematical understanding as well as a wide range of level of complexity in the questioning associated with the category.

Recent research into classroom discourse analysis in the mathematics classroom by Drageset (2014) reviewed a wide range of discourse analysis research from the ACT framework (Fraivillig et al., 1999) to Wood's (1998) funnelling and focusing. Funnelling is where teachers cognitively guide the learners to the solutions, whereas focusing places more emphasis on the learners thinking mathematically for themselves (ibid). Drageset (2014) describes how discourse analysis tools such as funnelling and focusing are used to look at teachers' practices too broadly to be helpful for teachers to improve their practice. Whereas the ACT framework, for example, looks at an element of teaching practice in greater detail which is "crucial for professional development as teachers have little use for general advice" (p. 289).

Distinguishing between open and closed questioning is a common way teachers are encouraged to classify their questioning (Hargreaves, 1984), however Watson (2003) criticises the simplicity of open and closed questioning in the teaching and learning of mathematics. Watson (2003) instead infers that opportunities to extend conceptual mathematical understanding are of greater importance when questioning learners in a mathematics classroom than simply considering whether a question is open or closed. Hargreaves (1984) attempts to overcome this issue with the notion of half-open questions, that is questions which require a yes or no response, but depending on how the question is posed will either encourage learners to elaborate on their answer or not.

Bloom's Taxonomy

Since the 1950s, many researchers have attempted to produce a hierarchy for the complexity of thinking skills (Gall, 1970), however it was the taxonomy of Bloom et al. (1956) (Figure 1.2) which experienced "phenomenal growth" (Bloom in Anderson & Sosniak, 1994, p.1), being used across a variety of countries and subject areas (Seddon, 1978) and as a result Bloom's Taxonomy became widely accepted as the optimal classification of questioning (Gall, 1970). Bloom's Taxonomy classifies learning objectives into a hierarchy of complexity in terms of learners' thinking and was applied widely in educational establishments to classify teachers' questioning (ibid). However Bloom's Taxonomy was not intended to be used as a "constructive way of planning and answering questions" (Morgan & Saxton, 2006, p.19). Instead it was intended simply to be a framework about knowledge which "helps us to see the kind of thinking we can set into action through questions" (ibid).

Figure 1.2. *Bloom's et al's 1956 Taxonomy (Image from GURO21, 2012)*

Despite Bloom being the most widely recognised taxonomy for thinking skills, there have been criticisms of its effectiveness to improve learning. One limitation of Bloom's Taxonomy cited by its critics, is that it classifies observable behaviours as opposed to describing how the learning is constructed from these classifications (Ormell, 1974), however Bloom et al. (1956) were aware of this limitation and in fact only intended the taxonomy to contribute to the development of a more complete theory of learning as opposed to providing the complete theory with the hierarchy itself. Yang (2006), however, contends that Bloom's Taxonomy is based on "unwarranted assumptions" (p.201) where "deficiency of explanatory power and inconsistencies in research results" (ibid) indicate that Bloom's Taxonomy is ineffective, as the level of a question cannot be judged in isolation of the participants and context of the lesson.

Despite Bloom's Taxonomy being internationally recognised, according to Anderson & Sosniak (1994), any educational changes as a result of the introduction of the taxonomy occurred mostly at policymaker level rather than having direct influence on the practice of classroom teachers. A further criticism of Bloom's Taxonomy was the omission of the term 'understanding' in the hierarchy (Furst, 1994). Furthermore it was felt that the taxonomy was presented as a hierarchy, thus implying that 'knowledge' automatically leads to intellectual abilities (Bereiter & Scardamalia, 1999). Bereiter and Scardamalia (1999) disagreed with this notion, claiming that traditional views of knowledge can be perceived on multiple levels of complexity and type, where not all forms of knowledge would necessarily lead to the construction of deeper learning and understanding.

The criticisms of the language used to describe each stage of Bloom's Taxonomy were addressed by Anderson et al. (2001) in a revised taxonomy (Figure 1.3).

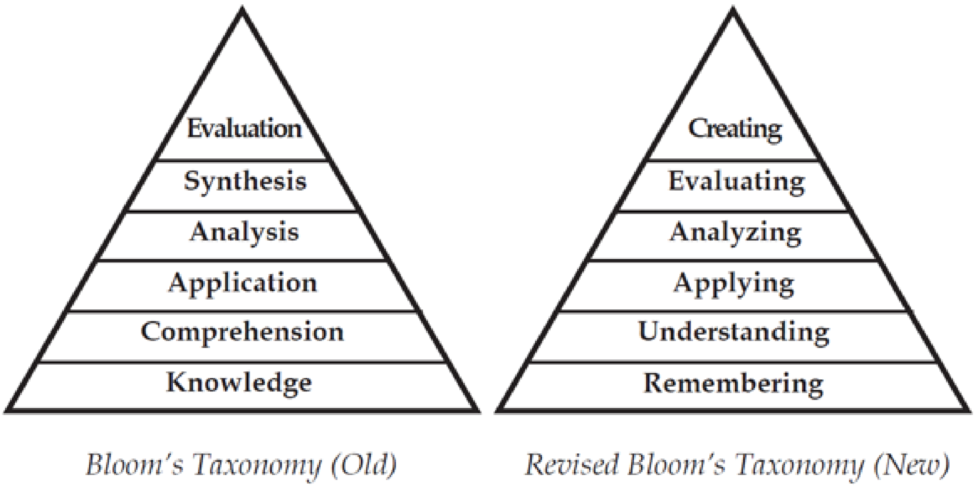


Figure 1.3. *Bloom's Original and Revised Taxonomies (Image from GURO21, 2012)*

Anderson et al. (2001) viewed understanding as a "widespread synonym for comprehending" (Kraftwohl, 2002, p.214) which was commonly used by teachers in their learning objectives (ibid) so the *comprehension* category was replaced with *understanding*. Besides the omission of the term knowledge and the addition of the term understanding there are other, more subtle differences between the two versions of Bloom's Taxonomy. Evaluating has been removed from the top tier, instead taking lower precedence in the hierarchy. This has allowed for *creating* to be deemed as the highest rank in thinking skills. The term *creating* replaced the term *synthesis* as it better reflected the thinking involved in this level of the hierarchy (Anderson et al., 2001). This could also be considered a reflection on the importance of learner-centred learning in the 21st Century classroom. Another, perhaps more subtle, difference is the replacement of the nouns in the original version with verbs in the revised version, perhaps highlighting the importance of active learner participation when teachers use the taxonomy in the classroom as opposed to simply the theory of what is expected to be happening.

Bloom's Taxonomy and Mathematical Understanding

As mentioned above, Bloom's Taxonomy presents a hierarchy of thinking skills, where remembering and understanding are considered to be lower-order thinking skills, while applying, analysing, evaluating, and creating are considered higher-order skills (Anderson et al., 2001). However is such a hierarchy necessarily applicable to the learning of mathematics? Watson (2007) claims that Bloom's Taxonomy "underplays knowledge and comprehension in mathematics" (p.114) as these can be interpreted at different levels of mathematical thought and states that it "does not provide for post-synthetic mathematical actions, such as abstraction and objectification" (ibid). Indeed some researchers would argue that mathematical understanding is not necessarily a linear progression (Sfard, 1991; Gray & Tall, 1991). For example in process-object theory, conceptual understanding in mathematics can be considered as both a process and a mental object, where encapsulation or reification of the object does not necessarily lead smoothly from understanding the mathematical process, rather it often requires a leap in understanding on behalf of the learner (see Sfard, 1991; Gray & Tall, 1992; Dubinsky & McDonald, 2001).

Process-object theories make the distinction between the act of doing procedural mathematics and the higher-level ability to consider mathematics as cognitive constructs, agreeing with Yackel's and Cobb's (1996) discovery that learners' explanations in an inquiry based classroom "increasingly involved describing actions on what to them were mathematical objects" (p.469), hence moving away from explanations which describe a followed procedure to seeing the structure of the mathematics as a whole and developing the ability to conceptualise rather than following by rote.

However, more generally than just through a process-object theory perspective, mathematical understanding is more than simply being able to follow a procedure. Many educational researchers have attempted to distinguish between the understanding in performing mathematics and the grasping of mathematical concepts (e.g. Michener, 1978; Skemp, 1976; Sfard, 1991). Skemp (1976) for example describes the difference as instrumental and relational understanding, where only relational understanding is considered to be true mathematical understanding. Michener considers this more conceptual understanding of mathematics as "an intuitive feeling for the subject, how it hangs together, and how it relates to other theories" (1978, p.1).

Fan and Bokhove (2014) on the other hand contend that there is a place in mathematics learning for algorithms, as they can contribute to higher-order thinking and mathematical understanding. This is as a result of how an algorithm is used as a cognitive process. For example, simply remembering an algorithm in order to use it requires lower-order thinking skills, however understanding how and why an algorithm works and evaluating the efficiency of algorithms, can pave the way to the learner creating their own algorithms which becomes a higher-order thinking skill (ibid). As discussed in a previous essay (Denton, 2013b), process-object theories do not necessarily imply a linear progression from following a procedure to formulating a concept. While the process is seen as the lower-order dimension, it does not mean that it cannot support the development of conceptual understanding (Sfard, 1991; Gray & Tall, 1992; Dubinsky & McDonald, 2001). According to Fan and Bokhove (2014) "[t]he problem is not in the algorithms themselves, but how to teach them effectively and, more, cognitively" (p. 491). That is, there is a place for the learning of procedures and algorithms in mathematics. Moreover, understanding can be derived from procedures when evaluating how an algorithm can solve a range of problems and when focus is given to the structure of the algorithm itself (Fan & Bokhove, 2014). Therefore it is possible for learners to be following a procedure, but for the teachers' questioning to elicit conceptual thinking from the learners for how and why the procedure works and how it can be adapted to solve a range of similar problems. Therein lies the importance of derivational thinking skills in learning and understanding mathematics.

In the original taxonomy (Bloom et al, 1956), the highest order learning objective was deemed to be evaluating, however in the revised taxonomy Anderson et al. (2001) introduced the classification of *creating* as the highest order thinking skill. In terms of mathematics, this term could be problematic to define, however it could be interpreted as the ability to derive new mathematics, through applying prior knowledge or adapting a known procedure in a new situation. This ability is vital in mathematics according to Bauersfeld (1993):

> Knowledge (in a narrow sense) will be for nothing once the user cannot identify the adequateness of a situation for use. Knowledge, also, will not be of much help, if the learner is unable to flexibly relate and transform the necessary elements of knowing into his/her actual situation (p. 4).

Bloom's Taxonomy could help the teacher establish social norms for developing learners' thinking in the classroom, but does not necessarily support teachers to develop sociomathematical norms specific for conceptual development in mathematics since the classifications could be applied across the curriculum rather than being specific to mathematics.

Taxonomies for Mathematical Thinking

Other researchers agree that Bloom's Taxonomy has limitations in its application to learners' mathematical thinking (for example Fan et al., 2004) and there have been attempts to devise mathematics specific taxonomies to overcome the criticisms cited above. One alternative taxonomy is the SOLO (Structure of Observed Learning Outcomes) Taxonomy of Biggs and Collis (1980), which proposes a sequence of unistructural, multistructural and relational understanding (Pegg & Tall, 2010), which Watson (2007) believes "can be used to devise questions which make finer distinctions than the vague notions of 'lower-order' and 'higher-order'" (p.115), however Figure 1.4 shows a visual interpretation of the SOLO Taxonomy as a hierarchy in a similar way to how Bloom's Taxonomy is often represented.

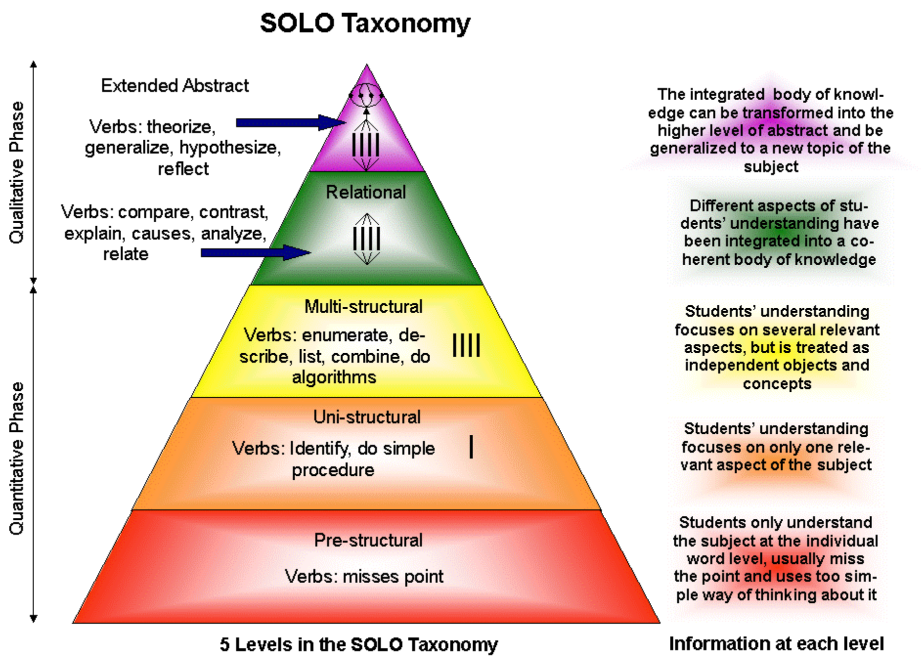


Figure 1.4. SOLO Taxonomy as a Hierarchy (Chan, 2010)

Furthermore Figure 1.5 demonstrates how the SOLO model can be aligned to Bloom's learning objectives. This highlights that, although there is not a one-to-one correlation between the stages of the two taxonomies, the SOLO model could provide the missing element of conceptual understanding in Bloom's taxonomy for the use of categorisation of thinking skills in mathematics.

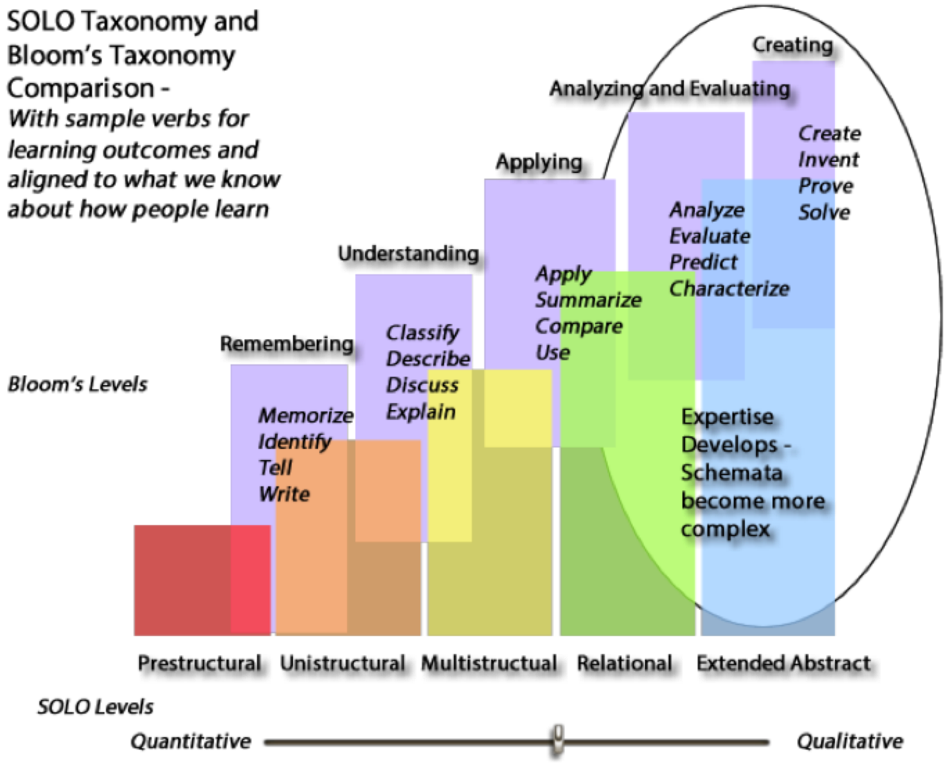


Figure 1.5. Comparison of SOLO and Bloom's taxonomies (Trekles & Sims, 2013)

The diagrams suggest that the SOLO model can introduce the missing element of conceptual understanding of mathematics not covered by Bloom's taxonomy, through consideration of structural thought. However, according to Chick (1998), "structure does not necessarily give an idea of mathematical value, but only of cognitive complexity" (p. 20), which is not enough on its own to assess mathematical performance (ibid).

Smith et al. (1996) developed the MATH Taxonomy (Mathematical Assessment Task Hierarchy) for constructing examination questions, which overcomes the issue above of a taxonomy which can be applied to the assessment of mathematical performance. In the MATH Taxonomy, the hierarchy is split into three groups (see Table 1.2) and difficulty is distinguished through the types of activity which require either a surface or deeper approach within each group (Smith et al, 1996), rather than a hierarchy of difficulty between groups.

Group A	**Group B**	**Group C**
Factual knowledge	*Information transfer*	*Justifying and interpreting*
Comprehension	Application in new situations	Implications, conjectures and comparisons
Routine use of procedures		Evaluation

Table 1.2. MATH Taxonomy (Smith et al., 1996, p.67)

Group A suggests procedural and instructional understanding as opposed to Group C which implies a more relational understanding. Group B concerns applying knowledge to new contexts and appears to be a bridging category between knowledge, represented by Group A, and justification, represented by Group C, that is, the ability to combine the understanding from the two categories to allow for the transference of knowledge in unfamiliar situations. Like the SOLO model, the MATH taxonomy seems to overcome some of the limitations of Bloom's Taxonomy in the application of mathematical thinking, however 'comprehension' is an ambiguous term, perhaps being the reason it was removed from Bloom's original taxonomy, and remains an ambiguous term in the MATH Taxonomy. A further limitation seems to be that an important classification is missing from the MATH Taxonomy, that of structural thought, which is central to the SOLO model and is key for developing mathematical thinking (Mason et al., 2009).

The MATH Taxonomy was subsequent to research by Marton and Saljo (1976a; 1976b) who first developed the terms *surface approach* and *deep approach* to learning at the same time as Skemp's (1976) instrumental and relational understanding. *Surface* and *deep* were approaches Marton and Saljo (1976a;

1976b) observed that were demonstrated by learners in their research. Interestingly Biggs, following the development of the SOLO Model, further developed the notion of adopting either a surface or deep approach to learning at higher education level (Biggs & Tang, 2007). Howie and Bagnall (2013) are very critical of the distinction between surface and deep approach to learning, particularly in relation to Biggs' SOLO Model, as they deem the "mantra of 'deep is good, surface is bad'" (Howie & Bagnall, 2013, p.394) as too simplistic, in that these could hold different meanings throughout the stages of the model. Howie and Bagnall (2013) argue that research into Asian approaches to learning are often categorised as surface, which implies a bad approach, and yet their outcomes in terms of academic results are very good.

The characteristics which determine whether a learner adopts a surface or a deeper approach to learning, are in part down to the approach taken by the teacher in encouraging connections in learners' understanding, as opposed to presenting mathematical ideas as a series of unconnected concepts (Howie & Bagnall, 2013). These are issues which could be addressed by teachers in their questioning by considering whether the questions they ask learners intend them to take a surface or deep approach to mathematical thought, however, just as Fan & Bokhove (2014) defend procedures and algorithms as having a place in the learning of mathematics, the place for a surface approach requires further research (Howie and Bagnall, 2013).

Another distinction in classifying mathematical thought is in *product-process* questioning (Muijs & Reynolds, 2011), where the product element is concerned with obtaining the result or answer whilst the process element of the questioning is focused more on explanation. However as discussed above in relation to process-object theory, in mathematics *process* is not necessarily considered higher-order thinking (Dubinsky & McDonald, 2001; Sfard, 1991) particularly if it is only concerned with the explanation of a procedure as opposed to justification of the methods used. For product-process questioning to be a meaningful way of classifying questions then Smith's et al.'s (1996) surface or deeper approach could be considered in conjunction with the process element.

Structural thought is also given due consideration in the seven mathematical foci of Andrews et al. (2005) for analysing teachers' behaviour (Table 1.3).

Conceptual	The teacher emphasises or encourages the conceptual development of his or her students.
Derivational	The teacher emphasises or encourages the process of developing new mathematical entities from existing knowledge.
Structural	The teacher emphasises or encourages the links or connections between different mathematical entities; concepts, properties etc.
Procedural	The teacher emphasises or encourages the acquisition of skills, procedures, techniques or algorithms.
Efficiency	The teacher emphasises or encourages learners' understanding or acquisition of processes or techniques that develop flexibility, elegance or critical comparison of working.
Problem solving	The teacher emphasises or encourages learners' engagement with the solution of non-trivial or non-routine tasks.
Reasoning	The teacher emphasises or encourages learners' development and articulation of justification and argumentation.

Table 1.3. Mathematical Foci (Andrews et al., 2005, p.11)

These foci describe "the intentions of teaching through classifying features of mathematical meaning and structure without assuming that learners necessarily do what is intended" (Watson, 2007, p.116). A criticism of the seven foci could be that the problem-solving focus is more of a means of achieving the other foci, for example engaging in non-routine tasks could be seen as developing derivational thinking, and engaging in non-trivial tasks could be taken as the development of conceptual thinking. Furthermore the focus of efficiency could perhaps be alternatively described as a reflective focus, that is, to evaluate or criticise approaches to the mathematics in order to develop the sociomathematical norm of mathematical efficiency.

Although the taxonomies described above could be used for the classification of questioning, they were in fact developed for a range of other purposes. For example Bloom's Taxonomy was developed to classify learning objectives; the MATH Taxonomy was developed for the classification of examination questions and Andrews' et al.'s (2005) Mathematical Foci were intended to analyse teachers' behaviour more generally than just in the questioning used with learners.

There are however structures which were designed specifically for the classification of questions used in the mathematics classroom. More recent research by Orrill (2013) specifically analysed the proportions of different question types posed by four teachers in video recorded mathematics lessons. Lessons were then transcribed and coded for question type, "where type was an indicator of the way in which the question was posed (thus, a hybrid of form and purpose)" (p. 288). The results are shown in Table 1.4, where it can be seen that the highest proportion of question type asked by every teacher was a 'fill-in-the-blank' question, closely linked to the closed categorisation previously discussed. With the exception of one teacher, very few questions were asked where learners had to assess an element of the task and very few questions required learners to justify their thinking.

Orrill (2013) considers *question type* to be just one dimension in classifying questions, alongside *representation*, *mathematics* and *other*. Each question was coded independently in each dimension and was labelled uncategorised if it did not fit within any category in that dimension. This explains the high proportion of uncategorised in the table, as the question presumably was not relevant to that dimension.

	Deborah	Cecilia	Barb	Kevin
Question type				
Self-answered	4 (3 %)	0	4 (2 %)	1 (1 %)
Fill-in-the-blank	84 (73 %)	98 (66 %)	116 (70 %)	70 (44 %)
Who is this	0	0	14 (8 %)	5 (3 %)
Follow-up	6 (5 %)	25 (17 %)	17 (10 %)	29 (18 %)
Open-ended	9 (8 %)	16 (11 %)	12 (7 %)	28 (18 %)
Assess idea	1 (1 %)	1 (1 %)	2 (1 %)	11 (7 %)
Justification/argument	2 (2 %)	7 (5 %)	2 (1 %)	7 (4 %)
Uncategorized	9 (8 %)	1 (1 %)	0	9 (6 %)
Representation				
Analyze something on screen	59 (51 %)	33 (22 %)	76 (46 %)	58 (36 %)
Make a prediction	9 (8 %)	18 (12 %)	5 (3 %)	15 (9 %)
Make a connection	10 (9 %)	23 (16 %)	25 (15 %)	27 (17 %)
Uncategorized	38 (33 %)	74 (50 %)	61 (37 %)	60 (38 %)
Mathematics				
Recall	32 (3 %)	23 (16 %)	22 (13 %)	8 (5 %)
Apply	39 (34 %)	27 (18 %)	36 (21 %)	16 (10 %)
Analyze/evaluate	6 (5 %)	12 (8 %)	3 (2 %)	9 (6 %)
Reflect	0	0	0	1 (1 %)
Uncategorized	38 (33 %)	87 (59 %)	106 (63 %)	126 (79 %)
Other				
Teaching the technology	31 (27 %)	31 (21 %)	20 (12 %)	17 (11 %)
Logistics	39 (34 %)	53 (36 %)	51 (31 %)	43 (27 %)
Uncategorized	44 (38 %)	64 (43 %)	90 (54 %)	99 (62 %)

Table 1.4. Table of results showing proportions of question types asked during whole class instruction by each teacher (Orrill, 2013, p. 291)

The taxonomies developed specifically for mathematics could give a starting point to support teachers to establish sociomathematical norms to develop conceptual thinking and justification in mathematics. Andrews' et al.'s (2005) seven mathematical foci and the SOLO model described above, consider the development of learners' structural knowledge specific to mathematics, and so could be used to establish sociomathematical norms. Andrews et al. (2005) in fact do this to some degree in their research, as they note that that mathematical efficiency was not given particular emphasis in the lessons they observed.

The categorisations in Smith's et al.'s (1996) MATH Taxonomy, for example factual knowledge, information transfer, and justifying and interpreting, could be considered social norms as they are all expectations which we would expect to be set in other subject areas (Yackel & Cobb, 1996), hence the categorisations, although relevant to mathematics, are not unique to the learning of mathematics. As a result, using the MATH Taxonomy would not necessarily be classed as addressing sociomathematical norms without specific reference to mathematical thinking in a particular question.

Higher-Order Questioning Vs Higher-Order Thinking

Holster (2006) states that questioning is "a vital mode of interactive discourse in the classroom" (p.4), however it is more important to consider questions which elicit higher-order thinking as opposed to identifying higher-order questions, as according to Kawanaka and Stigler (2000), "asking more higher order questions does not simply improve student learning" (p. 255). Coding of question types needs to consider the pedagogy behind the question in relation to the task rather than just the question itself (Holster, 2006; Kawanaka & Stigler, 2000) so a question cannot be viewed in isolation from whether it can be considered higher-order; it must be viewed in the context in which it was posed. Furthermore it is possible to pose a question which appears open or higher-order but to follow this question with subsequent questions which are closed or direct the learner to the answer expected by the teacher, thus likely inhibiting higher-order thinking.

Kawanaka and Stigler (2000) researched into the discourse in classrooms in Germany, Japan and the United States; they found that the majority of questions asked by teachers in all three countries were lower-order questions and the phase of the lesson was one of the contributing factors to when higher-order questioning was used. In their coding of lessons, they also found that questions relating to descriptions and explanations cannot be considered higher-order questions unless they produce higher-order thinking from the learner.

Recent research by Drageset (2014) considers how learners respond to a given question to be vital in the analysis of classroom discourse. The notion of examining the expected response or intended mathematical thinking as opposed to the question itself is an interesting one and supports the stances of Kawanaka and Stigler (2000), Yang (2006) and Holster (2006), that questioning cannot be analysed in isolation from the context of the classroom, the learners and the discourse.

Questioning in mathematics and eliciting meaningful responses is impacted by the sociocultural-mathematical norms in the classroom (Mason, 2014), that is, if the teacher asks simple questions requiring low level responses then learners will not develop mathematical autonomy. Similarly, according to Mason (2014), if the teacher does not vary the type of question they pose, then learners do not learn to pose questions themselves. Therefore questions which seem lower-order, for example "What am I going to ask you?", become a powerful tool for developing autonomy, as the learner may then begin to ask such questions independently of the teacher.

The Role of Teacher Knowledge

As described earlier in this chapter, the MATH Taxonomy classifies mathematical tasks in terms of the type of thinking expected from the learner. Chapman (2013) discusses the teachers' role in developing intended mathematical thinking:

> A teacher could turn an open-ended task into a closed one or a closed one into an open one. He or she could treat a task of high cognitive demand as a low level one or vice versa. (Chapman, 2013, p. 1)

This notion is echoed by Steele (2013) who states that the depth of teachers' understanding of the activity has implications for teaching in a way that will "either maintain the cognitive demand or proceduralize the task" (p. 261). However, this is not just dependent on the teachers' subject knowledge, but more the teachers' abilities to impart this knowledge in a manner that learners can construct new mathematical knowledge for themselves (Barwell, 2013; Chapman, 2013). Chapman (2013) lists factors which could inhibit mathematical thought, including the teachers' knowledge and preferred pedagogies. It is not enough to simply provide an activity which is designed to develop learners' understanding of mathematics; the teacher must also "optimize the learning potential of such tasks" (p. 1).

Campbell at al. (2014) distinguish between *mathematical content knowledge* and *pedagogical content knowledge*. The former refers to the teachers' knowledge of the syllabus and the underpinning mathematics associated with it. The latter refers to the teachers' understanding of mathematics teaching and learning, including recognising and addressing common learner misconceptions, and understanding learners' interpretation of the mathematics. Campbell et al. view these as two distinct types of knowledge, but recognise where the two are linked and found that, perhaps unsurprisingly, both types of teacher knowledge have a statistically significant impact on learners' attainment in pre-high school education in the United States.

Teachers need guidance for how to ask meaningful questions, as highlighted by Moyer's and Milewicz's (2002) research into trainee teachers' use of questioning, where they found the following common themes in observed discourse: *checklisting*; *instructing rather than assessing*; and *probing and follow-up*. Checklisting is what Moyer and Milewicz (2002) call proceeding to the next question regardless of the response and they found this was common in prospective teachers. Instructing rather than assessing was also a common theme, whereby trainee teachers would ask leading questions or simply tell the learner the answers. Finally there was some good practice with trainee teachers asking probing and follow-up questions, however this was, more often than not, questioning only an incorrect answer as opposed to the learner's conceptual thinking more generally. Thus the teachers' knowledge was in need of development.

Using Lesson Study to Develop Teacher Knowledge

According to Puchner and Taylor (2006), higher-order thinking skills can be developed in learners through teachers following a *lesson study* approach to their professional development. Lesson study is a Japanese teacher development concept where teachers observe each other's lessons, discuss the issues, plan together and then observe again (Lewis et al., 2006). A benefit of the approach is that it is driven by the teachers themselves, although a lesson study advisor is recommended (Puchner & Taylor, 2006). Knapp et al. (2011) provide case study evidence that this approach could improve the quality of questioning used by teachers in other cultures. This is because lesson study allows the teacher to analyse the learners' thinking, including any misconceptions they have and then plan for questioning which addresses these misconceptions (Olson et al., 2011).

However the lesson study approach requires a significant commitment to time being built into teachers' timetables (Olson et al., 2011) to allow teachers the time required to meet to plan and discuss as well as the time to be in one another's classrooms. With the current teacher shortages in mathematics classrooms in secondary schools in the UK (Thornby et al., 2016), and indeed in many other countries (Smithers & Robinson, 2013), this model is difficult to replicate in many other cultures.

Formative Assessment

As discussed in a previous essay on the use of questioning in mathematics (Denton 2013a), Black and Wiliam (1998) state that discourse in the classroom should be "thoughtful, reflective, focused to evoke and explore understanding, and conducted so that all pupils have an opportunity to think and express their ideas" (p.12). Black and Wiliam (1998) identify that thinking is inhibited by two main factors. Firstly that many teachers direct learners towards an answer and hence discourage the learners from developing their own ideas. Secondly learners are often not given enough "quiet time" (ibid, p.11) to think about a question before a response is expected. *Quiet time* in this sense is more commonly known as *wait time* (Rowe, 1986). Teachers often feel uncomfortable by a silent pause in class discussion which results in teachers answering their own questions in an attempt to maintain pace in the lesson (ibid). Black and Wiliam (1998) attribute these inhibitors as the reason why only a minority of the class participate in whole class discussion. Black and Wiliam (1998) state that learners are also inhibited if they believe that the teacher is looking for a predetermined response. Cobb and Yackel (1996) echo this in their research on establishing sociomathematical norms, in that they observed that learners often "infer the response the teacher had in mind rather than [to] articulate their own understandings" (Cobb & Yackel, 1996, p. 178)

In his research on the use of wait time in the classroom, Rowe (1986) found that teachers, across subjects and phase of schooling, typically pause for less than one second both after posing a question and after a response is given, not allowing learners the time to process the question asked by the teacher or the response given by another learner. Rowe (1986) concludes that wait time is crucial to allow students to both think through and expand upon responses. These conclusions are supported by research by Black et al. (2003). Yang (2006), however, is critical of the notion of wait time, as he contends on its own it does not guarantee that learners are using the time to think mathematically. Instead of an "absolute amount of pausing time" (ibid, p.198), Yang suggests ensuring questions are meaningful to the learner and are given a "subjective time duration" (ibid).

In order to overcome these inhibitors to learner participation, Black and Wiliam (1998) recommend that teachers should do the following: increase the time given to learners to respond to a question; provide opportunities for learners to discuss their ideas in pairs or small groups first; provide learners with a list of possible answers to choose from; and get the whole class to write down an answer, then select a few members of the class to share their responses.

Following Black's and Wiliam's (1998) research, formative assessment became a focus for schools, supported by the National Strategies' Assessment for Learning (AfL) materials (Ofsted, 2008). However in a review of AfL a decade later, Ofsted (2008) concluded that despite the resources which had been produced and shared with schools, and despite training given to schools and teachers on developing formative assessment techniques, teachers still needed to "develop their skills in targeting questions to challenge pupils' understanding, prompting them to explain and justify their answers individually, in small groups and in whole class dialogue" (ibid, p. 7).

Hodgen and Wiliam (2006) states that by "exploring and 'unpacking' mathematics, students can begin to see for themselves what they know and how well they know it" (p.5). In this later work, he exemplifies Black's and Wiliam's (1998) recommendations with mathematical specific strategies. These strategies include the use of mini-whiteboards by all learners in the classroom, thereby enabling all learners to think about and come up with a solution to a question posed, generating discussion from incorrect answers. Watson's and Mason's (1998) 'Show me…' questions have potential as an effective mini-whiteboard questioning strategy, to achieve rich discussion amongst learners and provide the teacher with opportunities to give formative feedback. Discussion can also arise when the teacher poses questions which have multiple solutions (Hodgen & Wiliam, 2006).

A Need to Identify More Effective Questioning Strategies

Black et al. (2006) believe effective questioning is essential to develop metacognition and self-awareness, so learners "can ask questions of each other and the focus can move from the teacher to the pupils" (p.128). Furthermore Yackel and Cobb (1996) describe the central role that the teacher plays in establishing sociomathematical norms and in developing the expectations of what constitutes an acceptable mathematical explanation and justification, mathematical difference, mathematical efficiency, and mathematical sophistication. However achieving the establishment of sociomathematical norms in the classroom "relies on teachers' abilities and willingness to ask questions that engage students in higher-level reasoning" (Orrill, 2013, p. 286).

This literature review has demonstrated that there exists a wealth of research on questioning (e.g. Mason, 2000; Fraivillig et al., 1999; Orrill, 2013; ...) and classification of thinking skills has been a focus of educational research for over half a century (Bloom et al.,1956; Gall, 1970; Andrews et al., 2005). Furthermore, Assessment for Learning (AfL) techniques have been a high priority for schools in recent years (Ofsted, 2008). However according to Ofsted (2012), despite this wealth of research, questioning still requires development in the mathematics classroom, agreeing with empirical research which has shown that mathematics teachers still use a large proportion of lower-order questioning (Orrill, 2013).

Orrill (2013) contends that mathematics teachers struggle with asking higher-order questions in the classroom which lead to higher-order thinking. He suggests that such struggle could be addressed through teachers working collaboratively through lesson study, however as described above, this process requires a commitment to time and staffing which, at the time of writing this thesis, are often scarce in mathematics departments (Thornby et al., 2016; Smithers & Robinson, 2013). What is needed, therefore, is further research "to identify and characterize more effective questioning strategies" (Orrill, 2013, p. 287) which are easily accessible to mathematics teachers and which they can put into practice in their classrooms with minimal training. Thus I decided to develop a new taxonomy to address this need. The empirical dimension of this thesis explores its effectiveness and the pilot is described in the next chapter.

2. Pilot Study

This chapter describes the pilot study for the thesis which was conducted as part of the Foundation Research Methods module towards the degree of Master of Science. The research design of the pilot study is outlined, including the research strategy, methodology and methods used, as well as sampling and ethical considerations. The results of the pilot study are presented and discussed, concluding that the proportion of surface questioning outweighs a deeper approach and finds that while Assessment for Learning (AfL) techniques increase the variety of question type used by teachers in their questioning, such techniques are in practice used more by teachers to increase learner participation rather than as tools to probe mathematical thinking. How the findings of the pilot study shaped the main study is also considered along with the limitations of the pilot study.

Pilot Research Hypotheses

The pilot study set out to trial an original version of the IMPaCT Taxonomy and researched the following hypotheses:

1. **A larger proportion of questions requiring a 'surface approach' are used in mathematics lessons than those requiring deeper thinking.**
2. **Using formative questioning techniques supports a wider range of intended mathematical thinking, compared to questions posed without these techniques**.

These hypotheses were investigated using a combination of the MATH taxonomy framework (Smith et al., 1996) and Andrews' et al.'s (2005) mathematical foci. Using these tools, a new framework for classifying questioning was formulated, an early version of the IMPaCT Taxonomy (Appendix 1) which classifies question types in mathematics into the following categories: factual; procedural; structural; reasoning; reflective; and derivational. Each of these categories were supported by question prompts proposed by Watson (2007) and Hodgen and Wiliam (2006), to analyse the intended purpose of the types of questions used in mathematics lessons and whether formative questioning techniques supported deeper thinking.

Research Design in the Pilot Study

A combined methodological approach was used to address these pilot hypotheses combining deductive reasoning (Hartas, 2010), using quantitative methods for statistical analysis of proportion of question types and techniques employed by teachers in observations, and inductive reasoning (ibid), utilising qualitative research through interviews to probe deeper into the teachers' intent and allow for new themes to emerge.

The teachers' voices were heard through radical listening (Clough & Nutbrown, 2007). This was achieved through semi-structured interviews which allowed the participants to speak freely about their experiences while at the same time ensuring that key questions were asked. Learner voice was heard through the more quantitative method of questionnaires. This methodological triangulation was used to ensure that the research data was viewed from as many perspectives as possible (Denscombe, 2007). In addition, if the outcomes corresponded then greater confidence could be had in the findings (Cohen et al., 2007).

The pilot study was conducted at a federation of two 11-16 single-sex schools which was my place of work at the time of conducting the pilot study. The federation was non-selective and located in a local authority borough where two out of the ten secondary schools were grammar schools. The proportion of students in the school who were from minority ethnic backgrounds or who spoke English as an additional language was above average. The context of the school is important here to define the constraints for comparison with the other school used in the main body of this research as there is a limitation that any findings from the pilot study could be unique to that federation of schools. This issue will be addressed in more detail later.

Sampling Considerations

For the pilot study, a sample of classes from both schools in the federation was chosen to ensure the results were as representative as possible (Cohen et al., 2007). A cluster sample of four classes with which to observe four teachers was chosen using both convenience and purposive sampling techniques. A degree of convenience sampling was necessary due to the federation's policy to minimise the cost of covering lessons, therefore the classes chosen had to be those which took place in my non-contact time. However it could also be considered to be purposive sampling, as I consciously aimed to include as broad a range of learner ages, gender and level of attainment as possible (Denscombe, 2007). The participant groups consisted of a mixed attainment Year 7 boys class; Year 8 girls, set 1 of 5; Year 9 boys, set 5 of 6; and Year 10 girls, set 1 of 7. While this included a range of abilities and genders, the girls in this sample were both older and higher attaining than the boys which needed consideration in the conclusions drawn from the pilot study.

Due to the time it would have taken to enter data for up to 120 participants, a stratified sample of 30 learner questionnaires was calculated (Table 2.1). This allowed for proportional groupings from the population (Robson, 2002) to ensure the results were a fair representation of the Federation. Systematic sampling was used to select the learners from each category as this method of sampling is considered an efficient form of probabilistic sampling, provided the questionnaires within each group were randomised first (Cohen et al., 2007), giving each member of a group an equal chance of being chosen, hence minimising the potential for bias in the sample.

Group	KS3 Male	KS3 Female	KS4 Male	KS4 Female
Number of students in whole federation	199	226	295	344
Calculated sample stratified by Key Stage and gender	5.61	6.37	8.32	9.699
Number of students for sample (rounded)	6	6	8	10
Number of completed questionnaires	28	25	21	24
Systematic sample of students (once data is randomised)	every 5th	every 4th	every 3rd	every 2nd

Table 2.1. Pilot Study Sampling Calculations (Denton, 2013a)

Ethical Considerations

All participants involved were informed of the purpose of the research, how the observations, questionnaires and interviews would be used and were assured that any information given was confidential (BERA, 2011). Participant teachers were also informed of their right to withdraw from participating at any time. To assure confidentiality and anonymity, the participants are referred to as Teacher X etc. and names were not requested on the student questionnaires (ibid). In line with school policy, the Head Teacher gave consent including, in loco-parentis, to conduct the learner questionnaires.

The Learner Questionnaires

An initial questionnaire (see Appendix 2), which was trialled on six learners, took between five and ten minutes to complete. As a result of this trial, in a revised questionnaire (see Appendix 3), further clarification was given in question 5 and two questions were removed: question 7 due to co-linearity with question 6, and question 9 because of both redundancy and reliability issues (Cohen et al., 2007).

The revised learner questionnaires were distributed at the end of each observed lesson and the class completed them in exam-like conditions at their own pace. They were collected in an envelope to ensure anonymity (Cohen et al., 2007). As a result of this method of conducting the questionnaires, the response rate of sufficiently completed questionnaires was 98.0%, excluding absentees (Denton, 2013a). Another benefit of this method was the opportunity to clarify the meaning of questions to the participants if needed, which was especially useful with the younger participants. Unfortunately, despite this advantage and having already made alterations, Question 7 remained problematic so was not analysed or included in the results.

SPSS was used to analyse the quantitative data collected from the questionnaires. This enabled statistical comparison of both distribution and proportions (Muijs, 2010) of which AfL techniques were used, in learners' opinions, in their lessons. SPSS was also used to analyse learners' perceptions of the importance of the final answer in mathematics, as opposed to the thinking process to arrive at a solution. This use of a quantitative method to collect learners' opinions enabled a larger number of participants to be sampled than in conducting interviews, as the closed questions data were converted to nominal or Likert scales for ease and speed of analysis. Another benefit of using questionnaires was that the data collected were anonymous, enabling participants to be honest about their experiences. The last question was open, to ensure a degree of flexibility in participants' responses by allowing them to elaborate on anything they felt necessary, however this question was answered by only four participant learners. As a result of this, any findings from these responses may not be representative of the population (Cohen et al., 2007), so those findings are not included in the results section of the pilot study.

Lesson Observations

Lesson observation data were collected in a "non-participant observer role" (Cohen et al., 2007, p. 259) where I sat in the lessons and scribed every question the teacher asked either the whole class, groups or individuals.

Although the main purpose of this was to minimise bias in the validity of the data collected, it was also became apparent that there was not the time to do anything other than transcribing the lesson anyway. The data were coded using the descriptive coding system (Miles & Huberman, 1994), shown in Table 2.2, and statistically analysed in Excel to compare the proportions of question type and depth of intended mathematical thinking in the observed lessons. Similarly a coding system was used for the questioning techniques observed (see Appendix 4).

Question Type *Adapted from Smith et al. (1996) and Andrews et al. (2005)*	**Surface Approach Coding**	**Deeper Approach Coding**
Factual	SUR-FAC	
Procedural	SUR-PRO	
Structural	SUR-STR	DP-STR
Reasoning	SUR-REA	DP-REA
Reflective	SUR-REF	DP-REF
Derivational	SUR-DER	DP-DER

Table 2.2. Descriptive coding system for question type (Denton, 2013a)

Inter-observer reliability was tested (Robson, 2002), and as a result it was decided that factual questioning could not take a deeper approach as it was simply a matter of recall of knowledge. Deep procedural questioning was also eliminated from the coding table due to the ambiguity of deep procedural being considered either derivational, if adapting a procedure to apply to a new situation, or structural, if the reasons why a procedure will always work are explored.

Participant Teacher Interviews

Semi-structured teacher interviews (see Appendix 5), allowing flexibility in terms of the adaptation of questions during the interview (Robson, 2002), were conducted with the teachers after the observed lessons, to explore their views on what the intended mathematical thinking of their questions had been

(Miller & Glassner, 2004). The disadvantages and limitations of using this method was that the interviews were time-consuming and that teachers might just have said what they thought I wanted to hear. This could be considered a threat to the validity of the data collected (Cohen et al., 2007), however participant teachers were assured of the non-judgemental nature of the observations and subsequent interviews and reassured that their lessons were not graded, which should have minimised this threat.

A descriptive coding framework was used for the analysis of the interviews to draw out emerging themes (Delamont, 2002). As topics emerged they were assigned a new code and then grouped together into themes (see Table 2.3).

<table>
<tr><th>Topic in Interview</th><th>Code</th><th>Theme</th></tr>
<tr><td>No hands up</td><td>NHU</td><td rowspan="4">AfL used to increase participation</td></tr>
<tr><td>Random methods to select</td><td>RAN</td></tr>
<tr><td>Use of mini-whiteboards</td><td>MWB</td></tr>
<tr><td>Wait time</td><td>WT</td></tr>
<tr><td>Linking question type to topic
Linking question type to perceived “ability”</td><td>TOPIC
ABILITY</td><td>Perceived factors dictating question type</td></tr>
<tr><td>Estimates of proportions surface/deeper
Teachers intent was different from outcome</td><td>PROPORTION
INTENT</td><td>Planning for questioning</td></tr>
</table>

Table 2.3. Emerging themes from interviews (Denton, 2013a)

Pilot Study Findings – Surface versus Deeper Approach

In the observed lessons, the vast majority (88%) of questions posed by the teachers required a surface approach to the learners’ mathematical thinking; however there was variation in the ratio of surface to deeper questioning between the four lessons observed (see Figure 2.1).

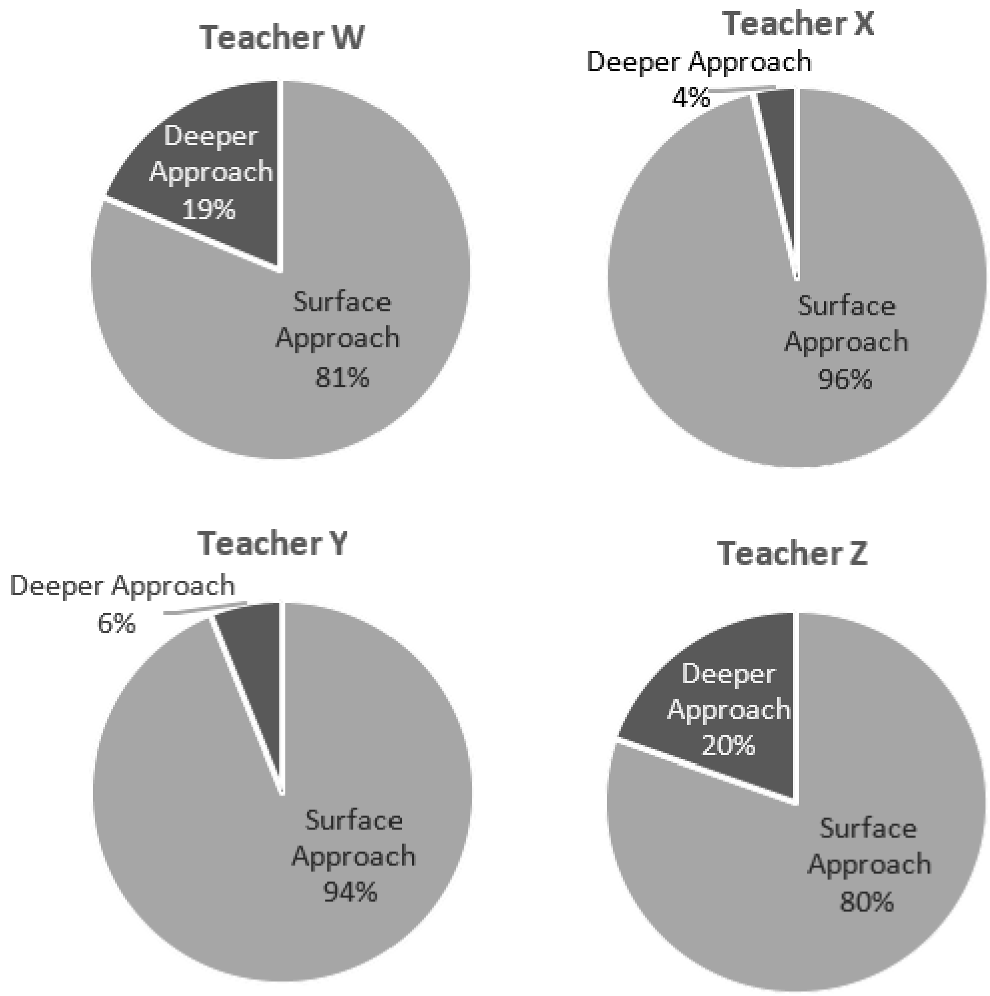


Figure 2.1. Percentage of surface and deeper questioning for each teacher

Teacher W was not surprised however that over four fifths of the questioning was surface level as this had been planned with a specific purpose in mind:

> **Teacher W**: The vast majority of the questions at the start were a surface approach to get the lesson going and involve everyone.

Whilst I do not dispute the need for surface questioning in mathematics, particularly in terms of establishing and maintaining pace in the lesson, there seemed to be missed opportunities in some lessons for more probing and structural questioning which could have given depth to learners' understanding. For example in the Year 9 lesson on multiplying decimals a far greater emphasis was given to the procedure of adding zeros and moving the decimal point when multiplying by powers of ten than strengthening learners' understanding of place value with respect to how many times bigger or smaller their answer needed to be, derived from multiplication facts they already knew. More planned questioning on the structure of the mathematics involved, rather than the following of rote procedures, could, in my view, have achieved a deeper understanding of this concept.

In the teacher interviews following the observations, all participants except Teacher Z underestimated the proportion of surface questioning in their lesson, for example:

> **Teacher X:** I'd estimate 80% [surface] this lesson due to the type of topic; different topics would be different amounts. Geometry and Algebra need more probing, deeper questioning, but not really Number. BODMAS [order of mathematical operations] could be deeper.

A common theme which emerged from the interviews was a belief that the topic being taught impacted the complexity of questioning the teacher employed, as seen in the extract above. However, the topics which teachers believed to require a deeper approach to questioning did not correlate between teachers. Teacher X viewed the order of mathematical operations as a topic which lends itself to deeper questioning, however Teacher Z perceived the exact same topic as requiring a more surface approach.

The limitation that the girls' classes in the sample were both older and higher attaining on average than the boys also did not seem to have an effect on the depth of the questioning employed. The two classes with the highest proportion of deeper questioning were the Year 7 boys and the Year 10 girls classes, suggesting that age and attainment were less of a factor on the depth of the questioning employed in the lesson than the effect of the teacher.

The data collected in the pilot study do not support the notion that the level of complexity of questioning is determined by the mathematical topic or concept being taught. In their observed lessons, both Teachers Y and Z were teaching specific 'functional' activities which lend themselves to more open questioning

(DfES, 2007), however Teacher Y only asked 6% deeper level questions compared to 19% for Teacher W who was teaching simultaneous equations without any contextual links. The results from the pilot study suggest that the depth of intended mathematical thought in the questions asked is in fact more attributable to the teacher than the topic, however a greater number of observations would be required to verify this theory, due to the small sample size in the pilot study and only one lesson observed for each participant teacher. This does however correspond with Chapman's (2013) findings that even if a rich mathematical task is used with learners, it is how the teacher elicits the understanding of the mathematics from that task which dictates the effectiveness of the learning outcomes.

Analysis of Types of Questions Observed

The quantitative data collected from lesson observations, which compares the percentages of surface and deeper approaches for each question type (see Figure 2.2), also supports the qualitative observations that probing the structure of the mathematics on a deeper level was a rare occurrence compared with derivational questioning which took a deeper approach 50% of the time. Reasoning and reflective questioning were coded as taking a deeper approach in over a quarter of the questions posed in these categories. As discussed above, factual and procedural questioning were only coded as surface level.

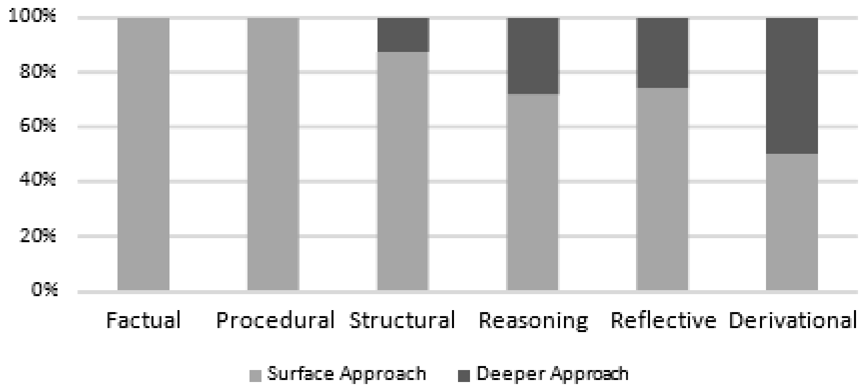


Figure 2.2. Percentages of surface and deeper approach

It was in fact the derivational category that had the highest percentage of deeper level questions (see Figure 2.2), nearly double the percentage for reasoning and reflective. This could however be misleading as the derivational category in its entirety only made up 2.5% of all questions asked in the observed lessons (see Figure 2.3). The high proportion of deeper level questioning in this category could simply be down to derivational questioning lending itself to deeper thinking through the application, adaptation and transference of prior knowledge.

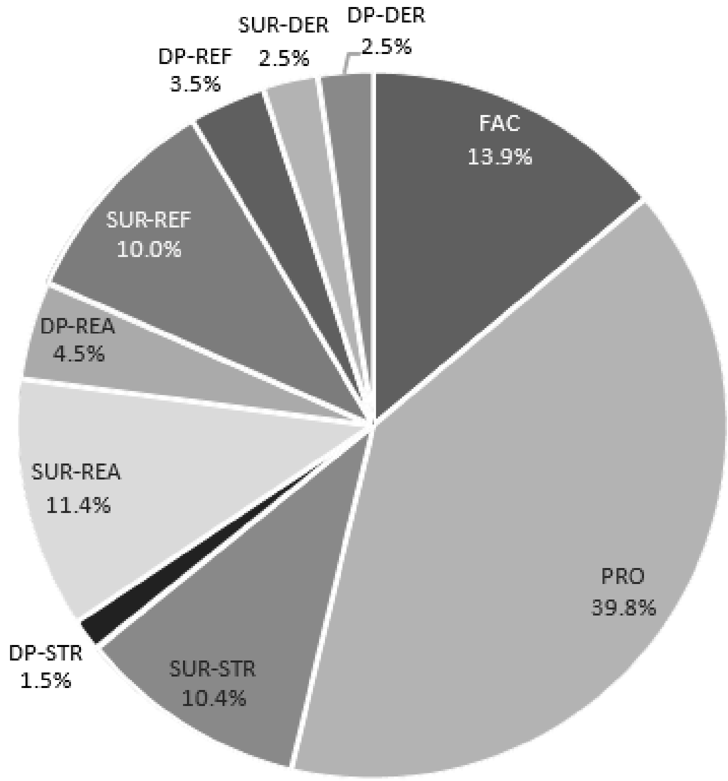


Figure 2.3. Overall percentages for each category

Questions asked in the observed lessons which required structural thinking made up 12% of all questions observed. However only 12.5% of those questions required deeper thought, totalling only 1.5% of all questions asked being deep structural (see Figure 2.3).

From Figure 2.3 it can also be seen that 53.4% of all questions were either factual or procedural. This could result in learners not having a structural understanding of the concept. For example the following questions were asked on a number of occasions to the whole class, small groups and individuals in the Year 9 lesson on multiplying decimals:

> **Teacher X:** How many decimal places have we got in all? So where will you put the decimal point?

Later in the lesson, many learners in the class concluded that 0.4 x 0.06 = 0.24 instead of 0.024. This reliance on procedure as opposed to understanding the structural knowledge of place value could have led to such a misconception or perhaps learners are simply calculating 4 x 6 and choosing a related decimal solution. Either way, these learners did not seem to consider the question structurally nor did they seem to reflect on their response by considering the magnitude of their answer or by putting it into a familiar context.

Furthermore, the absence of structural questioning in the same lesson could also have contributed to why questioning on reasoning only seemed to expect a surface level response:

> **Teacher X:** How did you get that [2.4 x 0.2]?
>
> **Pupil A:** 24 x 2 is 48…That has one decimal place [points to 2.4] and so does that [0.2], so the answer has two decimal places.

The teacher walked away from this discussion seemingly content with the learner's response as their thinking was not probed any deeper at the time. Such missed opportunities for deeper reasoning "rather than elaborating surface procedural knowledge" (Watson and Barton, 2011, p. 77) could have been avoided through not just asking *how* a procedure was used but also *why* it works structurally. The teacher demonstrated to learners that a procedural explanation and justification is acceptable in that classroom and the learner responded accordingly. This links back to the warning from Yackel and Cobb (1996), that if teacher only asks questions which require lower-order thinking, then superficial answers will become a classroom norm.

The analysis of learner questionnaires indicates that learners are taught the importance of the thought process in mathematics as opposed to simply the end solution (Figure 2.4). This is evidenced by the majority of learners responding that the process is of greatest importance, while only 6.7% of the sample believes the final answer is of greatest importance. However the qualitative analysis of lesson observations show that 71.9% of questioning on reasoning is on a surface level, expecting the *how* and not the *why*.

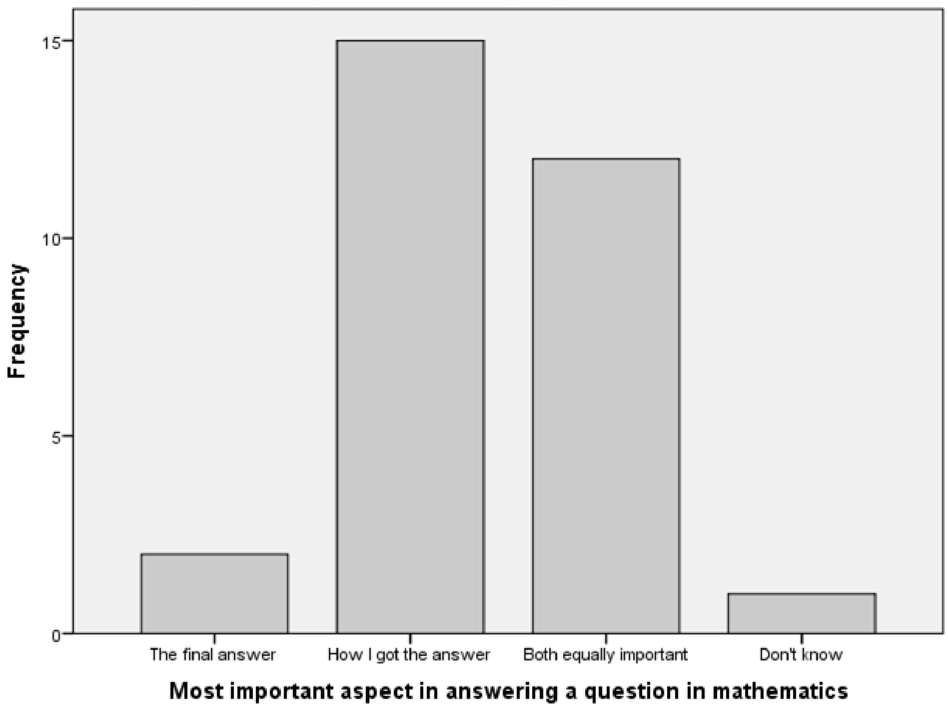


Figure 2.4. Learners' perception of the importance of process (Denton, 2013a)

The analysis of the learner questionnaires also indicates that learners are asked to explain their answers in the pilot school as 76.7% of the sample responded that they we asked to explain their answers frequently (Table 2.4). However, only 26.7% were asked to explain another learner's answer on a regular basis (Table 2.5).

Teacher asks learners to explain their answer

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Every Lesson	8	26.7	26.7	26.7
	Most Lessons	15	50.0	50.0	76.7
	Sometimes	5	16.7	16.7	93.3
	Never	1	3.3	3.3	96.7
	Don't know	1	3.3	3.3	100.0
	Total	30	100.0	100.0	

Table 2.4. Expectations for learners to explain their answers (Denton, 2013a)

Teacher asks learners to explain an answer someone else has given

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Most Lessons	8	26.7	26.7	26.7
	Sometimes	17	56.7	56.7	83.3
	Never	2	6.7	6.7	90.0
	Don't know	3	10.0	10.0	100.0
	Total	30	100.0	100.0	

Table 2.5. Expectations for learners to explain other learners' answers (Denton, 2013a)

As discussed in the literature review explaining someone else's answer could lead to deeper reasoning as the learner does not necessarily know *how* the other learner arrived at the solution but could explain *why* the answer must be correct (Yackel & Cobb, 1996; Martino & Maher, 1999). This supports Watson's and Barton's (2011) stance that questioning should draw on the "mathematical knowledge and experience we have, and on the ways we have individually encapsulated it" (p. 77) and Hodgen's and Wiliam's (2006) stance that by "exploring and 'unpacking' mathematics, students can begin to see for themselves what they know and how well they know it" (p.5)

In the teacher interviews, the perception emerged that factual, procedural and surface-reasoning questioning, that is, the expectation to explain *what* you did to get the answer as opposed to *why*, are question types for low attainers, while deeper questioning is more relevant for higher attainers. This could not be corroborated with the data collected from the lesson observations, as comparisons were only made between the hypotheses' variables and not between levels of attainment, gender and age groups of classes. This was due to the fact that any observed differences could be attributable to differences in the characteristics of the participant teachers as opposed to the differences in the learners they were teaching. This is an aspect which was worthy of further investigation in the main body of this research.

Assessment for Learning Techniques

From the quantitative analysis of lesson observations, 18.9% of questions posed by the teacher were asked in conjunction with a formative learning technique. Consequently, the number of questions available for analysis for the second hypothesis in the pilot study is greatly reduced from 201 to 38; however a data set above 30 in size can be considered large enough for statistical analysis (Hogg & Tanis, 2014).

The data linking questioning and the employment of AfL techniques suggest that when using AfL techniques, either explicitly or implicitly, teachers ask a broader range of question types (see Figure 2.5). The percentage of factual and procedural questions decreased significantly when employing these questioning techniques to just below 16%, compared to 63% of questions posed being factual or procedural without the use of formative techniques.

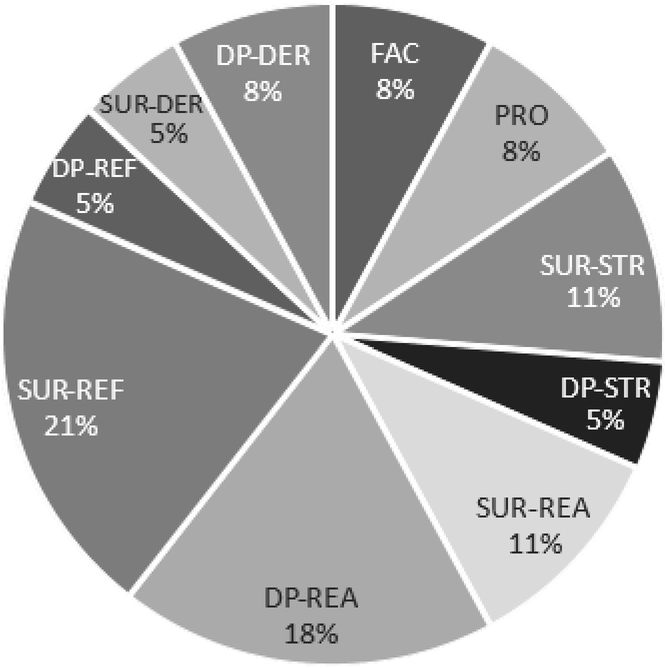


Figure 2.5. Percentages for each category when AfL techniques are employed

However, the increase in the range of question type emerged to be an unintentional by-product of using formative techniques. In the interviews, the teachers stated that most of the AfL techniques they used in lessons were intended to increase learner participation and monitor whether learners were on or off task. This was particularly the case with the 'no hands up' technique:

> **Teacher W:** I ask the least likely to understand to assess the ability (sic) of the class. I use the [questioning] techniques to inform participation rather than deep thinking.
> **Teacher X:** I randomly choose people to answer, or ones who look like they're not listening!
> **Teacher Y:** I do both [hands up and no hands] to make sure I get feedback from everyone.
> **Teacher Z:** I do make a conscious effort for no hands up, but it can slip when Year 7 are excited!

The findings presented in Figure 2.5 suggest that using formative questioning techniques has a positive effect on the range of the type of questioning and the depth of intended mathematical thought. Furthermore, the literature review showed that the 'no hands up' strategy gives the learner thinking time (Black et al., 2003) also allowing for deeper thinking, which Teacher X subscribes to:

> **Teacher X:** I use no hands up every lesson and get students to discuss on tables...it gives them time to think about it and I stand back and love to watch the animated discussion...almost arguing over it!

However, in the observed lessons, 'no hands up' questioning tended to be quick-fire, going against the importance of the thinking time required by Black et al. (2003).

As discussed in the literature review in Chapter 1, 'show me' questions lend themselves very well to the use of mini-whiteboards which in turn allows for formative feedback (Watson & Mason, 1998). However in the teacher interviews, there emerged a barrier in the school to teachers being able to take the risk of using mini-whiteboards both with older students, as they thought the learners may find it a childish resource, and with students with challenging behaviour, due to the fact that the lesson may become harder to manage. The barriers teachers face in using AfL questioning techniques, in particular mini-whiteboards, is investigated further in the main study.

Reliability and Validity of the Pilot Study

A number of steps were taken to reduce bias and ensure the reliability and validity of the pilot study. For the questionnaire the response rate was analysed and, due to the method of conducting the questionnaires, this was found to be particularly high which gives some reassurance of the reliability of the data collected as such a high proportion of the learner voices was heard. However, a test-retest of the questionnaire was also carried out with one class which showed that some of the questions were answered significantly differently on the two occasions which impacts the validity of the learner voices in this research. Consequently only questions with a correlation higher than 0.65 from the test-retest have been used in the findings of this pilot study.

As the sole researcher on this study, I both observed and interviewed the teachers involved as well as analysing and drawing conclusions from the data. This limitation of the study could lead to bias (BERA, 2011), and could have an effect on the validity and objectivity of the data collected and the conclusions drawn. To minimise this threat, inter-observer reliability was tested with another teacher coding one of the lessons to check the percentage agreement between the two observers (Robson, 2002). As discussed above, this resulted in the factual and procedural categories only being coded as surface level due to ambiguity in the categories of deep procedural and derivational or structural thinking. Once the deep factual and deep procedural categories were eliminated, the inter-observer reliability increased from 79.6% to 85.7%. The percentage agreement between surface and deeper questioning, at 98.0%, was more reliable than the agreement for question type. This suggests that depth of questioning is easier to categorise than type of questioning.

A test of the internal validity of the results could have been achieved by observing a teacher again after a couple of weeks and comparing the statistical analyses. This however was not possible as the teachers had all discussed the questioning categorisations in the interviews following the lessons, so any differences could have been attributable to adapting their questioning in line with these categorisations as opposed to exposing weaknesses in the replicability of the data and hence the validity of the study.

Conclusions from the Pilot Study

In the pilot schools, surface questioning in mathematics significantly outnumbers a deeper approach; however there is a variation in the proportion for different teachers. While the teachers attribute a variation in the proportion of surface level questions to the level of attainment of the class and the topic or concept they are learning, the findings from the pilot study suggest that it is in fact the teachers themselves who influence the depth and type of questioning employed.

The variety of question type increases and the depth of questioning is more prevalent when formative techniques are employed, even if that was not the teacher's motivation for using AfL techniques. A more in-depth, qualitative analysis of whether this relationship between the use of formative assessment and the type and depth of questioning holds will be investigated in the main study.

The main action research element of this thesis aims to explore whether working with teachers on the new IMPaCT Taxonomy has the potential to increase the proportion of deeper approaches to questioning and to broaden the types of question they pose in lessons. The pilot study gives some anecdotal evidence related to this question from the inter-reliability tester who said observing and coding the lesson using the early form of the taxonomy made her consider her own questioning techniques in the classroom.

Implications for the Main Study

The inter-observer reliability testing suggests that the categorisation of questioning was consistent, however while the results are considered reliable for the lessons observed, it was not possible to assess the internal validity so there is no observational evidence that the results are repeatable for other topics or classes (Stake, 2005). In the main study therefore it is essential that more than one lesson from each teacher is observed before interviews take place.

Although the interviews added to the validity of the pilot study by giving evidence from the teachers' points of view of how typical these lessons are, these interviews were not recorded and only notes were taken. This means that other information may have been missed. This data could have been made more reliable if time had been available to record the interviews and analyse the transcripts (Denscombe, 2007). Although transcription is time consuming, it will be necessary to the reliability of the main study to record and analyse the teacher interviews in this way.

The major weakness to the pilot study was both the reliability and the validity of the learner questionnaires (Denton, 2013a). Despite carrying out an initial test of the questionnaires, some questions were still not answered consistently by the participant learners. As this suggests that not all of the learners' responses are a fair representation, learner questionnaires will not be used in the main part of the study in order to devote more time to coding and analysing lessons.

In the next chapter I will explain how the results of the pilot influenced the revised version of the IMPaCT Taxonomy.

3. The IMPaCT Taxonomy

This chapter describes how the literature review on the emergent perspective and classifying questioning, along with the results of the pilot study, were combined to form the Intended Mathematical Processes and Cognitive Thought (IMPaCT) Taxonomy. A visual representation of the IMPaCT Taxonomy is presented, which was designed to support teachers to understand how their questioning affects the type and depth thinking of their learners. Finally, this chapter describes how the IMPaCT Taxonomy is intended to support teachers in establishing the sociomathematical norms of *mathematical efficiency*, *sophistication and elegance*, developing an understanding of *mathematical difference,* and understanding what constitutes an acceptable *mathematical explanation* and *justification* in their classrooms.

The literature review established that it is often more helpful to distinguish between a surface and deep approach to questioning in mathematics (Smith et al, 1996; Marton & Saljo, 1976a; 1976b) rather than the more simplistic notion of open or closed questions (Watson, 2003). The IMPaCT Taxonomy, therefore, determines whether questions are higher-order or lower-order, by considering whether or not they require learners to take a surface or deeper approach to their mathematical thinking. However, in the IMPaCT Taxonomy, this is considered in terms of what mathematical thinking the teacher intended, as Watson (2007) argues that what a teacher intends and what a learner perceives are not necessarily consistent.

From the literature review it was found that, from the emergent perspective, to establish sociomathematical norms in the classroom, the questioning needed to be specific to mathematics (Yackel & Cobb, 1996) in order to elicit mathematical thinking. Therefore, Bloom's taxonomy was found to be helpful in establishing social norms, but not necessarily sociomathematical norms. As a result, the IMPaCT Taxonomy draws more on the research of Andrews et al.

(2005) and their seven mathematical foci: *conceptual; derivational; structural; procedural; efficiency; problem-solving;* and *reasoning*. However the IMPaCT Taxonomy differs from these foci in three main ways:

1. The *conceptual* focus was removed, following a process-object and emergent perspective stance that conceptual understanding can be developed through both procedural and structural thinking.
2. *Efficiency* has been replaced with *reflective* to allow this category to also include the sociomathematical norms of *mathematical difference* and *mathematical sophistication*. The reflective category therefore covers the teachers' questions which encourage learners to evaluate mathematical efficiency, difference and sophistication through considering alternative approaches to tackling a mathematical problem and critically examining their effectiveness in the given situation.
3. *Problem-Solving* has been removed as all the foci could be considered within a problem-solving environment, causing potential for overlap in coding the categories. Problem-solving will be considered, but more in relation to whether an inquiry-based lesson has an impact on the nature and depth of teachers' questions.

As discussed in Chapter 2, in the pilot of the original taxonomy (Denton 2013a), it was decided that factual and procedural questions could always be considered surface level as, despite the level of complexity of the question, if the learner has simply been taught a procedure to follow then deeper thinking is not required. Following a more in-depth literature review, it was decided to investigate Fan's and Bokhove's (2014) idea that it is how procedures are used and discussed which contributes to higher-order mathematical thinking. If procedures were used and discussed in this way by the teachers, then, in the IMPaCT Taxonomy, these questions fall into either the structural thinking category, for why an algorithm works, or the derivational thinking category, for how a procedure could be adapted for new situations. This begged the question: could all the other categories be considered at both a surface level and deeper level?

Derivation, by the very definition of the word, requires the learner to adapt their prior knowledge and apply what they have learned to new and unfamiliar situations, which requires learners to think beyond the surface. In addition, if a teacher's question considers the structure of the mathematics, it is requiring conceptual thought whatever the level, so will always encourage a deeper approach to thinking. The prompts adapted from Watson's (2007) analytical instrument for the structural classification in the early version of the IMPaCT Taxonomy (see Appendix 1), for example analyse, generalise, classify, compare, also suggest that only higher-order thinking is required to think about a problem structurally. If a teacher asks a structural question which is known to the learners then it could be classified a factual recall question.

The refined IMPaCT Taxonomy (Figure 3.1) classifies questioning in mathematics into the original six categories of: factual; procedural; structural; reasoning; reflective; and derivational. However, as can be seen in the Venn diagram representation of the taxonomy, I deemed that only the reasoning and reflective categories can take both a surface and deeper approach to learners' thinking.

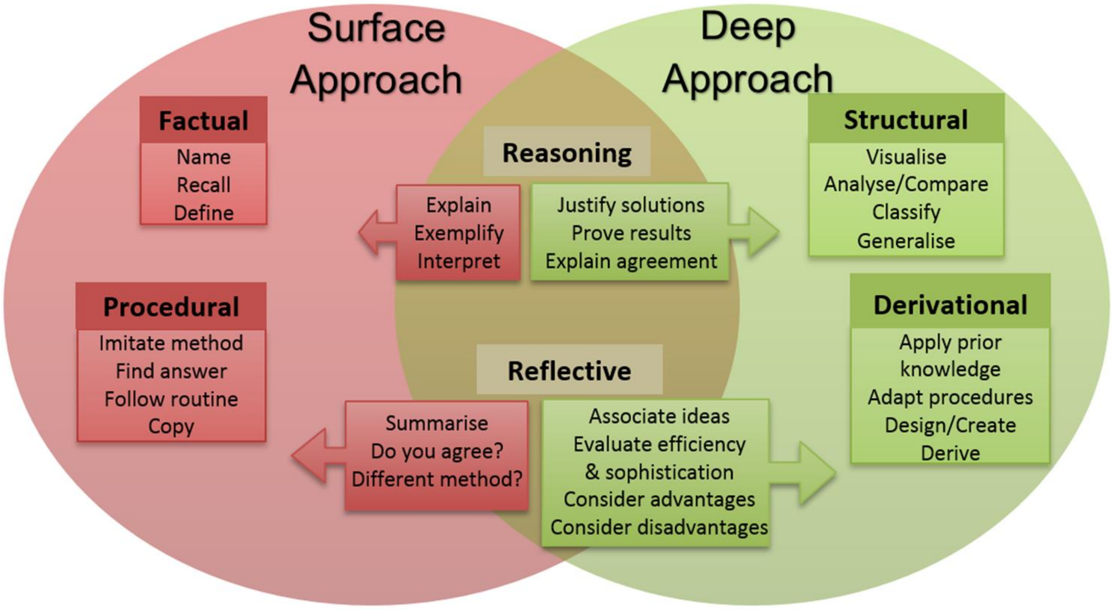


Figure 3.1. The IMPaCT Taxonomy

The categories in the IMPaCT Taxonomy do not form a hierarchy as such on their own, as the taxonomy considers the depth of the intended mathematical thinking in addition to the type of question, however factual and procedural questions can only be classified as surface level, and structural and derivational can only be classified as deeper level. The reflective and reasoning categories could be tackled at a surface or deeper level, which follows the findings of Yackel & Cobb (1996) from the literature review. For example with reasoning, a learner may have been asked to simply explain what they did in terms of following a procedure which would be considered surface level, whereas if they were asked to justify or prove their answer then a deeper level of thought would be required to reason in terms of the structure of the mathematics.

This also links in with the distinction between social norms and sociomathematical norms; a social norm may be established in the classroom which encourages learners to explain their thinking which would be considered surface level as outlined above, however the sociomathematical norms of encouraging learners to evaluate the mathematical efficiency, difference or sophistication of their approach, and indeed to compare those to alternative approaches offered by their peers or the teacher, could be considered a deeper level approach to cognitive thought.

In the pilot, it was only the researcher and the inter-observer who had to work with the taxonomy, however in subsequent action research, I decided that the participant teachers needed to engage with the IMPaCT Taxonomy, so the updated visual representation of the taxonomy in the form of a Venn diagram (Figure 3.1) was designed to support teachers in classifying their questioning into both surface or deeper level as well as considering the six question type classification categories. The ultimate aim of the IMPaCT Taxonomy was to support teachers to understand how their questioning supports their learners' mathematical thinking.

4. Research Design

This chapter outlines the methodology, the qualitative and quantitative research methods, the sampling methods, and the action research strategy used in the empirical research for the main study of this thesis. These are presented in conjunction with the literature used to inform the decisions taken. The limitations of each of the research methods and their potential threats to the validity and reliability of the results are then considered. Finally the ethical issues in relation to this research are discussed, along with how this study follows ethical guidelines for conducting research in an educational setting.

Research Questions

As stated in the introduction, this thesis aims to answer the following questions:

1. **What factors affect the type and depth of questioning used by mathematics teachers?**
2. **Does working with the IMPaCT Taxonomy affect the type and depth of questioning used in mathematics lessons?**
3. **Does the IMPaCT Taxonomy affect mathematics teachers' understanding of how their questioning impacts on their learners' mathematical thinking?**

These questions are all concerned with the connection between two or more variables (Kerlinger, 1970), however there are both quantitative and qualitative methods for testing these relationships (Cohen et al., 2007).

Research Methodology

> Theory is seen to grow out of practice and to feed back to inform and guide practice (Cobb & Yackel, 1996, pp. 175-176).

As a keen researcher and a practising teacher, Cobb's and Yackel's (1996) quote resonates with me particularly. Pedagogical research should not simply be carried out away from the classroom as, through observing practice, the theory of good pedagogy can emerge, the implications of which can then be used to shape future teaching practice. To this extent "the relation between theory and practice is reflexive" (ibid, p. 175). If one subscribes to this notion of reflexivity between teaching pedagogy and practice, then it stands to reason that the research methodology would employ both inductive reasoning, for theory generation, and deductive reasoning, for theory verification.

There exists a wealth of research on combined or mixed methods, as it is considered rigorous to combine methods in educational research (e.g. Hartas, 2010; Robson, 2002; Maxwell & Loomis, 2003; Cohen et al., 2007). However when it comes to methodology, there seems an expectation to choose between an inductive or deductive approach depending on whether you adhere to a positivist or an interpretist paradigm respectively (Denton, 2015). Gorard and Taylor (2004) argue that the methodology employed in a research project needs to be appropriate to the research questions being explored. Furthermore, Robson (2002) contends that qualitative research can verify theory just as quantitative research can generate theory, which raises the following question:

> [W]hen social inquirers mix methods, are they also mixing philosophical assumptions, and should they? (Greene, 2005, p.275).

Taking a pragmatist or situationist stance, Hartas (2010) states that "paradigmatic purity is not a prerequisite to the completion of a research project" (p.278), implying that paradigmatic positions can be brought together (Denton, 2015). This stance is supported by Greene (2005) who states that "flexibility, creativity, resourcefulness – rather than a priori methodological elegance – are the hallmarks of good mixed-method design" (p.277). Furthermore, Cohen et al. (2007) write that triangulation through mixing methods may indeed utilise both positivist and interpretive approaches

(Cohen et al., 2007). However, in the early 1980s, when the debate between the merits of qualitative and quantitative approaches was at a peak (Gunasekare, 2015), purists believed that these different paradigms involved assumptions which were incompatible (Smith & Heshusius, 1986).

Lincoln and Guba (see Cohen et al., 2007) contest the notion of theoretical and methodological triangulation as they believe this notion is incoherent in terms of epistemology, as "[n]o two theories, it is argued, will ever yield a sufficiently complete explanation of the phenomenon being researched" (Cohen et al., 2007, p. 144). Maxwell and Loomis (2003), on the other hand, argue that mixed method inquiry is more than simply mixing data collection methods, listing the following key components in making decisions on how to conduct a piece of research: the aims of the inquiry; the conceptual framework; the research questions themselves; the specific methods employed; and the validity of the study.

This thesis combines methods using a fully integrated methodology throughout the research for a holistic mixed methods inquiry. For instance interviews are often regarded as a qualitative research method (Hoepfl, 1997), however by planning a semi-structured interview schedule both deductive and inductive reasoning are possible. This is because the structured questions aim to verify theory, while allowing the interviewee to steer the subsequent discussion can generate new theory (Denton, 2015).

Combined Methods Approach

The juxtaposition of conceptual and methodological approaches in a combined method inquiry is both complex and frequently evolving (Denton, 2015). This is highlighted by Opie (2004) who claims that the "relationship between methodology and methods and knowledge and truth is controversial" (p.21). Although a combined methods approach is used commonly in educational research, the debate over the appropriateness of mixing methods continues (Gunasekare, 2015). However I share the stance of Johnson and Onwuegbuzie (2004), that the key decision on which methods should be employed should be dictated by the research questions. Therefore the following research strategies and methods were used for each research question in this thesis:

1. **What factors affect the type and depth of questioning used by mathematics teachers?**

   The questions teachers asked over a series of lessons were collected through recording and observing those lessons. These questions were analysed according to several factors:

   i. The participant teacher;
   ii. the level of attainment of the class;
   iii. the stage in the lesson;
   iv. questions which were asked in conjunction with one of the following AfL questioning techniques: mini whiteboards; 'no hands up' and 'wait time'; discussion in pairs or small groups before taking feedback (Hodgen & Wiliam, 2006);
   v. the topic being taught.

   The first three factors were investigated quantitatively through statistical analysis of the proportions of type and depth of questioning asked. The final two factors were analysed more qualitatively as outlined later.

2. **Does working with the IMPaCT Taxonomy affect the type and depth of questioning used in mathematics lessons?**

   A small-scale action research strategy was appropriate for this research question because the overall intentions were to improve mathematics teachers' current questioning in my school in my role as a Lead Practitioner of Mathematics, with the ultimate aim of improving learners' experiences in mathematics lessons. To achieve this, training was given to participant teachers on the importance of establishing sociomathematical norms in their classrooms in terms of the impact on learners' mathematical thinking and how using the IMPaCT Taxonomy could help establish these norms. Quantitative data collected from lesson observations at three stages in the action research were compared to establish whether, after having training on working with the IMPaCT Taxonomy, the depth and variety of teachers' questioning used in the classroom increased.

3. **Does the IMPaCT Taxonomy affect mathematics teachers' understanding of how their questioning impacts on their learners' mathematical thinking?**
   After a series of baseline observations, participant teachers were interviewed to ascertain their thoughts and opinions on questioning prior to the intervention. Following the intervention, teachers' opinions were collected to make a judgement on whether the actions had brought about a change in teachers' approaches to questioning. This data collection took the form of post-intervention teacher interviews. Some questions were identical in the pre- and post-intervention interviews in order to compare whether the participant teachers' views had changed as a result of the intervention. It was intended that the qualitative data collected from this research method would also help to triangulate any findings from the first and second research questions.

As demonstrated above, a combined methods approach is relevant due to the type of data collection and analysis required to address each research question (Gorard & Taylor, 2004) as some questions require quantitative statistical analysis, whereas others require a more qualitative analysis of participants' opinions. A combined methods approach is more commonly referred to in research literature as a mixed methods approach, where both quantitative and qualitative methods are used for exploring the relationships between the variables (Kerlinger, 1970). The participant teacher interviews were analysed qualitatively. The questions transcribed from the lesson observations were analysed mainly in quantitative terms to generate summary statistics, however the ensuing discourse between teacher and learner was also analysed qualitatively where possible and where appropriate. I expected that the intended mathematical thinking of a question would not be apparent until the learner's response and the teacher's reaction to that response were observed. This is an example of how qualitative forms of data can give more depth to the quantitative analysis (Creswell et al, 2006).

Sampling Methods

A convenience sample of four teachers was selected with a similar ratio (1:1) of male to female as the whole mathematics faculty (2:3), and a range of teaching experience similar to the range of the whole faculty. Although this does not guarantee representativeness, it does attempt to maximise it. These teachers were also chosen as they were keen to be involved in the study for their own professional development. According to Sapon-Shevin and Schniedewind (1991), this participant teacher buy-in is essential to ensure that the teachers fully participate in the development of the research. The participant teachers' profiles can be seen in Table 4.1.

Teacher	Gender	Age range	No. of years teaching	No. of years at the school	Last lesson observation grade
P	Female	20-29	4	2	Good
Q	Female	40-49	13	5	Outstanding
R	Male	20-29	2	1	Good
S	Male	30-39	7	2	Good

Table 4.1. Profile of the participant teachers in the main study of this research

To attempt to maximise the validity of the study, five classes were chosen; the classes were all from the same year group to eliminate the variable of the age of the learners which was not a factor chosen for investigation. Four of the classes were higher attaining learners, one from each of the teachers, to reduce the variable of attainment when comparing the effect that the teacher has on the type and depth of questioning employed. One of the four teachers was also observed with a lower attaining class, which provided the opportunity for some interesting comparisons of the questions used by the same teacher in each type of class, and provided data for the second factor in the first research question. Unfortunately, it was only possible to use one lower attaining class in this research as the other participant teachers did not feel comfortable being video recorded with a lower attaining class.

Lesson Observations

Three one-hour lesson observations per participant class were carried out from June to October 2015 to estimate the starting point in terms of the current depth and variety of questioning used by the participant teachers. The original intention was to carry out all of the observations before the school summer holidays in order to begin training on the IMPaCT Taxonomy in September, however it was not possible to achieve this in that timescale, so training did not start until after the October half term break.

The fifteen lessons were either observed, in similar manner to the pilot study, or video-recorded using the IRIS Connect system used in the school. This is a portable fish-eye camera which can capture the majority of the classroom visually when positioned in the back corner of the classroom. The teacher wore a microphone which was remotely linked to the recording. All the questions asked by the teachers were transcribed, then coded and the frequencies of the types and depth of questioning were calculated in a similar manner to Orrill's (2013) empirical research of the classification of questioning (see Table 1.4). These were then compared to the findings from the pilot study.

Following developmental work on the IMPaCT Taxonomy with the participant teachers through faculty meetings, which is described in detail in Chapter 6, a further one-hour lesson observation per class was carried out in February 2016 to compare the differences between the frequencies of each type and depth of questioning before and after the intervention. Between the first and second round of observations, the school's subscription to IRIS expired and, due to school budget restrictions, it was not able to be renewed. An alternative recording method was devised using a regular video camera. To enable the teachers' questions to be heard over the general classroom noise throughout the lesson, the teachers wore a microphone linked to a speaker used for hearing impaired learners, with the speaker placed beside the video camera to pick up the questions. On the whole this was successful; there were occasions when the level of classroom discourse was so high that some of the teachers' questions or the learners' responses were inaudible, although this was rare. Nevertheless, the effect of the change in recording method is considered in the analysis and discussion in Chapters 7 and 8. The February observations allowed for the monitoring of any effect the IMPaCT Taxonomy was having on the participant teachers' questioning and allowed time to intervene further if required before a further two lessons were recorded in April and May 2016 at the end of the data collection for this thesis.

The questions collected from each lesson observation were analysed statistically according to the teacher, the level of attainment of the class and according to in which third of the lesson the questions were posed. In theory, this should have been a 20 minute time period, however, due to lesson changeover times, most lessons started approximately 5-10 minutes after the bell, which needed to be taken into account. Thus the length of the lesson from when learners were settled at the start until they started to pack away at the end of the lesson, was divided by three. This particular partitioning to the lesson was chosen to see if there is a difference in the type and depth of questioning employed in starter activities, main learning activities and plenary activities in this school. That is, does the introductory phase of a lesson differ to the middle section of the lesson or the summation phase of the lesson, as found by Kawanaka and Stigler (2000). To test that the differences in proportions of both type and depth of questioning were statistically significant, the z-test was used to test the null hypothesis that any difference could be attributed to chance (Warner, 2016).

Some questions asked by teachers in the lesson observations are of course uncodable in terms of the IMPaCT Taxonomy, therefore only questions relating to mathematical thinking were transcribed and coded. Questions relating to classroom management and organisation were not included in the analysis. Rhetorical questions and statements which were not be phrased as a question but had a clear intent for mathematical thinking were also transcribed and coded using the IMPaCT classifications. Where a question was repeated by the teacher, it was only coded once, unless the response accepted by the teacher gave grounds for the question to be classified differently.

The question type and depth asked in conjunction with the AfL questioning techniques listed above were qualitatively analysed to corroborate the findings from the pilot study and to exemplify how these techniques increase the range and depth of intended mathematical thinking. Occasions where the teacher used follow up questions which inhibited deeper thought, or did not seem to provide enough 'wait time' before providing the learners with an answer, were also highlighted, along with situations where the teacher appeared to be looking for a predetermined response which may have inhibited thinking (Black & Wiliam, 1998; Cobb & Yackel, 1996).

The dialogue from the recorded lessons was also analysed for the ensuing discourse between teacher and learner. This takes into account Hargreaves' (1984) notion of the need to clarify the intended response by analysing the dialogue which follows. This also addresses Yang's (2006) view that questions cannot be classified in isolation from the context of the lesson and its participants.

One danger of learners being aware of being videoed is that the presence of a camera might have stifled their responses (Swann, 1994). Furthermore as the microphone was attached to the teacher, the learners' responses were only audible if the teacher was close enough to the learner speaking at any one time. It was important therefore to remind the teacher to move towards learners when engaging in discourse with them.

There is potential for observational bias in this analysis of discourse from the observed and filmed lesson, as the dialogue is "subjective in interpretation" (Burton et al., 2014, p.118), that is the researcher may focus on comments which support the research questions rather any which negate them, and may not be open to new themes emerging in the data. This potential for bias will be addressed through inter-observer reliability testing which is outlined later in this chapter. The effects of these limitations will need consideration in the interpretation of the discourse and any conclusions drawn (Swann, 1994).

Participant Teacher Interviews

At the start and end of the research, participant teacher interviews were carried out to collect the teachers' views on how they found planning for questioning using the taxonomy to allow me to probe deeper into the participants' thoughts (Silverman, 2006). These interviews were audio recorded and transcribed. Semi structured interview schedules (see Appendices 3 and 4) were devised with key questions and prompts to probe deeper where required (Kajornboon, 2005). By this I sought to ensure that the key issues were explored but that flexibility was maintained (Robson, 2002). This flexibility also has the potential to reveal unexpected themes (Zhang &

Wildemuth, 2009), which could then in turn alter the focus of the research questions.

Transcripts from the teacher interviews were analysed using a simple descriptive coding framework (Miles & Huberman, 1994) for thematic analysis to draw out emerging themes (Delamont, 2002) of teachers' perspectives on questioning types and techniques. This data collection was intended to evidence the impact of the action research intervention in terms of how the IMPaCT Taxonomy develops questioning from the teachers' perspectives.

The interviews were conducted individually with the intention that less experienced staff would not be self-conscious of their knowledge compared to more experienced teachers. A limitation of a one-to-one interview however could be that the interviewee might just say what they think the interviewer wants to hear, that is in offering "social desirable answers" (Hartas, 2010, p. 258). To minimise this threat to validity (Cohen et al., 2007), the participant teachers were assured of the non-judgemental nature of the interviews and that anything they said would remain confidential and they would remain anonymous in the thesis write up (BERA, 2011). Another possible issue with these interviews could be that, because only four teachers were interviewed, the data collected may not reflect the opinions of the whole mathematics faculty. To minimise this, a range of experience and teaching styles was purposively chosen in the sample of participant teachers.

A further negative aspect of interviews is that it is very time consuming to convert the audio recordings made in the interviews into transcript form (Britten, 1995). There is also the issue of potential bias in the analysis and interpretation of the data, even when converted to transcript form, as the researcher may seek out themes which support the research questions and ignore others which perhaps do not support them (Johnson, 1997). To address this, the inter-coder reliability tester was also asked to analyse the codings and the themes drawn from the transcripts to verify that all relevant topics had been identified.

Research Strategy

As a practitioner researcher, my primary interest is in effecting change in my own school. Furthermore, due to the nature of the second and third research questions, an action research study was the most appropriate strategy for this research, with the aim to bring about change in the depth and variety of questioning used in mathematics classrooms. This research is on too small a scale to be considered experimental, as differences between a test group and a control group could be attributable to other factors as opposed to the teachers employing the questioning taxonomy. Furthermore, on ethical grounds, any training given to participant teachers on questioning had to be open for any mathematics teacher in the faculty to attend. Therefore a control group was not used in this empirical research.

As discussed earlier in this chapter, the second and third research questions suit an action research study. Furthermore, the research questions fit with the cyclic approach associated with action research (Coghlan & Brannick, 2005), that is, planning, acting, observing and reflecting. Figure 4.1 demonstrates the action research cycle in relation to this research:

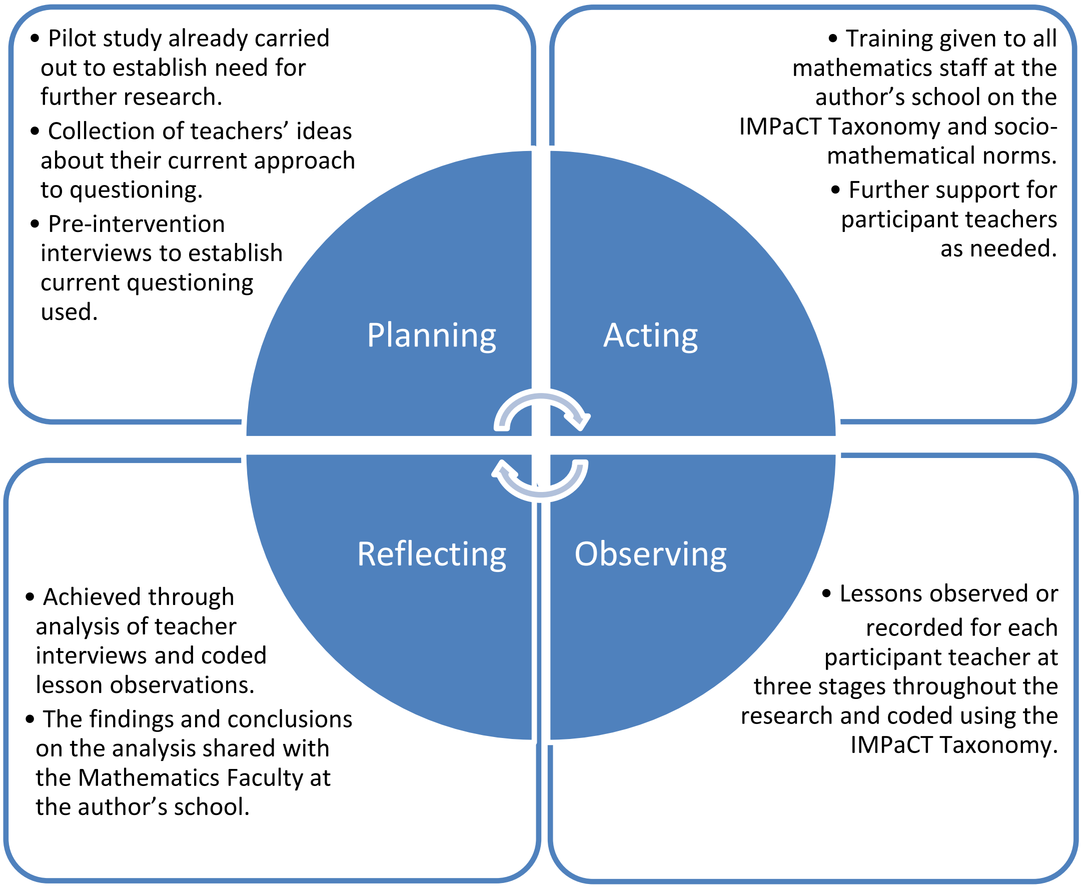


Figure 4.1. Exemplification of how this research follows an action research strategy

However, the cyclic nature of action research can result in multiple cycles. For this thesis, the action research took place over two of these cycles, which are further exemplified below:

Action Research Cycle 1

The planning element, in the first cycle of the action research strategy, was the literature review and pilot study, to ascertain the need for further investigation and intervention in this area. This formed the initial plan for the main study around the refined research questions. Five Year 10 classes were chosen to take part in this research, taught by four different teachers. Three lessons per class were observed or filmed to ensure the results were representative for that class or teacher, followed by pre-intervention teacher interviews. This first cycle of the action research is represented in Figure 4.2.

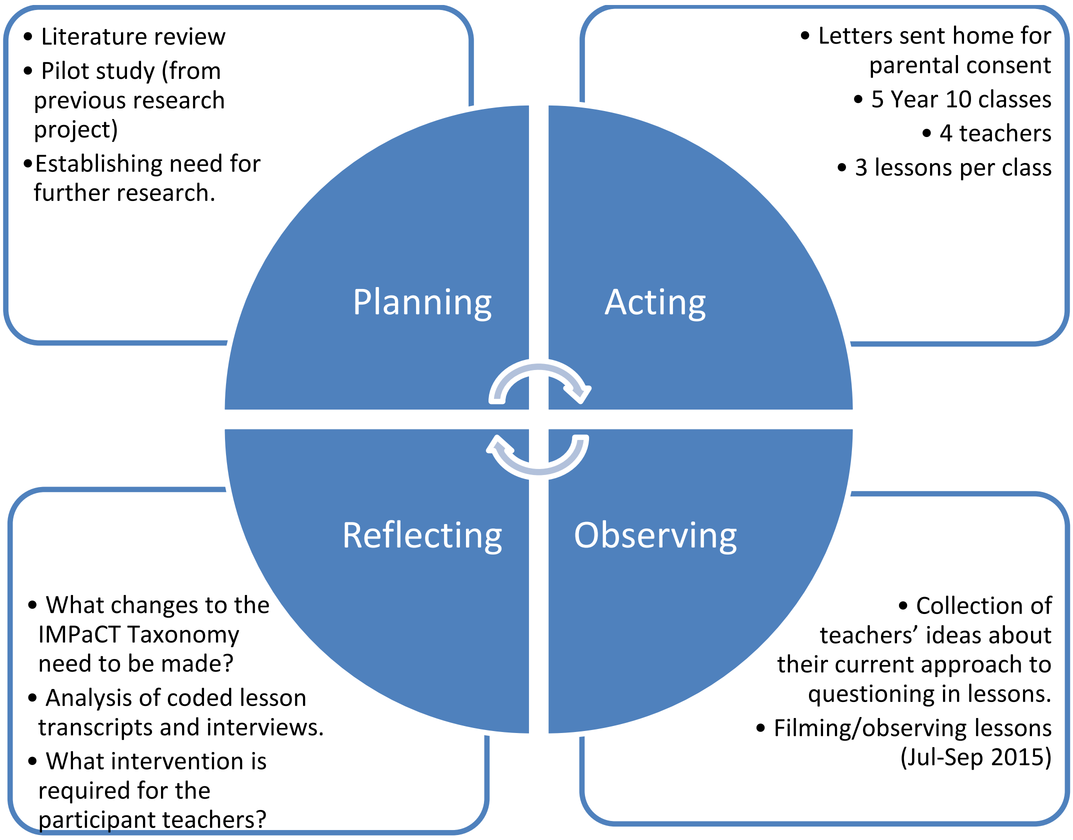


Figure 4.2. Action Research Cycle 1

Action Research Cycle 2

The second action research cycle was based on the results of the first cycle and any new themes that emerged as a result. To this extent, although this research is answering set questions, it is open to inductively take the research in unexpected directions based on the previous findings (Cohen et al., 2007).

Action research methods usually involve some degree of monitoring at all stages within each cycle (Elliot, 1991). This was specifically planned for in the second cycle of this action research in order to monitor the progress in bringing about the desired change. To achieve this, one lesson for each participant class was recorded and coded during the second cycle and was compared to the original findings. By this time the classes were in Year 11.

These interim monitoring observations highlighted where further intervention and training for the participant teachers was required (see Figure 4.3). At the end of the second cycle, two more lessons from four of the classes were observed or recorded, coded and analysed to draw final conclusions for this thesis.

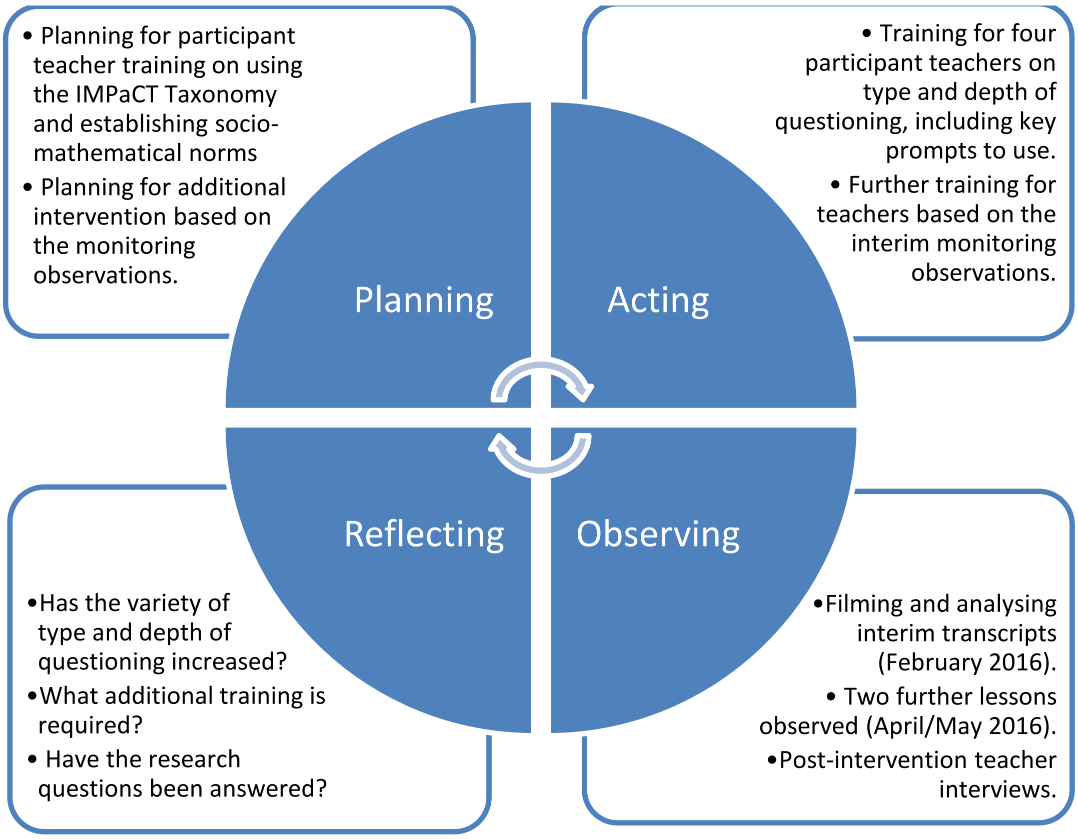


Figure 4.3. Action Research Cycle 2

As this research involves classes over two academic years, observations with these classes were not possible after the second stage due to students going on study leave before the classes' final GCSE examinations. This also impacted on the recording of lessons for Teacher S who had two participant classes and was unable to record the final two lessons in the time available. This unfortunately limits the analysis of the questioning used in the lower attaining class following the intervention with only one lesson on which to base the findings.

Maximising Validity and Reliability

The purpose of mixing methods in research is to optimise the strengths and limit the weaknesses of the research (Lindsay, 2013) through a "reduction of *inappropriate certainty*" (Robson, 2002, p.370). Rossman and Wilson (1985) state three main ways this is achieved

1. **Corroboration**
   This concerns the triangulation of data, where different methods are analysed to test agreement with one another, to give more confidence in the findings. Denzin (1989) refers to this as either triangulation within methods (the stronger the agreement, the more replicable the study) or between methods triangulation, where convergence between two or more methods are analysed (Cohen et al., 2007).
2. **Elaboration**
   This is where one method can be utilised to probe deeper into a phenomenon exposed by another method.
3. **Initiation**
   This considers whether alternative methods can offer a "fresh perspective" (Hartas, 2010, p. 278) on the research.

More rigour can be achieved through triangulation, through enhancing the interpretability of the data collected and supporting of the development of explanations (Robson, 2002).

The combined methods approach used in this action research was therefore designed to triangulate the findings between teacher interviews and researcher observations by ensuring that data were collected "from as widely different perspectives as possible" (Denscombe, 2007, p.135). However due to the unreliability of the data collected from learners in the pilot study, despite changes being made to the learner questionnaire (see Denton, 2013a),the learners' voice was not heard in the main study of the research unless something emerged in the observations which required further exploration with the learner concerned.

To increase reliability, this research also needed to be replicable (Bashir et al., 2008). Therefore the teachers were observed several times to test internal consistency reliability (Cohen et al., 2007). Since three lessons per class were recorded or observed before and after intervention, it was easier to compare whether differences could be attributable to attainment, topic, teacher etc., that is, the themes which emerged from the pilot study. This should have helped to optimise the internal validity which was not possible in the pilot study with only one lesson per teacher to analyse. Furthermore, if the outcomes from the qualitative post-intervention teacher interviews corresponded with the quantitative data from observations, then greater confidence could be had in the validity of the findings (Cohen et al., 2007). The aim of this triangulation, therefore, is to maximise "the repeatability" (Stake, 2005, p.454) and the trustworthiness of the research findings.

To address Yang's (2006) criticism that classifying the level of complexity cannot be done in isolation from the context in which it was posed, I planned that all coding would be made in relation to the class and topic being taught, that is, based on their level of attainment and prior knowledge. Inter-observer reliability was tested by employing another non-participant colleague to code a lesson and compare their results with my analysis. This was done in the same way as the pilot study (Denton, 2013a), with the exception that it was from a recording of a lesson.

Ethical Considerations

All the research in the main body of this thesis was carried out in accordance with BERA (2011) guidelines. Ethical approval was granted by the Centre for Education Studies at the University of Warwick for this research (Appendix 8).

The Head Teacher at the participant school gave consent to conduct the research including, in loco-parentis, consent to video and observe lessons (BERA, 2011). Prior consent of students and their parents, although not necessary according to the school's use of technology policy, was obtained initially as an opt-in letter (Appendix 9), but as the response rate was so low, a second opt-out letter (Appendix 10) was sent home to all learners in the participant classes. This assured parents and learners that individual learners would not be identified in any way in the write up and gave them the option to sit out of the camera view if they did not want to participate in the video recording. As a result, one of the potential participant classes was removed from the research, due to a high proportion of the learners not being comfortable with being filmed.

The purpose of the study, and how observational data and information given in semi-structured interviews would be used, were explained to the participant teachers. The teachers were informed of the voluntary nature of their participation and their right to withdraw at any time (BERA, 2011). They were also assured that any observations would be to capture their questioning and any dialogue between them and their learners and were not intended to grade or judge their teaching in any way. All participant teachers signed a form confirming they had understood this information (Appendix 11). One teacher was not comfortable being filmed, so those lessons were observed in person in the same manner as in the pilot study. This was to ensure the teacher was not put under additional stress as a result of this research. This could have an effect on the findings in the study. To explore the nature of any effect, one lesson with another participant teacher was both observed and recorded. Following this both of the transcripts from the observation and the video recording were coded and analysed to gain a view on how much using both these data collection methods could impact on results.

The confidentiality of the research was explained to the teachers and that to maintain anonymity they would be referred to as Teacher P, Q R and S. Likewise learners would be referred to as Student A etc. from any dialogue following teachers' questions to ensure that individual learners are not identifiable.

5. Findings from Action Research Cycle 1

This chapter presents the findings drawn from the analysis of the results collected from the first cycle of the action research. The lessons in this cycle, which were observed and recorded, are referred to as the baseline observations, as they were used to estimate a starting point for the participant teachers with respect to the type and depth of questioning they employed in their lessons. This chapter presents the results of these baseline observations in relation to the previous findings from the pilot study in Chapter 2, and presents the results of the internal consistency reliability testing and the inter-observer reliability testing as outlined in Chapter 4. The findings from the pre-intervention teacher interviews are also presented to demonstrate the participant teachers' perspectives on questioning prior to any interventions. Finally this chapter discusses the problems encountered in the first cycle of this action research project and the steps which were taken to overcome these difficulties.

The Pre-Intervention Headline Statistics

In the 15 observed or recorded lessons, a total of 1182 questions were transcribed and coded, a mean average of 78.8 questions per lesson. The number of questions coded per lesson ranged from 26 to 124. The teachers' individual variation and mean average of number of questions per lesson differed greatly:

- Teacher P – 26 - 49 questions (mean average 41 questions per lesson)
- Teacher Q – 79 - 112 questions (mean average 94 questions per lesson)
- Teacher R – 61 – 124 questions (mean average 94.7 questions per lesson)
- Teacher S – 57 – 118 questions (mean average 82.2 questions per lesson)

The number of questions posed by Teacher S for his higher and lower attaining classes were similar at 81.3 and 83 respectively. The minimum and maximum number of questions for Teacher S were both with his higher attaining class.

The lower number of questions transcribed and coded for Teacher P could be attributed to the different data collection method used for this teacher, that is, live observations rather than video recorded lessons. However justifications for keeping Teacher P's data in the research are included later in this chapter, following the reliability and validity testing of these two methods of data collection.

The breakdown of the percentages of type and depth of questions coded can be seen in the table in Table 5.1.

Question Type	% S1	% S2	% S3	Total % Surface	% D1	% D2	% D3	Total % Deeper	Total %
Factual	6.4%	9.1%	5.6%	21.2%					21.2%
Procedural	8.6%	10.2%	7.8%	26.6%					26.6%
Reasoning	4.1%	5.8%	2.9%	12.7%	2.5%	2.5%	1.7%	6.8%	19.5%
Reflective	3.9%	5.4%	5.0%	14.3%	1.9%	1.9%	1.3%	5.2%	19.5%
Structural					3.3%	3.5%	3.8%	10.6%	10.6%
Derivational					1.1%	1.1%	0.6%	2.8%	2.8%
Total %	23.0%	30.5%	21.2%	74.7%	8.9%	9.1%	7.4%	25.3%	100.0%

%S1, %S2, %S3 denote the number of surface level questions in each third of lessons as a percentage of the total questions asked. Similarly, %D1, %D2, %D3 denote the number of deeper level questions in each third of lessons.

Table 5.1. Percentages of questions in the baseline observations in each category.

Depth of Questioning

The percentage of deeper level questioning, shown in Figure 5.1, was higher in the main research school than the pilot school (see Figure 2.1), which could be attributable to the higher calibre of teachers involved in the main research in terms of the their lesson observation grading from the respective schools. In the pilot schools, two of the teachers were graded as requiring improvement compared to all good and outstanding in the main research school. There were also some changes to the IMPaCT Taxonomy since the pilot study, as discussed in Chapter 3, which would also have had an effect on the results.

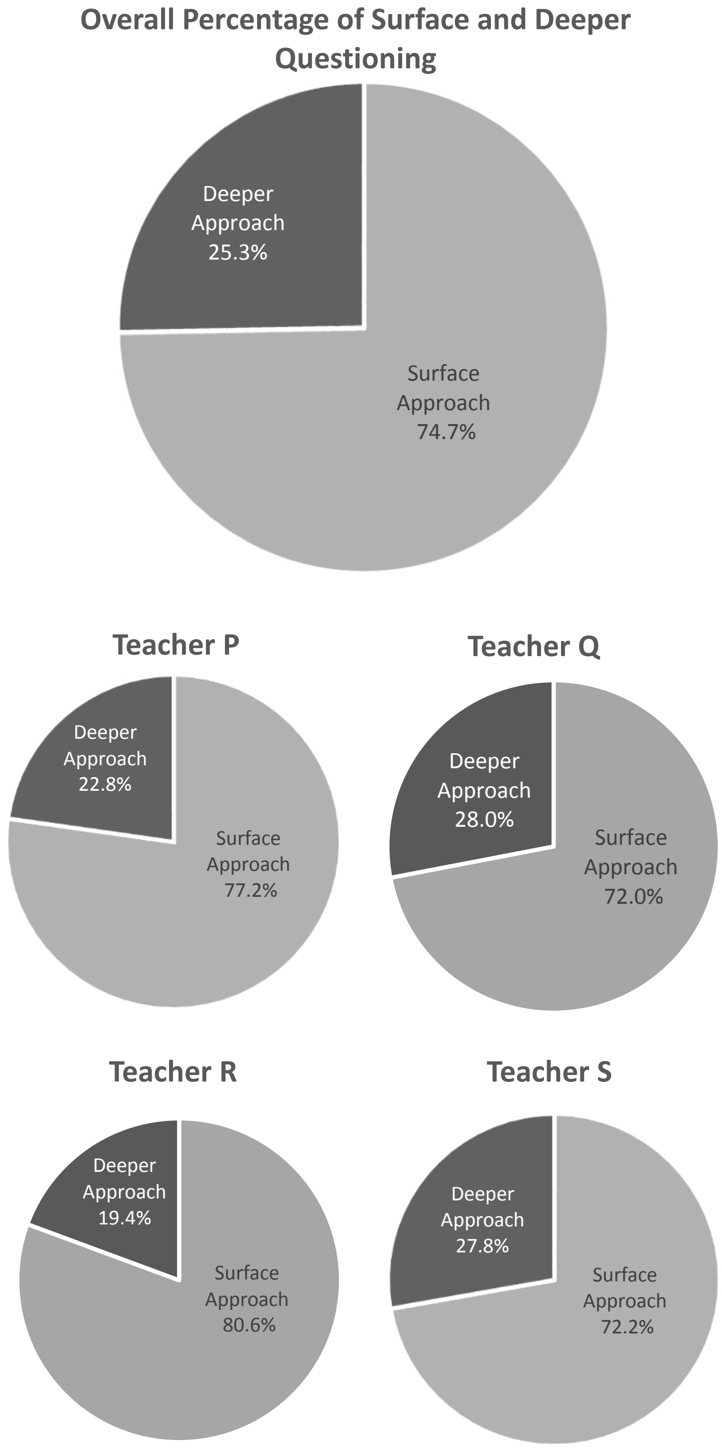


Figure 5.1. Baseline Observations - Percentage of Surface and Deeper Questioning

Although Teacher S had approximately the same number of questions posed on average for both his classes, looking at the breakdown for the two classes

(Figure 5.2), it can be seen that there were approximately 10% more deeper approaches to questioning in the higher attaining class than the lower attaining group.

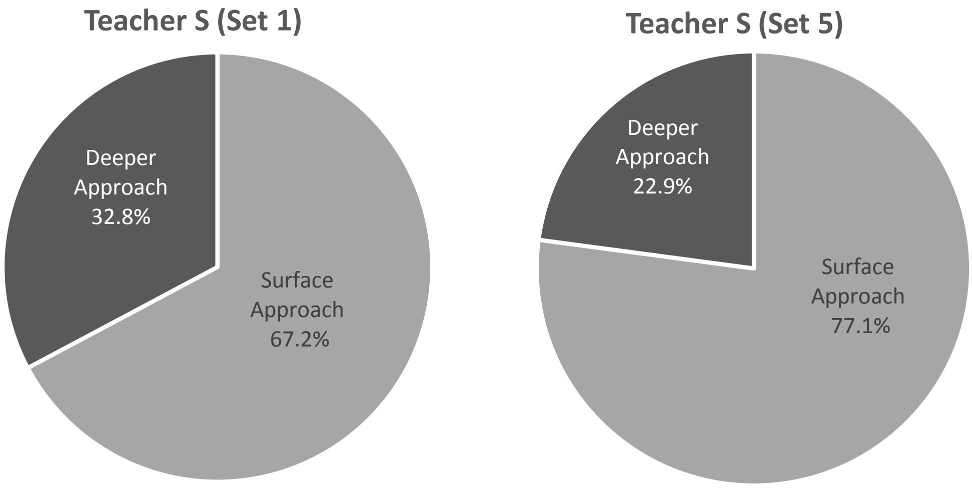


Figure 5.2. Comparison of percentage of surface and deep questioning for each class for Teacher S

Type of Questioning

The most common type of question asked was procedural and the least common question type was derivational, agreeing with the pilot study findings (see Figure 2.3). The number of reasoning and reflective questions were not dissimilar to the number of factual questions asked, however for both the reasoning and the reflective categories, approximately two thirds of the questions asked only required a surface level approach to learners' thinking (see Figure 5.3). There was a substantial difference in the spread of these proportions across the individual teachers. Teachers P and R asked a greater proportion of factual and procedural questions compared to Teachers Q and S. The latter two teachers also asked more reasoning and reflective questions. The fact that Teacher S' statistics are calculated over six lessons, due to the fact that he had two classes in the sample, is taken into account through comparing mean averages and proportions. Figure 5.3 also shows how the proportion of deeper questioning within the reasoning and reflective categories is also greater for Teachers Q and S, with the exception of Teacher P in the reflective category, although this was based on a very small number of questions asked in this category, so the high proportion is unlikely to be statistically significant (Hartas, 2010).

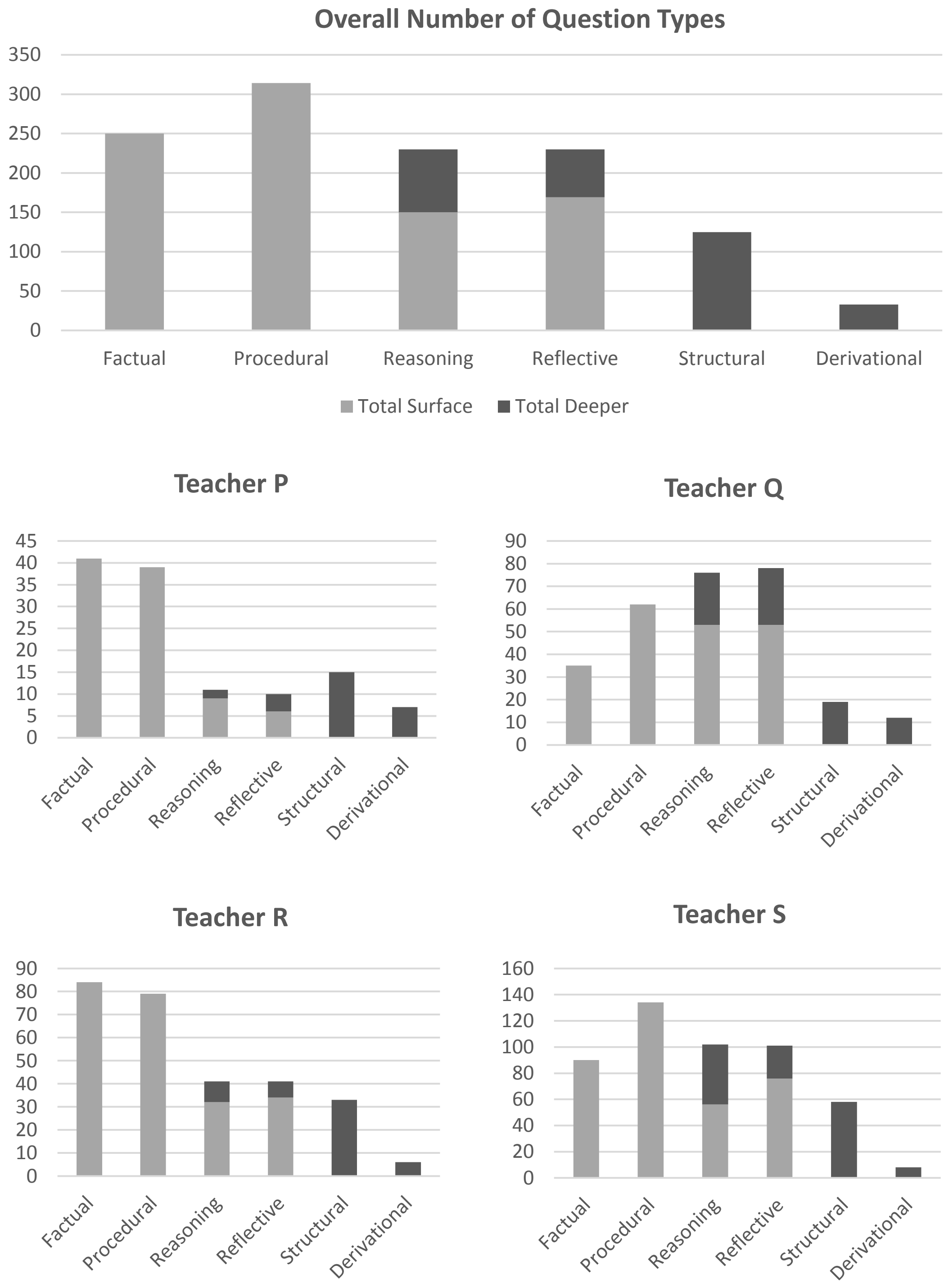


Figure 5.3. Baseline Observations - Number of questions per category

The numbers of questions in each category differed greatly for Teacher S between his two classes (see Figure 5.4). This can be seen particularly in the number of procedural questions being nearly double over the three baseline observations for the lower attaining class compared to the higher attaining class. The number of factual, reflective, structural and derivational questions were all very similar for the two classes. The biggest difference can be seen with the reasoning category, with again nearly double the number, in the higher attaining class, approximately half of which were considered deeper level questions.

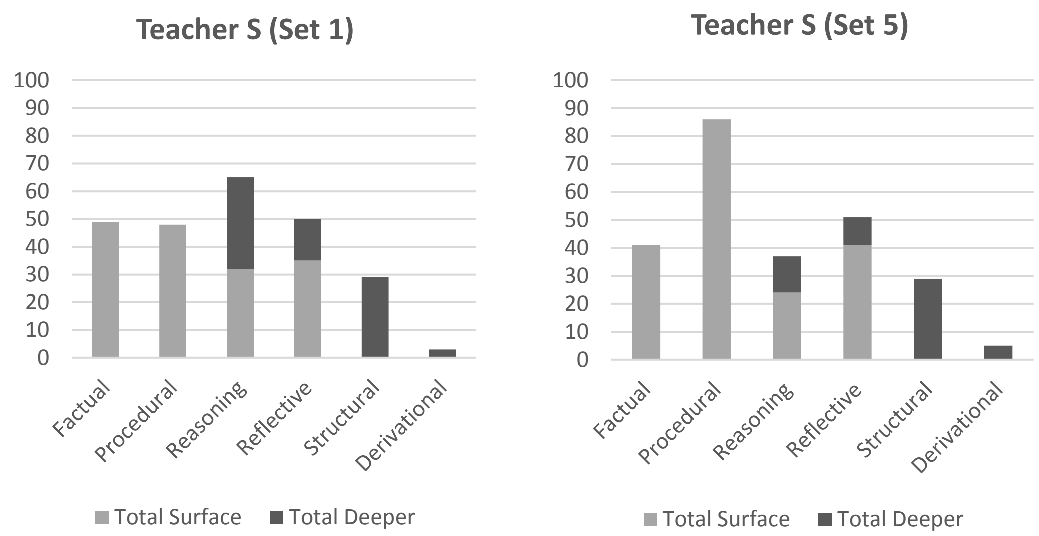


Figure 5.4. Baseline Observations – Comparison of numbers for Teacher S

The pie charts in Figure 5.5, while not showing the percentages of surface and deeper for the reasoning and reflective categories, allow for better comparison of the proportion of question types as they do not take into account the relatively low number of questions asked by Teacher P compared to the other teachers and the fact that Teacher S was tallied over 6 lessons rather than three.

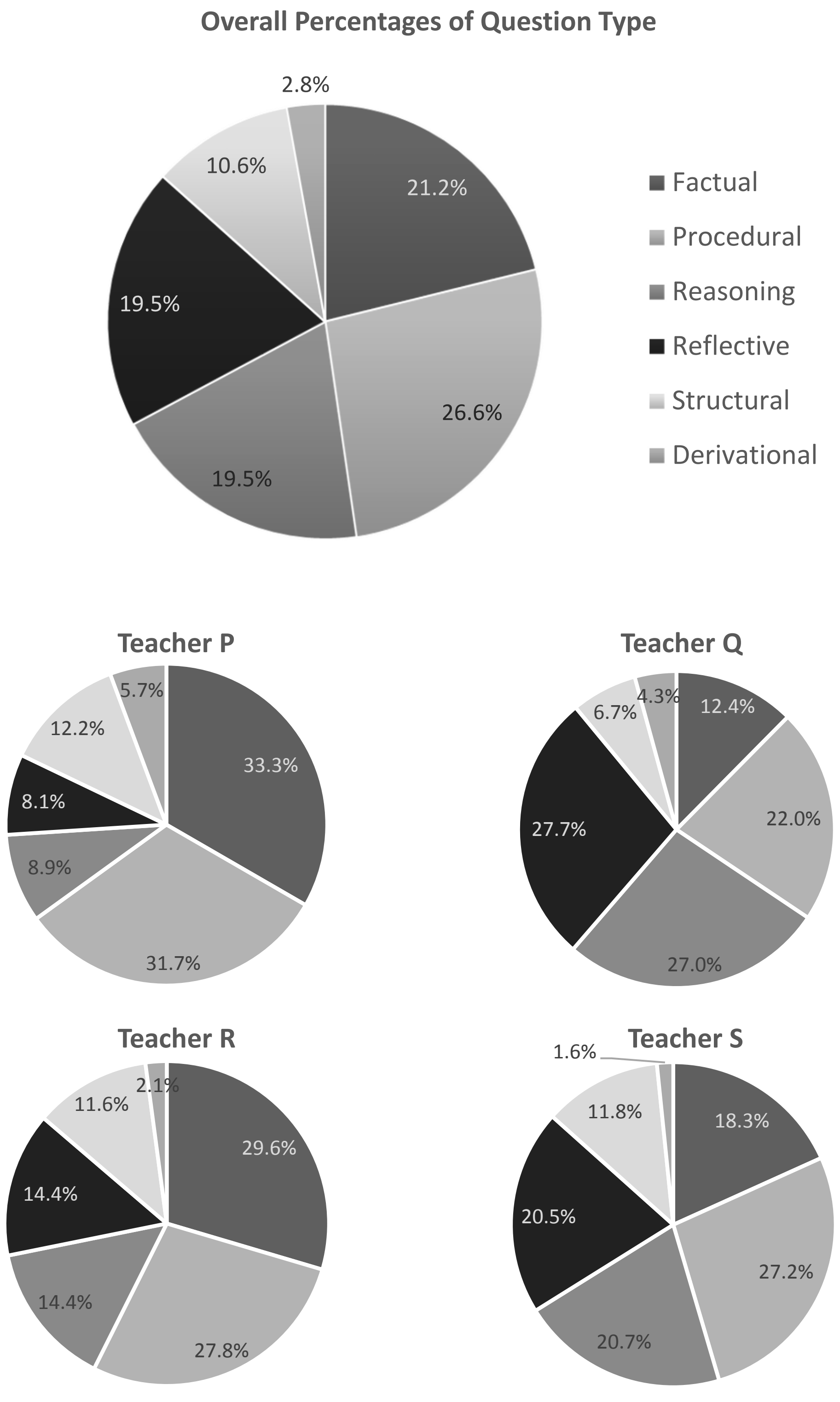


Figure 5.5. Baseline observations – Percentage of type of question asked by each teacher

Overall, factual and procedural questions made up nearly half of the questions asked in the baseline observations, although this was still less than in the pilot study. There were more reasoning and reflective questions asked than in the pilot study. Again due to the profile of the teachers involved in the two studies, these statistics are not particularly surprising. There was a slightly greater proportion of structural questions asked in the pilot study observations compared to the baseline observations in the main study, however the majority of the structural questions in the pilot study were considered surface, which is not a category in the updated IMPaCT Taxonomy, so comparisons of structural questions between the pilot and baseline observations have limited value. Likewise, there were only 2.8% of derivational questions in the baseline observations compared to 5% in the pilot. However it should be noted that this drops to 2.5% when only the deeper derivational questions are considered from the pilot study.

Interestingly, in the baseline observations, Teacher P, who asked the greatest proportion of factual and procedural questions, also asked the highest proportion of derivational questions. Although in terms of numbers of derivational questions, this only equates to 2.3 derivational questions per lesson, compared to 4.0 derivational questions per lesson for Teacher Q.

From the individual charts in Figure 5.5, it can also be seen that Teacher Q had the greatest variety in the types of question posed. Teacher Q also employed more AfL techniques in her lessons, particularly with the use of mini-whiteboards, which could support the findings from the pilot study that using AfL techniques increases the variety of questions asked in lessons. Teacher S had the second greatest variety of type of question posed, however when analysing this separately for the two classes (see Figure 5.6), while the data from the higher attaining class look similar to the spread for Teacher Q, the most experienced of the participant teachers, the lower attaining class' data mirror the spread for the less experienced Teacher P and Teacher R.

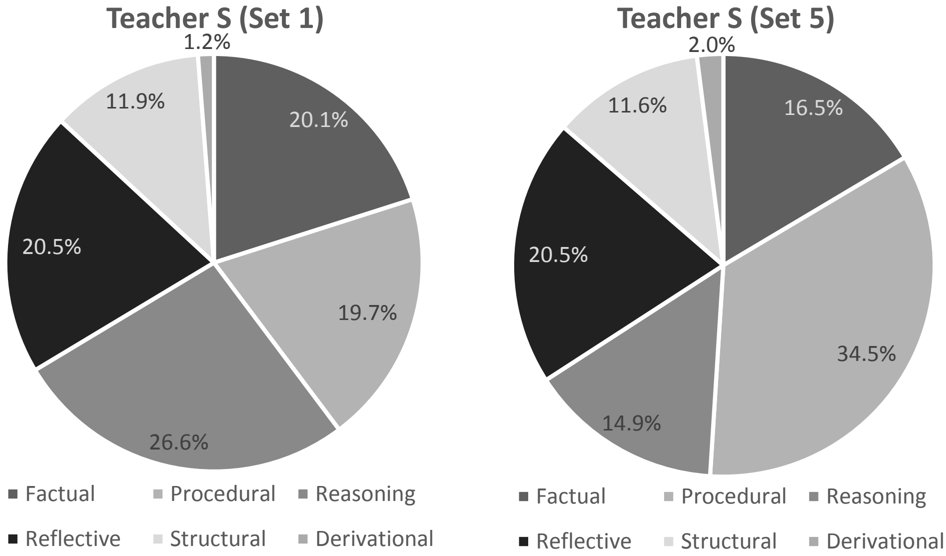


Figure 5.6. Comparison of percentage of question type for Teacher S with two classes

Depth of Questioning by Stage of the Lesson

To address the first research question on how type and depth of questioning is affected by the stage of the lesson, the questions were analysed in relation to which third of the lesson they were posed. In terms of the surface to deeper questioning ratio, in the overall analysis (see Figure 5.7), a similar proportion of deeper questions were posed at the start and end parts of the lesson, with a lower proportion posed in the middle phase of the lesson, however this did vary according to the teacher. Teachers P, Q and R asked the greatest proportion of deeper level questions at the start of each lesson, whereas Teacher S asked the largest proportion of deeper questions at the end of the lesson. No teacher asked the highest proportion of deeper questions in the middle section of the lesson in their combined baseline observations. Teacher Q had the closest proportion of surface to deeper level questioning over the three phases of the lessons.

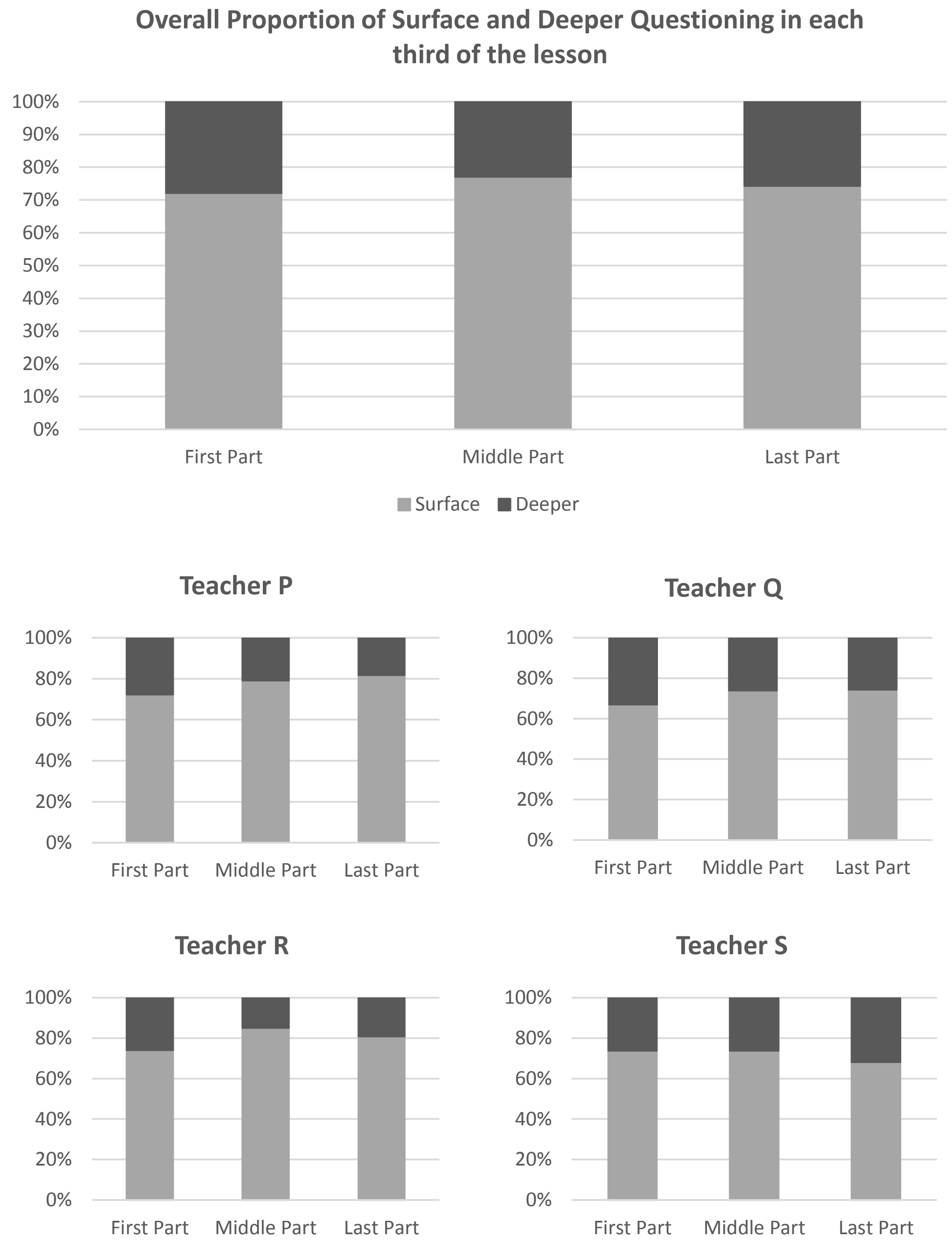


Figure 5.7. Baseline Observations - Proportion of surface and deeper questioning in each third of the lesson

In terms of individual lessons, in seven out of the 15 lessons observed, the highest proportion of deeper level questions was asked in the first part of the lessons. In five lessons, the highest proportion of deeper questions were asked in the last part of the lesson, and in three of the 15 lessons, the highest proportion of deeper questions were asked in the middle section. In terms of the number of deeper questions asked in each phase of the lesson compared to deeper questions asked overall, the largest number occurred in the first part of the lesson in six of the observed lessons. The highest number occurred in the middle section in four lessons and the highest number occurred in the end part of the lesson in four of the lessons. In one of the 15 lessons, there was a joint highest number of deeper questions in the first and middle phase of the lessons.

Teacher S asked a similar proportion of surface level questions in the middle part of the lesson for both his classes, however there was a large difference in the proportion of surface level questions asked in the first part of the lesson, where the proportion was over 20% higher for the lower attaining class, and was the highest proportion in this stage of the lesson out of all of the classes.

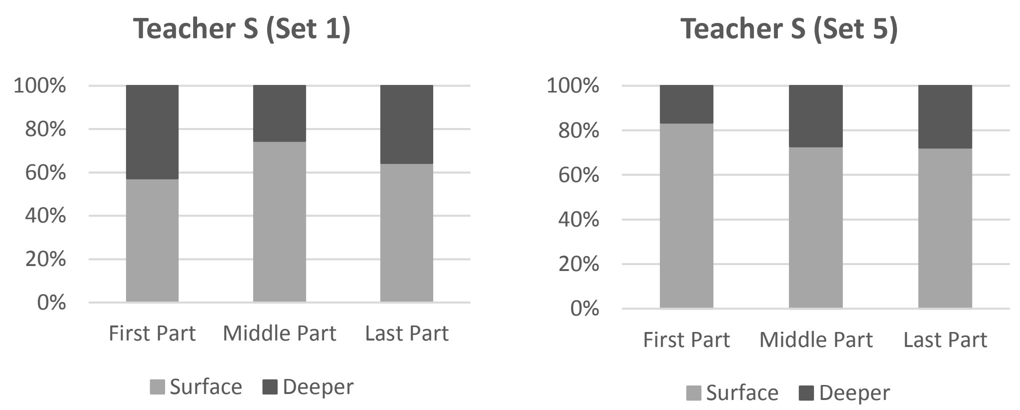


Figure 5.8. Baseline Observations – Proportion of Surface and Deeper Questioning in each third of the lesson for Teacher S with two classes

Type of Questioning by Stage of the Lesson

The proportion of question type was very similar in the first and middle part of each lesson (see Figure 5.9). It differed in the last part of the lesson with respect to the proportion of reflective questioning which was larger in the last part of the lesson compared to the first and middle sections.

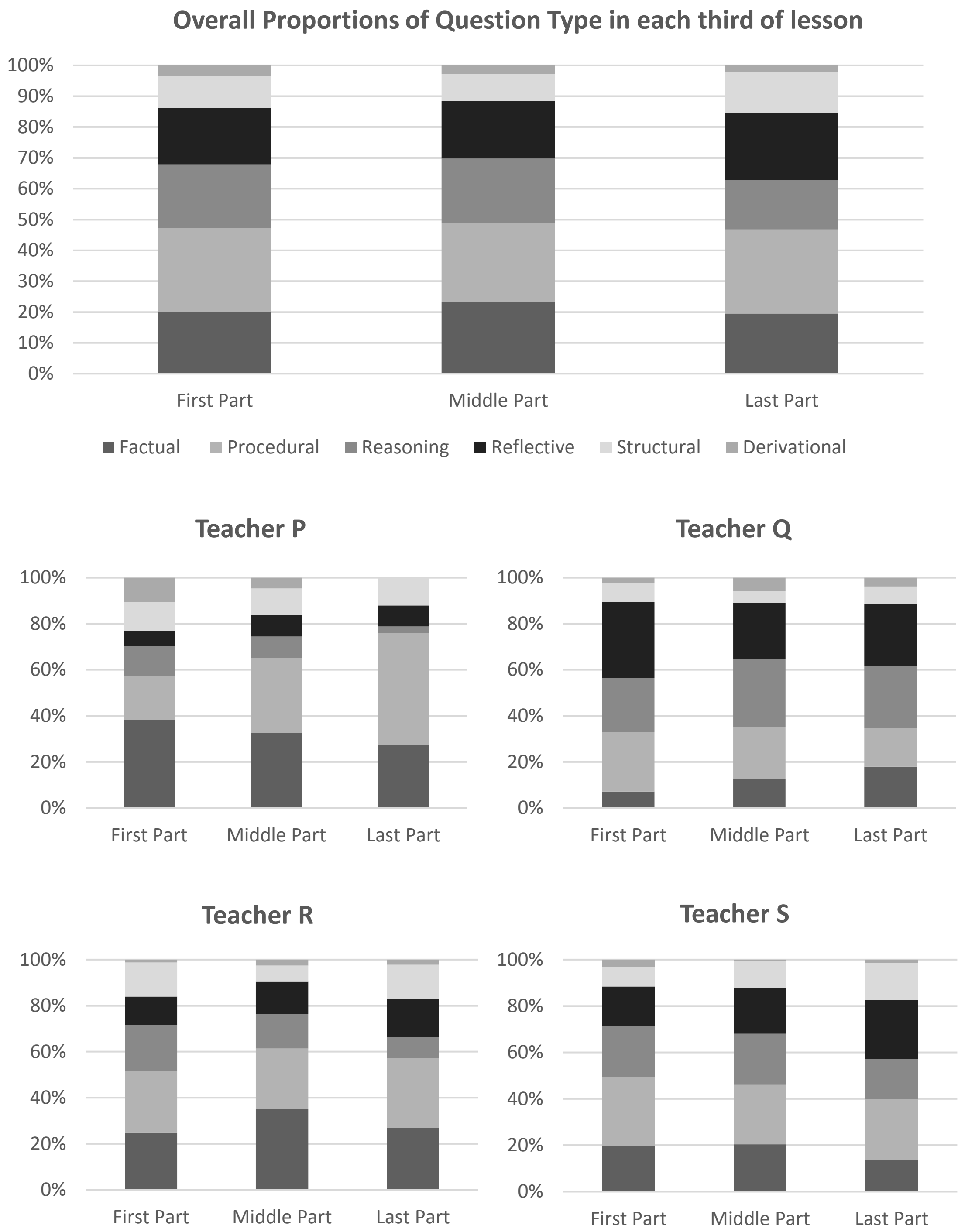


Figure 5.9. Baseline Observations - Proportion of question type in each third of lesson

The spread of question type in each third of the lesson varied greatly per teacher as is also illustrated in Figure 5.9 above. Teacher P's spread of question type varied the most over the three lessons, with the proportion of factual, reasoning and derivational questions decreasing in each third, while reflective questioning increased, and procedural questioning increased substantially. All teachers, apart from Teacher Q, asked the largest proportion of reflective questions in the last part of lessons.

Teacher Q was also different from the other three teachers when it came to reasoning questions, where she asked the highest proportion in the middle section of the lesson, compared to the other three teachers who asked the highest proportion at the start of the lessons. Looking at the breakdown for Teacher S for his two classes (see Figure 5.10), this was also the case with his higher attaining class, where the highest proportion of reasoning questions were posed in the middle section of the lessons, however for his lower attaining class, the increase in reflective questioning over the three lesson phases, and the decrease in reasoning questions over the three phases, more closely follows the patterns of Teacher P and Teacher R, the less experienced teachers.

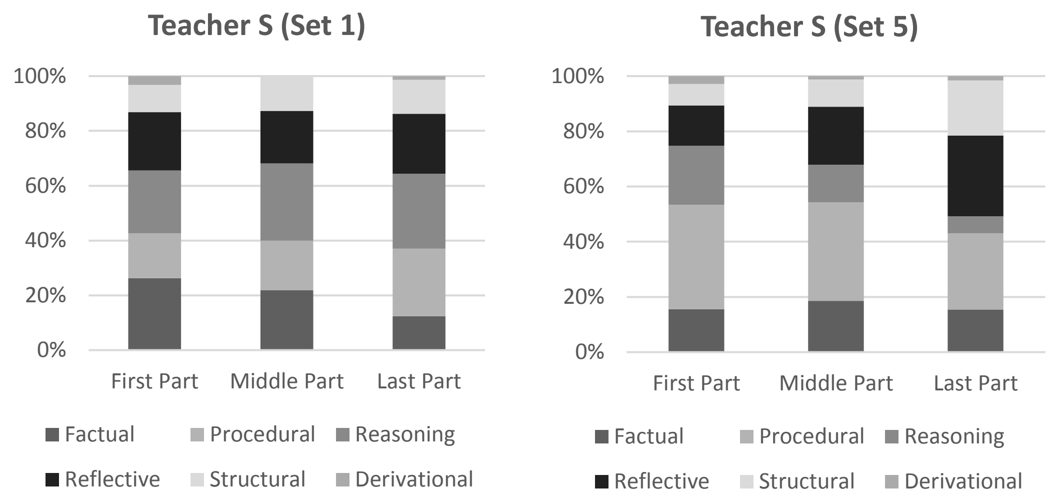


Figure 5.10. Proportion of question type in each lesson third for Teacher S with two classes

Depth of Approach to Reasoning and Reflective Questioning by Stage of the Lesson

Looking in more detail at the proportion of surface and deeper questioning within the categories of reasoning and reflective questions for each third of the lesson, Figure 5.11 illustrates that overall, a larger proportion of deeper reasoning questions were asked at the start and end of lessons compared to the middle section in the baseline observation lessons. This is in contrast to reflective questioning; although reflective questioning increased in proportion over the course of the lessons compared to the other types of question asked, the depth of the reflective thinking intended actually decreased as the lessons progressed.

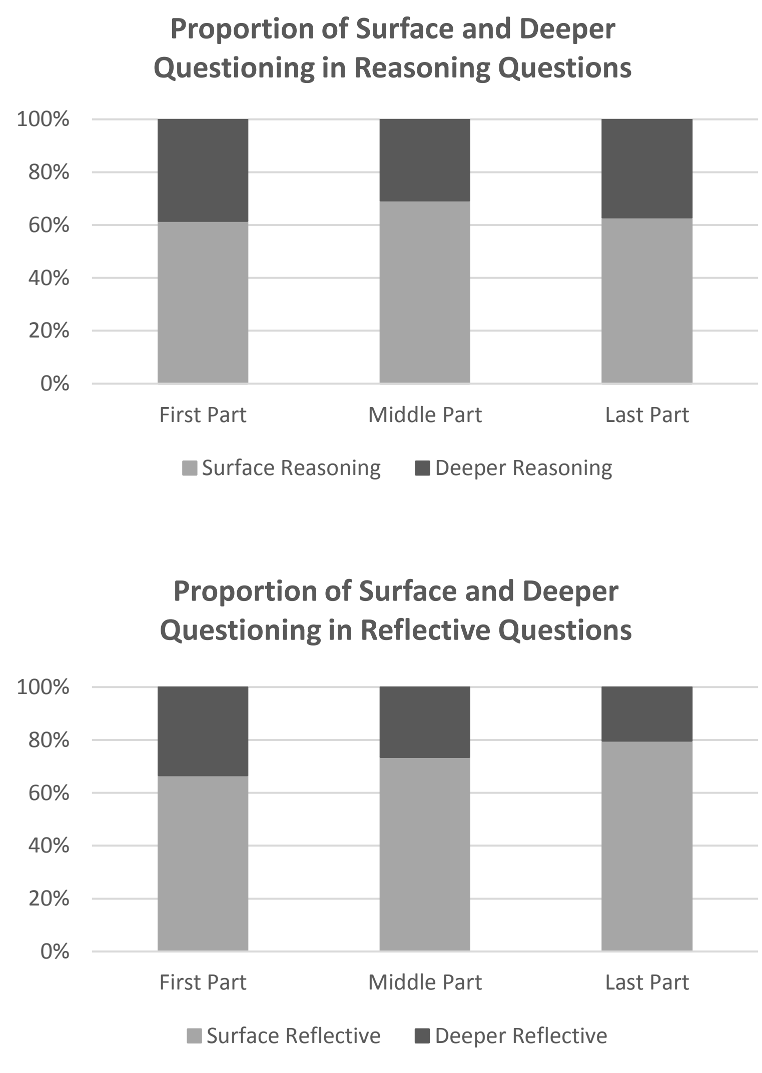


Figure 5.11. Proportion of surface and deeper approaches in reasoning and reflective questioning in each third of the lesson

Evidence of Social Norms and Sociomathematical Norms

There was evidence of social norms being established in the observed lessons. For example, Teacher Q explicitly reminded learners that not only was it acceptable to make mistakes in the mathematics classroom, there were learning opportunities from those mistakes:

> **Teacher Q**: It's ok to make a mistake. There's a saying, if you've never made a mistake, you've never made anything. I've made lots of things obviously!

Teacher R, assured his class in a similar way in one of the pre-intervention recorded lessons:

> **Teacher R**: No question is stupid!

Evidence of sociomathematical norms being established were less common in participant classrooms in the pre-intervention lesson observations. One example was observed with Teacher S with his higher attaining group in his lesson on circle theorems, where he explicitly stressed the need to give reasons from a structural perspective and not simply describe the method employed. This could be interpreted as moving from the social norm of developing explanation to the sociomathematical norm of what constitutes an acceptable mathematical explanation. Interestingly, this was less evident with Teacher S's lower attaining group, when a procedural response to the question "Why is it 8x + 6?" was praised with "Brilliant", despite the question being answered with *how* rather than *why*.

From the qualitative analysis of the lesson observations, one factor limiting the depth and variety of mathematical thinking in the pre-intervention observations was where the teachers interjected quickly with follow-up questions following a potentially deep approach question, or simply answered their own question to move a learner on more quickly. For example, in a lesson where learners were considering how to make a funnel of a set volume with the least amount of material, Teacher Q asked the following list of questions in quick succession with no opportunities for the learners to think in between:

"What are we going to try?"

"We're going to fix the volume to be a litre. What do we need to know about that?"

"Does anyone know how to calculate the volume of a cone?"

"What are we fixing this at?"

"How are we going to measure the dimensions of this? What are we going to use to measure it?"

In another example, Teacher Q followed the potentially deep reflective question of "Looking at that now, do you want to revisit what you said?" by the more factual or procedural question "What is this angle alternate to?" before the learner had time to reconsider.

There was evidence of the other teachers also following potentially deep questioning with procedural follow-up questions before allowing learners the opportunity for deeper thinking.

Teacher S: How are G and H possible? [I.e. what mathematical theorem is required to calculate angles G and H?]

This question was followed immediately by the teacher asking for the value of G, thereby inhibiting the class from the deep reasoning question he originally posed. Opportunities for reflective thinking for learners were potentially lost by Teacher S when he explained why a mathematical problem was easier or more challenging than other problems encountered, as opposed to allowing learners to make the distinction for themselves.

Despite Teacher Q having the highest proportion of deeper level questioning, there were occasions where the discourse became instructional, therefore procedural:

Teacher Q: [Your] slant height is 4.2, do you agree?

Student A: Yes

Teacher Q: The diameter is 6.4 so what is the radius?

Student A: Err…3.2

Teacher Q: So you need to do that squared then add it to that squared then square root it.

However, on other occasions, in the same lesson, the dialogue was more open:

Student B: Is there a way to calculate the height? That would be so convenient!

Teacher Q: Yeah it would! Well figure it out!

Student C: Is the radius that bit [points to radius of the net]...or like this? [Makes the net into a cone]

Teacher Q: Well you tell me!

Another limiting factor was where the teacher appeared to be looking for a pre-determined response to a question posed, for example, this very brief exchange between teacher and learner:

Teacher P: Have a look around, what's wrong with yours?

Student D: Don't know.

Teacher P: It doesn't go through the origin.

In the pre-intervention observations, learner engagement was improved when the teacher did not know the answer to the problem that the learners were trying to solve. This occurred at the end of Teacher S' second recorded lesson, when the highest attaining learners in the class were solving a problem which the teacher had not previously solved himself. Instead of directing the learners to where he knew they should go, his questions were far less leading:

Teacher S: So what do we know? Can we work that one out there?

Student E: Oh yes, you can...145, then that equals 35. Does it?

Teacher S: Where do we go next?

Student E: If that equals 90, then that also must equal 90, if that's 65.

Teacher S: We don't know that yet, can we use that?

Student F: Yeah we can, because that equals 90 as it's a right angle.

Student E: Then 90 take away 35 equals... [writes on board]

Student F: Then you do that [points to board]

Teacher S: What are we missing?

Student F: We can work that one out.

Teacher S: Therefore we can work out this. Can you see that triangle?

[All 3 look and board thinking]

Teacher S: What are we missing girls? Where do we go next? When we spot it will seem dead easy! What's the answer...let's work backwards [brings up answer on the interactive whiteboard]

Student E: But how did they get that?

Student F: Oh, it's angles in a segment!

As the teacher did not seem to immediately know the answer himself, the students had more time to think and play an equal part in the discussion.

Evidence of AfL Questioning Techniques

Using the AfL questioning technique of allowing learners time to discuss in small groups, could have limited the procedural questions which followed a potentially deep reflective and derivational starting question. This technique was employed by Teacher Q where the question “How could you change the problem if someone was having difficulty; how could you make the problem easier or more challenging?” was given to small groups to discuss before feeding back their ideas to the whole class, allowing all learners the time and opportunity to think for themselves rather than being guided in a pre-determined direction.

The use of wait time was observed in some of the pre-intervention lessons, for example:

> [a=5, b=3 c=-2 written on board]
>
> **Teacher S:** If I asked you to work out what is 2a + 4b, what would that be?
>
> [Time given to think individually before teacher takes feedback with no hands up]
>
> **Student G:** It’s 22
>
> **Teacher S:** Student H, do you agree it’s 22?
>
> **Student H:** Yes
>
> **Teacher S:** Why is it 22? Student J?

Unfortunately, Student J was unable to answer, so Teacher S explained himself. Perhaps, given more time to think and ‘bouncing’ the question around a little further, this explanation could have been elicited from the learners.

Another AfL technique employed by Teacher S which allowed for structural thought, was the following list of inequalities presented on the board with 30 seconds to discuss in pairs which were true and which were false:

5 > 3

4 < 2

6 ≤ 6

7 < 10

14 > 7

6 < 8

Again, the time to discuss allowed learners to consider the problems for themselves before discussing with the whole class.

Teacher R employed the use of 'randomly' selecting a student to answer, as a means of encouraging all learners to make the use of wait time to think through the problem:

> **Teacher R:** Please have a look. I don't know who I'm going to ask yet.

This was implicit in other lessons too where Teacher R would ask a question directed at particular individuals, although without the wait time before nominating a learner to answer.

Teacher P was the only teacher to use mini-whiteboards in one of her baseline observations. Of the 12 questions which were asked using the mini-whiteboard, half of them were a deeper level approach. While this is much higher than the 22.8% overall for her three lessons, no conclusions can really be drawn from this due to the small data set.

Avoiding Bias in the Action Research

The overall proportion of questions which required a deeper approach was larger in the baseline observations than in the pilot study. This is not necessarily down to the level of competency of the teachers from each school; in the pilot study the teachers were not aware that the focus of the lessons was their questioning, as the intention was to get a view of the situation without bias. For the main study of this thesis however, it could be considered biased by not informing the participants of the focus of the observations as the purpose here is to determine if working with the IMPaCT Taxonomy has an effect on teachers' questioning. If teachers were not aware of the focus of the observations in the baseline observations in the first cycle of the action research, but were for the second cycle due to the intervention taking place, any difference could be attributable to the fact they were more aware of their questioning rather than the effect of the IMPaCT Taxonomy.

Reliability and Validity Testing

As Teacher P was not comfortable in being video recorded for the purpose of this research, it was agreed that her lessons would be observed in the same manner as in the pilot study. This raised the question of how comparable her statistical analysis would be compared to the other three participant teachers. In the pilot study all lessons were observed in person which allowed for direct comparison, however this was not an option for the main study due to the sheer number of lessons needing observing and the cover implications that would bring. Another option would have been to exclude this teacher from the research, which was decided against as she was keen to be involved and as discussed in the literature review, teacher buy-in is essential in this kind of action research (Sapon-Shevin & Schniedewind, 1991).

Therefore, a comparison needed to be made between the coding of an observed lesson and a recorded lesson to examine the impact that the two methods of lesson observation had on any findings. Teacher Q's third lesson in the first cycle of the action research was both observed and video recorded (see Appendix 12). The difference in the number of questions missed in the live observation was quite substantial, with only 74 questions transcribed in the observed lesson compared to 112 in the recorded lesson. A full breakdown on question types and depth for both methods of data collection can be seen in Table 5.2.

Observed Lesson Analysis

Question Type	Surface Approach	Deeper Approach	Totals
Factual	13	0	13
Procedural	17	0	17
Reasoning	13	2	15
Reflective	12	8	20
Structural	0	2	2
Derivational	0	7	7
Totals	55	19	74

Recorded Lesson Analysis

Question Type	Surface Approach	Deeper Approach	Totals
Factual	16	0	16
Procedural	24	0	24
Reasoning	17	4	21
Reflective	28	10	38
Structural	0	5	5
Derivational	0	8	8
Totals	85	27	112

Table 5.2. Comparison of observed and recorded analysis for the same lesson

The difference in the number of questions coded was a concern as to whether Teacher P's lessons could be analysed alongside the other three participant teachers. Two justifications were made for keeping Teacher P's lesson analyses. Firstly, Teacher P asked significantly fewer questions in lessons than Teacher Q. In her three baseline observation lessons, Teacher P asked 48, 26 and 49 questions. This is a 35.1, 64.9 and 33.8 percentage difference respectively to the 74 questions observed in Teacher Q's lesson. This equates to an average percentage difference over the three lessons of 44.6%. This difference in the quantity of questions asked will have had an impact on the accuracy of the questions transcribed for Teacher P's lessons as, with substantially fewer questions posed by the teacher, it was easier to keep up with transcribing the questions in the lesson, making it less likely that questions and follow up questions were missed.

The second justification for keeping Teacher P's analyses in the research was that, when Teacher Q's observed and recorded lesson transcripts were compared to look at percentages of questions asked from each category, those percentages were in fact quite similar for both types of data collection methods as can be seen in Table 5.3. The biggest difference in percentages was 6.9% in the reflective category, although this category was still the most common classification for both methods.

Observed Lesson Analysis

Question Type	Surface Approach	Deeper Approach	Totals
Factual	17.6%	0.0%	17.6%
Procedural	23.0%	0.0%	23.0%
Reasoning	17.6%	2.7%	20.3%
Reflective	16.2%	10.8%	27.0%
Structural	0.0%	2.7%	2.7%
Derivational	0.0%	9.5%	9.5%
Totals	74.3%	25.7%	100.0%

Recorded Lesson Analysis

Question Type	Surface Approach	Deeper Approach	Totals
Factual	14.3%	0.0%	14.3%
Procedural	21.4%	0.0%	21.4%
Reasoning	15.2%	3.6%	18.8%
Reflective	25.0%	8.9%	33.9%
Structural	0.0%	4.5%	4.5%
Derivational	0.0%	7.1%	7.1%
Totals	75.9%	24.1%	100.0%

Table 5.3. Comparison of percentages observed and recorded analysis for the same lesson

Compared to the large difference in the amount of questions transcribed for the two data collection methods, the percentage of question depth for each was also surprisingly similar. To the nearest five percent, both the observed and the recorded analyses had the same percentage of deep and surface approach to questioning.

To check the reliability of the coding, a lesson was coded twice with a four week gap between the two codings to see if the same observer would code the lesson in the same way on two separate occasions. A four week gap was left between the two codings to ensure that the coding used from the first observation could not be recalled, however to ensure that the questions were coded in relation to the context of the lesson (Yang, 2006), here, the context being the learners and how their responses were accepted by the teacher, the recorded lesson was observed again before the questions were coded for the second time. This testing showed reasonably consistent coding on the two occasions, with a 1.4% difference in the percentage of questions coded deeper and surface level. The question type categories varied a little more. The procedural classifications differed by 2.7%, the reasoning and derivational categories differed by 1.4%, and the factual, reflective and structural categories were coded exactly the same on both occasions, although the percentage of surface and deeper questions within the reflective category differed by 1.4%. Although this shows that I am reasonably consistent in my coding of lessons, perhaps a different coder would not agree with my codings.

The inter-observer reliability testing was conducted on the same lesson as the dual live observation and recorded lesson. The categories in the IMPaCT Taxonomy were explained to the second observer and the background context of the class, including prior knowledge and attainment. The transcribed questions from the recording were given to the observer. The inter-observer then watched the recorded version of the lesson and coded the questions. The percentage agreement of whether the questions were coded as surface or deeper level was 92%, however an exact match of both question type and depth was only 76%. This was less than ideal, however the IMPaCT Taxonomy is designed to support teachers to vary the type and depth of questioning, so as long as that is achieved, the one-to-one agreement of the inter-coder reliability testing is less important than using the inter-observer reliability testing to check that the consistency between the pre- and post-intervention codings are reliable. As discussed in Chapter 4, there is the potential for researcher bias as I am testing the taxonomy which I devised. Therefore, it was decided to use the inter-coder reliability testing to check the other coder found a similar change to me in terms of the percentage of type and depth of questioning pre- and post-intervention. Therefore inter-coder reliability testing is revisited in Chapter 7 to test this potential bias.

Listening to the Teachers' Voice

Four interviews were conducted with the participant teachers which were audio recorded and transcribed. These were then analysed to look for themes in teachers' approaches and their views on questioning in mathematics. A simple descriptive framework was used for this analysis, using the codes listed in Table 5.4 to describe any topic which appeared more than once across the four transcripts. New codes were added to the framework as new topics arose. The fourteen codes used by the end of the analysis were then grouped into five themes, as shown in Table 5.4.

A lack of available time to effectively plan for questioning was a common theme raised amongst all the participant teachers, which was a similar theme in the pilot study in Chapter 2. Most of the teachers felt that their questioning would be improved with more time dedicated to it, suggesting that the teachers understood the impact of their questioning on learners' thinking. However time planning the activities in the lesson took precedence over planning for questioning:

> **Teacher P:** I would like to sort of be better at questioning. Obviously that's to do with time restraints. It's not easy, you don't always get the time to really sit down and think about what questions or what you're going to talk to the pupils about. Most of the time you have to do it quite quickly so you just think about actually giving them the work rather than how are you actually going to deliver it.

Furthermore, a general theme which emerged was that planning for questioning was something that teachers used to do but they believed they now had the skills to develop their questioning as the lesson progressed, rather than needing to plan specifically for questioning. This throws doubt on whether the participant teachers truly understood the benefits to learners' thinking through planning for questioning.

Topic in Interview	Code	Theme
Questioning is not identified on teachers' lesson plans.	LP	Time constraints to plan for questioning effectively
Questioning, if planned for, is done in the teacher's head only.	HEAD	
Proportion of time spent considering questioning when planning a lesson is low.	PROP	
Teachers believe spending more time on planning for questioning would make a difference to learning.	QfL	
Teachers do not classify questions using Bloom's or any other taxonomy.	CLASSIFY	Limited understanding of benefit of plan for questioning
Planning for questioning is something trainee teachers and newly qualified teachers do.	NEW	
Teachers 'go with the flow' with regards to questioning in lessons.	FLOW	
Closed questions are used to assess knowledge and understanding.	CFA	Perceived factors dictating type and depth of questioning
Open questions are used to introduce a new topic.	OFI	
Questioning is dictated by topic.	TOPIC	
Teachers view open questions as requiring higher-order thinking.	OHO	
AfL questioning techniques are used to increase participation.	PART	AfL techniques used for other purposes than developing thinking skills
No hands up is used to assess all learners.	ASSESS	
Teachers like to use mini-whiteboards but there is a shortage of them available with working pens.	LACK	Limited resources to support use of AfL

Table 5.4. Emerging themes from the pre-intervention interviews

Most participant teachers only classified questioning in terms of being open or closed during the interviews, and it emerged that the teachers perceived open questions as a means of introducing a new topic and closed questions as a means of assessment. Teacher R, however, was more critical of the notion of open and closed questions than the other participant teachers and made the connection with higher-order thinking skills:

> **Teacher R:** There's no point in asking an open question that's not really in depth. Well, there is still a point, but it doesn't develop the higher-level thinkers. Whereas a hard closed question could be more appropriate in that circumstance.

Teacher Q had Bloom's Taxonomy displayed in her classroom in the past, and used the taxonomy to classify questions to a certain degree, although not formally and not written on lesson plans. Instead she described preferring to classify her questioning in terms of probing the learners' thinking:

> **Teacher Q:** I get them to work out what they can work out with what the information is and then with the particular task and then I ask them, the question is actually how much more information would they require to be able to answer the question fully. So it's really probing actually what they know.

In line with the interview analysis from the pilot study, the participant teachers in the main study also believed the topic being taught dictated the type and depth of questioning used in lessons. The pilot study suggested that it was the teacher as opposed to the topic that had the biggest impact on the type and depth of questioning. Further analysis of the extent to which topic influences questioning can be found in Chapter 7.

The use of AfL techniques to increase participation was again a theme which arose in both the pilot study and the pre-intervention interviews. Although the participant teachers in the main study cited the use of AfL techniques to be able to assess the whole class, they did not mention their usage as a means of developing or extending learners' thinking.

The final theme that emerged in the teacher interviews was the lack of resources to support AfL questioning techniques, especially with respect to limited access to mini-whiteboards with working pens. In the pilot study, the teachers had expressed concern over using mini-whiteboards with older or more challenging learners, however this was not the case in the main study; the teachers wanted to use mini-whiteboards with all their classes, but with a limited budget, the mathematics faculty were unable to replace the whiteboard pens.

Issues Encountered in Cycle 1

Several issues arose in the first cycle of the action research with regards to using the filming equipment. While IRIS is a very sophisticated recording system, the microphone is only attached to the teacher so not all dialogue in the classroom is captured. While this is not a major problem in terms of only being interested in the dialogue which follows teachers' questioning, it was difficult at times to hear learners' responses from the back of the classroom when the teacher was stood at the front during whole class teaching. There was also the issue in two of the classes with hearing impaired learners for the filming equipment not to interfere with the learner's hearing equipment. To overcome this the teacher met with both students concerned before their respective lessons to ensure it did not impact negatively on them.

In several lessons it was clear that the learners were conscious of the filming taking place to the extent that one teacher felt it necessary to tell the class to 'act normally', however it is difficult to judge if this is any bigger impact than having the researcher in the room to observe (Cohen et al., 2007).

A further issue that emerged was whether the questions posed by the teaching assistant should be taken into account. For logistical reasons more than anything it was decided not to include the questions, the biggest reason being that all her questions were not completely audible because she was not wearing a microphone, as the IRIS recording equipment used in this stage of the research only uses one microphone which had to be worn by the teacher as described above.

6. Training the Teachers on working with the IMPaCT Taxonomy

This chapter outlines the intervention which took place with the participant teachers following the analysis of the baseline assessments and the teacher interviews. The chapter discusses the issues that were learned in the pilot school with respect to training teachers on using the early version of the taxonomy and how these obstacles were overcome in the main research. The methods used to engage the participant teachers with using the IMPaCT Taxonomy to reflect and improve upon their own practice are outlined and links are made back to some of the key literature review findings on the classification of questioning.

Lessons Learned from the Pilot School Training

To disseminate the findings from the pilot study in my previous school, I delivered a mathematics in-service training session to the whole department. One of the activities, which the teachers were asked to do, was to sort a list of some of the questions which I had observed over the four lessons as part of the research for the pilot study (Appendix 13). Most of the members of the mathematics department found this a very challenging task and their classifications were quite different to those I had chosen. One member of staff even commented that he found it easier to apply Bloom's Taxonomy to the questions I had given them than to classify them using the IMPaCT Taxonomy. The teachers worked in three groups on this task and all three groups sorted the questions very differently from both the other groups and from my own coding of these questions. This disparity between my coding and the teachers' coding in the training was surprising, as the inter-observer reliability testing in the pilot study had shown that the result of two different observers watching, transcribing and coding a lesson independently, produced a percentage agreement of 85.7%. In fact, in the inter-observer reliability testing, the percentage agreement for coding surface and deeper approaches to questioning was as high as 98.0%. Following discussion around this issue in the training in the pilot school, it was decided that this discrepancy was not necessarily because the IMPaCT Taxonomy was a less effective means of classifying questions, but that the surface and deep element of the taxonomy meant that it was not possible to classify questions away from the context of the learners and the lesson, agreeing with Yang (2006) that classifications cannot be made in isolation from the context. Perhaps this applies more so with the mathematics specific categorisations in the IMPaCT Taxonomy than the more generic classifications in Bloom's Taxonomy.

Recognising the Importance of Context

There are a number of factors in relation to context which need to be taken into account when classifying questions using the IMPaCT Taxonomy, including the learner's prior knowledge, for example whether a question is procedural because the content has been previously met by the learner or derivational because it has not been encountered before, requiring the learner to independently adapt or apply their prior knowledge. Furthermore, as discussed in the literature review, it is often required to consider the learner's response to a question, and indeed the manner in which a teacher accepts that response, to determine whether the question is higher-order or lower-order. In addition, the presence or absence of follow up questions from the teacher could also help in the classification. For example, a question may in the first instance appear to probe deep structural thinking, but if the teacher is quick to accept a surface level answer from the learner or offers leading follow up questions, then the learner does not have the opportunity for this structural thought. In this scenario, it seems the teacher posing the question is only expecting step by step procedural thinking.

Once again this links back to the findings in the literature review, and Yang's (2006) criticism of Bloom's Taxonomy being ineffective if the questions are classified in isolation from the context. This finding also follows Holster's (2006) and Kawanaka's & Stigler's (2000) notion that in classifying question type, consideration needs to be given to the pedagogy behind the question in relation to the task before it can be decided whether it can be considered higher-order.

For these reasons, it was decided not to train the teachers in the main research through sorting questions away from the context of the lesson from which they were taken. For the purpose of addressing the research questions in the main research, it was not in fact the aim for teachers to necessarily be able to classify individual questions accurately themselves. Rather the ultimate aim for the participant teachers was to develop an understanding of the different types of thinking that could be elicited through varying the type and depth of questioning employed and to come away with the skills and tools to achieve this outcome.

Introducing the IMPaCT Taxonomy

The intervention dimension to this research took place in the *acting* stage of the second action research cycle (see figure 4.3). This first phase of the intervention in the main study followed the baseline observations and pre-intervention interviews. Participant teachers were shown the IMPaCT Taxonomy in the form of the Venn diagram and discussed the role of different question types and how they could be considered to expect surface level or deeper level thinking. Focus was given to the exemplification words for each category (see Figure 3.1) and how this is only intended as guideline to illustrate each classification and is not by any means an exhaustive list of question stems.

From the pre-intervention interviews, it was found that all the participant teachers were familiar with Bloom's Taxonomy, however this was to varying degrees. In this first phase of intervention, the IMPaCT Taxonomy was shown alongside Bloom's Taxonomy and the similarities and differences between the two were discussed, with particular emphasis on the mathematics specific nature of the IMPaCT categorisations as opposed to the non-subject specific classifications in Bloom's Taxonomy. Three out of the four participant teachers expressed the difficulty they had had in the past in using Bloom's Taxonomy to plan for questioning, in particular how it relates to mathematical thinking. Teacher P had particularly strong opinions on how Bloom's Taxonomy was not relevant to mathematical questioning and aired her frustration at being required in her teacher training to reference Bloom's Taxonomy in her lesson plans.

It was discussed that, unlike Bloom's Taxonomy, the IMPaCT Taxonomy was not intended as a hierarchy as such. While it is important to ask questions expecting a deep thought process from the learners, and indeed these could be considered higher-order, surface questions also have an important place in the mathematics classroom, for example in maintaining pace or developing an ability to use algorithms. It was stressed that the intention of the IMPaCT Taxonomy was to see a greater variety in the type and depth of questions posed in the classroom so as to give learners the opportunity for different types of mathematical thinking.

Developing a Climate for Deep and Varied Thinking

The second phase of the training on using the IMPaCT Taxonomy was sharing the key findings from the literature review and the pilot study with the participant teachers. Particular emphasis was given to establishing sociomathematical norms in order to create a climate for learning where deeper and more varied thinking can take place through supporting learners to recognise the importance of mathematical difference, efficiency, elegance and sophistication (Yackel & Cobb, 1996), as well as mathematical explanation and justification (ibid) and a climate where making mistakes is acceptable (Kazemi & Stipek (2001). Participant teachers received a summary of this research for their own reference in future lesson planning (see Appendix 14).

The findings from the pilot study, which suggested that using AfL techniques increased the variety and depth of questioning, were shared with the participant teachers, alongside the list of techniques which were observed in the pilot study (see Appendix 4). The participant teachers were aware of Black's and Wiliam's (1998) research on formative learning and how AfL techniques had been developed from Black's and Wiliam's research, although some of the specific techniques were not familiar to the teachers. Any such terms were clarified with the participant teachers with exemplification of where they could be incorporated into lessons.

Using the Baseline Analysis and Findings

In the final phase of the intervention, before the post-intervention observations took place, each participant teacher was given an individual summary sheet of the analysis of their three baseline observations (see Appendix 15). This included a tabulated breakdown per lesson of the type and depth of questions used in each third of the lesson, plus the total numbers of both types and depth of questions per lesson. The summary sheet also combined the three lessons to produce a table of the total questions over the three lessons broken down into type, depth and stage of the lesson. Finally the combined data was represented in the following charts for a visual representation for each teacher on the type of questioning which is likely to be seen in their classrooms:

- proportion of surface and deeper questions posed;
- proportion of surface and deep questioning in each third of the lessons;

- proportion of surface and deeper approaches to both reasoning and reflective questions, compared for each third of the lessons;
- number of each type of question posed, with differentiation between surface and deeper approaches;
- proportion of type of questions posed.

The participant teachers found this analysis of their own lessons very interesting and were keen to compare their analyses with the other participant teachers. Furthermore, Teacher Q found it reassuring that she was consistent in her approach to teaching both across the three lessons and in each stage of the lesson as shown in her analysis.

Discussion ensued regarding how the topics taught in the three baseline lessons may have affected the outcomes. For instance, two of the three lessons observed for Teacher S with his higher attaining group were on circle theorems, a topic which the participant teachers felt lends itself well to reasoning questioning. The evidence from the pilot study was shared with the participant teachers in the main research that it was the teacher as opposed to the topic which dictated the type and depth of questioning. However the statistics presented as a tabulated break down per lesson was a good way of seeing if the topic did impact the individual teachers' outcomes. This will be explored in more depth in Chapter 8.

In this part of the training, the participant teachers used the IMPaCT Taxonomy to identify their own areas of strength with regards to questioning, and identified the areas in which they needed to develop their questioning in future. Non-participant teachers were invited to this training, however it was up to the participant teachers whether they chose to share their summary statistics with their colleagues.

Participant teachers were asked to consider examples of questions they could have asked in the reasoning and reflective categories to promote deeper thinking on the part of the students. In the topic of circle theorems it was decided that, while simply asking learners to state the theorem they used was considered giving reasoning for their result in terms of an examination question, it did not require deeper thinking on behalf of the learner unless they were asked to justify their reasoning and, perhaps for the higher attaining learners, prove the result.

Although Teacher Q had the highest proportion of deeper questioning, she identified that small changes in the approach to some of her questions could have made the proportion of questions expecting deeper thinking even greater. For example the question “Can you use Pythagoras to find the height [of the cone]” for the participant class was classified as a procedural question, whereas changing this to “How could you find height of the cone?” would require the learner to use derivational thinking as they would be making the decision themselves to apply their prior knowledge in a new context.

Similarly, the participant teachers considered how procedural questions could be adapted to encourage learners to think about the structure of the mathematics, which could then in turn prepare learners for more derivational thinking by equipping them with the skills to adapt procedures to solve more unfamiliar problems. Using the deeper reasoning skills of justification and proof, and deeper reflective questioning to encourage learners to associate ideas, could also pave the way for more derivational thought.

Intended Outcomes of the Intervention

The intended outcome of the training was that the participant teachers increased their awareness of the effect their questioning has on the thinking of their students, and provided the teachers with a practical tool to support them to vary the type and depth of their questioning. The analysis of the post-intervention interviews in the next chapter will provide the main source of evidence as to what extent this outcome was achieved. The second and third rounds of lesson observations in the second action research cycle evidence, whether or not working with the IMPaCT Taxonomy increased the depth and variety of questioning in mathematics lessons.

7. Findings from Action Research Cycle 2

This chapter presents the findings from the second cycle of the action research. Firstly the analysis of the interim lesson observations to monitor the participant teachers' progress on using the IMPaCT Taxonomy are presented and the necessary actions taken as a result are outlined. The overall findings from the second cycle are then presented which combine all the data collected from the post-intervention lesson observations and recordings, that is a combination of the monitoring lessons from February 2016 and the final lesson observations in April and May 2016. This enabled the same number of lessons to be analysed pre- and post-intervention. Where further intervention was required following the monitoring of progress within this action research cycle, an analysis of the final two lessons alone was carried out to assess the impact of the additional intervention. Finally the themes found from the analysis of the post-intervention teacher interviews are presented.

Headline Statistics from the Interim Monitoring Observations

In February 2016, five lessons were either observed live or recorded, that is, one lesson for each participant class. As discussed in Chapter 4, these lessons were recorded with a different method to the pre-intervention recordings. From these five observations, 406 questions were coded and analysed. These interim monitoring observations evidenced an increase in both the proportion of deeper questioning employed overall by participant teachers and the variety of intended mathematical thinking. A full breakdown of the percentages can be seen in Table 7.1.

Question Type	% S1	% S2	% S3	Total % Surface	% D1	% D2	% D3	Total % Deeper	Total %
Factual	2.2%	3.2%	3.4%	8.9%					8.9%
Procedural	5.4%	9.9%	4.7%	20.0%					20.0%
Reasoning	3.4%	3.2%	1.0%	7.6%	5.4%	2.0%	4.2%	11.6%	19.2%
Reflective	3.7%	3.2%	2.0%	8.9%	2.2%	3.4%	5.4%	11.1%	20.0%
Structural					12.8%	4.2%	4.4%	21.4%	21.4%
Derivational					2.2%	4.4%	3.9%	10.6%	10.6%
Total	14.8%	19.5%	11.1%	45.3%	22.7%	14.0%	18.0%	54.7%	100.0%

%S1, %S2, %S3 denote the number of surface level questions in each third of lessons as a percentage of the total questions asked. Similarly, %D1, %D2, %D3 denote the number of deeper level questions in each third of lessons as a percentage of the total questions asked.

Table 7.1. Percentage of questions in each category in the interim monitoring observations.

Monitoring the Depth of Questioning

Overall the percentage of deeper questioning rose from 25.3% in the baseline observations to 54.7%. Although the latter percentage is only based on 5 lessons as opposed to 15 lessons in the baseline observations, a z-score of 10.58 indicates that the proportion of deeper questions in interim monitoring is significantly greater than in baseline ($p<0.001$). This was a very promising indication that the intervention on working with the IMPaCT Taxonomy had made an impact on the depth of questioning of the participant teachers.

There was, however, a substantial difference between the teachers. Teachers P, Q and S all used deeper questioning in the majority of their questioning in the interim monitoring observations. The largest increase in percentage came from Teacher P with a difference of 40.2% and a percentage increase of 176.3%. Teacher Q also more than doubled the percentage of deeper level questioning from the baseline observations to the monitoring observations as can be seen in Table 7.2. Teacher S increased the number of deeper questions with both participant classes. Although the percentage of deeper questions used with his lower attaining class remained lower than that used with the higher attaining class, the former did, however, have a higher percentage increase with regards to their respective starting values.

Teacher	Baseline % Deeper	Interim % Deeper	Actual Difference	Percentage Increase	*z*-score
P	22.8	63	40.2	176.3	4.93977
Q	28.0	66.4	38.4	137.1	7.96807
R	19.4	20.9	1.5	7.7	0.223888
S (Set 1)	32.8	58.6	25.8	78.7	4.436021
S (Set 5)	22.9	42.3	19.4	84.7	3.158611

Table 7.2. Percentage of deeper and surface questioning in the interim monitoring observations.

For Teachers P, Q and S, the z-test statistics indicated that the null hypothesis can be rejected ($p<0.001$). Teacher R only increased by 1.5% in the monitoring observations, a 77.2% lower percentage increase than the next lowest percentage change. The z-score of 0.22 indicates that the baseline and interim proportions of deeper questions are not significantly different in this instance ($p>0.2$) as there is a greater than 20% probability that this difference occurred by chance. On speaking to Teacher R after the lesson was recorded, he had felt the lesson had not gone to plan, and was particularly affected by the late arrival of a number of students due to a languages examination. This can be seen in the analysis of each third of the lesson for Teacher R, where 46.2% of questions were coded deeper level in the first third of the one hour lesson. According to the lesson transcripts, the latecomers arrived 23 minutes into the lesson; the impact of this can be seen in Figure 7.1, where the percentage of deeper level questions at this stage in the lesson dropped to just 5.6%.

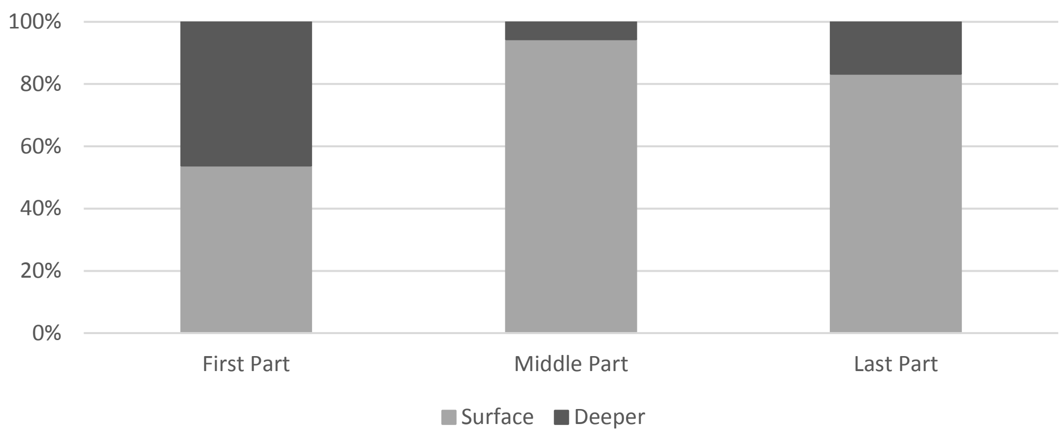


Figure 7.1. Proportion of deeper and surface questioning in each stage of the lesson for Teacher R in the interim monitoring observation.

Monitoring the Type of Questioning

In terms of the proportion of question types, the interim observations provided an early indication that working with the IMPaCT Taxonomy had an effect on the variety of intended mathematical thinking for the learners. This can be seen in Figure 7.2. In the baseline observations, the factual and procedural categories made up 47.8% of all questions asked, as opposed to just 28.9% in the interim observations, with a *z*-score of 6.65 implies that this is a statistically significant difference ($p<0.001$). The largest decrease can be seen in the factual category and the largest increase in the derivational category, with the proportion of derivational questions surpassing the proportion of factual questions. Again, the *z*-test indicates that it is very unlikely that these differences are attributable to chance ($p<0.001$). The differences in the reasoning and reflective categories are not statistically significant ($p>0.2$), although this does not take into account the proportion of surface and deeper questions within these categories which is explored in greater depth in the post-intervention analysis.

The spread of question type for individual teachers again differed greatly, with Teacher R not asking any reasoning or derivational questions in the interim monitoring lesson and the factual and procedural categories made up 60.5%, an increase of 3.1% since the baseline observations, although this rise could be attributed to chance (*p*>0.2). In fact, for Teacher R, the variety of question type was actually less in the interim monitoring lesson than in the baseline observations. There was, however, an increase of 4.7% in the percentage of structural questions posed by Teacher R, although again this increase is not statistically significant (*p*>0.2).

As can be seen in Figure 7.2, both Teacher P and Teacher Q made good progress in lowering the number of factual and procedural questions posed, both statistically significant ($p<0.001$) and significantly increased the number of structural questions asked ($p<0.01$ for Teacher P and $p<0.001$ for Teacher Q). Teacher S achieved a much broader spread in the type of questions in the interim observations.

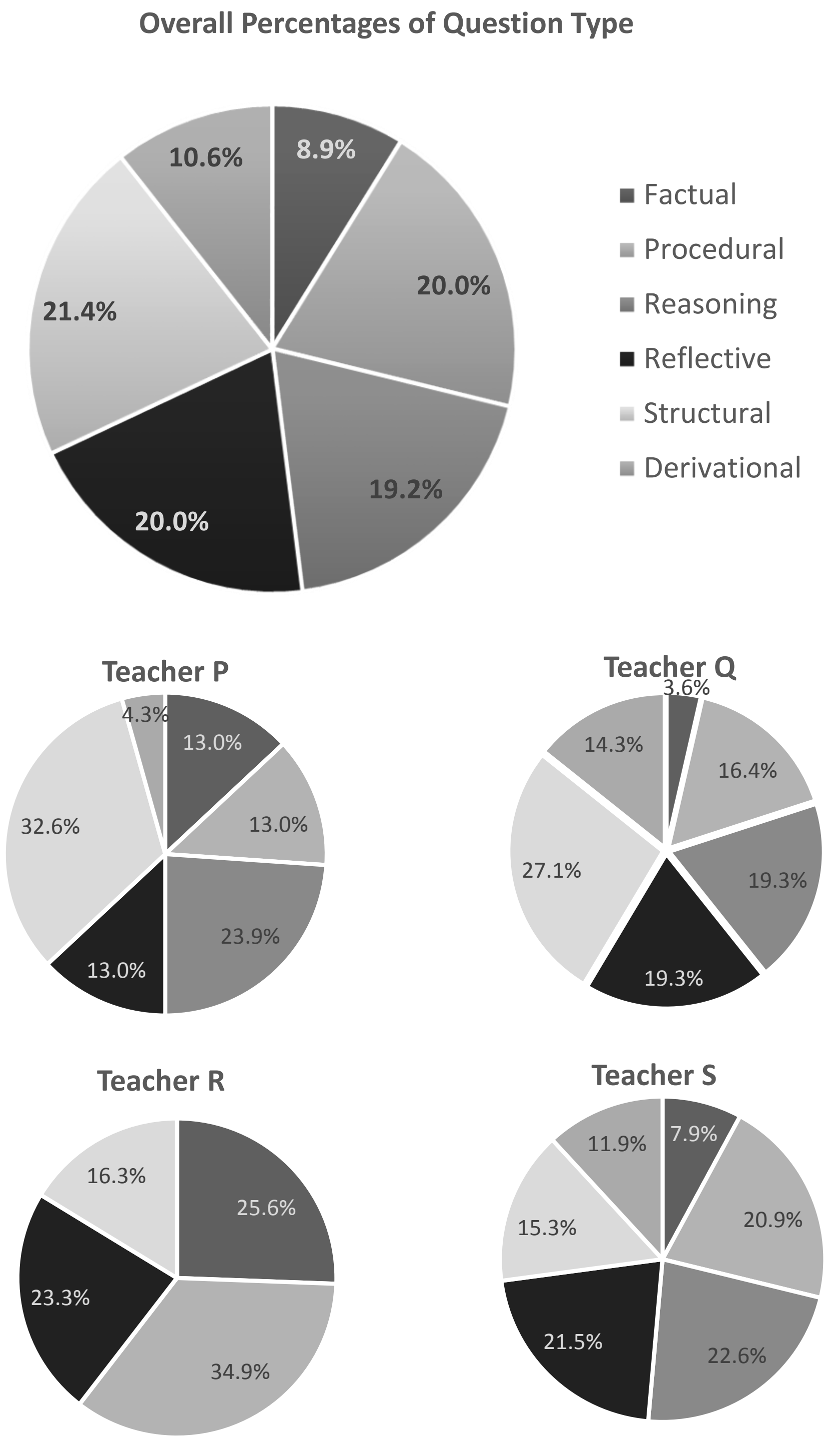


Figure 7.2. Percentages of question type in the interim monitoring observations.

Figure 7.3 shows the breakdown of the percentages of question type for each of the different attainment classes for Teacher S. Although more factual and procedural questions were used with the lower attaining group, more structural and derivational questioning were also utilised.

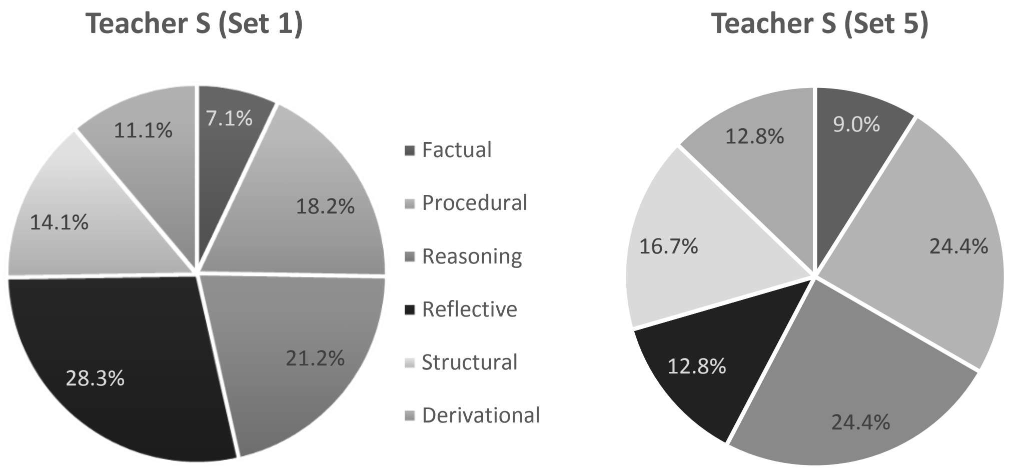


Figure 7.3. Percentages of question type for each class of Teacher S in the interim monitoring observations.

Actions Taken Following the Interim Monitoring Observations

Following the analysis of the interim monitoring lessons, each participant teacher was given a breakdown of the data collected in their lesson and a copy of their baseline data in order to make their own comparisons. Teacher R was disappointed with his progress, although not surprised. Additional time was spent with Teacher R analysing the types of question observed in his lessons to date and further discussion was had about the categories in the IMPaCT Taxonomy and the key question stems in the Venn diagram (see Figure 3.1) to support planning for a greater variety of questioning.

Headline Statistics from the Final Observations

In April and May 2016, a further eight lessons were either observed or recorded, transcribed and coded. Unfortunately, it was not possible to record the last two planned lessons with the lower attaining class of Teacher S. In the 12 observed or recorded lessons available for analysis, a total of 905 questions were transcribed and coded, a mean average of 75.4 questions per lesson. The number of questions coded per lesson ranged from 36 to 140. The teachers' individual mean average of questions per lesson differed greatly:

- Teacher P – 54.7 questions per lesson (increase of 13.7)
- Teacher Q – 117.7 questions per lesson (increase of 23.7)
- Teacher R – 51.7 questions per lesson (decrease of 43)
- Teacher S (Set 1) – 77.7 questions per lesson (increase of 4.5)

If Teacher R's interim monitoring lesson is omitted from these statistics, the overall mean rises to 78.4, just 0.4 less than in the baseline observations, however the number of questions posed by Teacher R is still substantially less than in the baseline assessments.

The breakdown of the percentages of type and depth of questions coded can be seen in the table in Table 7.3.

Question Type	% S1	% S2	% S3	Total % Surface	% D1	% D2	% D3	Total % Deeper	Total %
Factual	4.8%	3.3%	4.9%	12.9%					12.9%
Procedural	6.2%	8.3%	5.0%	19.4%					19.4%
Reasoning	3.2%	2.7%	1.7%	7.5%	6.0%	3.4%	3.8%	13.1%	20.7%
Reflective	3.4%	2.9%	2.1%	8.4%	4.0%	4.9%	3.9%	12.7%	21.1%
Structural					9.1%	3.3%	4.9%	17.2%	17.2%
Derivational					2.8%	3.6%	2.2%	8.6%	8.6%
Total	17.6%	17.1%	13.6%	48.3%	21.8%	15.2%	14.7%	51.7%	100.0%

%S1, %S2, %S3 denote the number of surface level questions in each third of lessons as a percentage of the total questions asked. Similarly, %D1, %D2, %D3 denote the number of deeper level questions in each third of lessons as a percentage of the total questions asked.

Table 7.3. Percentages of questions in each category in the post-intervention observations.

Depth of Questioning Post-Intervention

Overall, the percentage of deeper level questions following working with the IMPaCT Taxonomy rose from 25.3% to 51.7%, an increase of 26.4 percentage points and with a *z*-score of 12.64, indicates that the percentage of deeper questions post-intervention is significantly greater than pre-intervention (p<0.001).

Teacher	% Deeper Pre-intervention	% Deeper Post-intervention	Actual Difference	Percentage Increase	z-score
P	22.8	48.2	25.4	111.4	4.657217
Q	28.0	60.6	32.6	116.4	8.725988
R	19.4	29	9.6	49.5	2.2084
S (Set 1)	32.8	55.8	23	70.1	5.181773

Table 7.4. Percentages of surface and deeper questioning in the post-intervention observations.

Table 7.4 shows that Teacher Q had the highest rise in percentage points for the proportion of deeper level questioning. Despite the setback in the interim monitoring observations, Teacher R still experienced nearly a 50% rise in deeper questioning over the course of the action research, the impact of the intervention being statistically significant (p<0.05). If the interim monitoring lesson is excluded from the analysis, then the percentage of deeper questions rises to 32.1%, 12.7 percentage points more than the baseline observations and a percentage increase of 65.5% with respect to his starting point.

Type of Questioning Post-Intervention

Factual and procedural questions combined made up approximately one third of all questions asked post-intervention. The most common type of question asked in the post-intervention observations was reflective and the least common question types were factual and derivational. The largest percentage change in the proportion of each question type was derivational with a percentage increase of 207% and the z-test indicates that the proportion of derivational questions post-intervention is significantly greater than in pre-intervention (p<0.001), although it remains the lowest question type (see Figure 7.4).

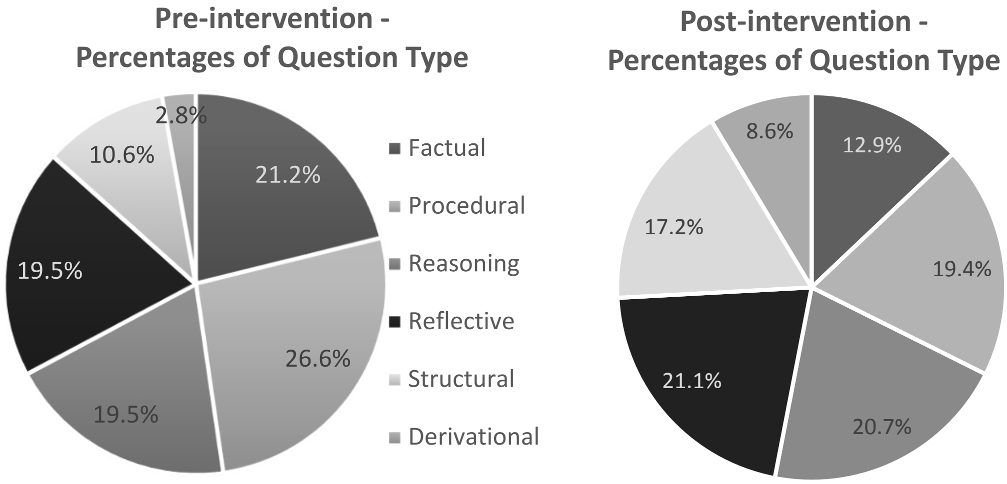


Figure 7.4. Overall percentages of questions type in the post-intervention observations.

41.8% of all questions posed in the post-intervention observations, appeared to intend either reflective or reasoning thinking. Although this was only a 2.8 percentage point increase since the baseline observations, the noticeable difference was the percentage of surface and deeper questions within each of these question types. The reasoning category had 63.2% deeper level questions post-intervention, compared to 34.8% in the pre-intervention observations. An even larger difference was seen in the reflective category where it rose from 26.5% deeper level at the start of the action research to more than double this figure at 60.2% post-intervention. Both of these increases are very unlikely to have occurred by chance (p<0.001).

Figure 7.5 shows how the variety of question type changes for each participant teacher from the start to the end of the main study for this research. Teacher P greatly increased the proportion of questions requiring reasoning and reflective thinking and, conversely, decreased the proportion of factual and procedural questioning used. Teacher Q increased the variety of questions by asking less procedural questions and instead asking more questions which required structural or derivational thinking. Teacher S asked far fewer factual questions and fewer procedural questions and increased the proportion of questions posed which were structural and derivational in nature.

Despite increasing the percentage of deeper questioning, the percentage of factual and procedural questions used by Teacher R actually stayed constant over the course of the research. Further analysis shows that this was achieved by increasing the percentage of deeper level questioning within the reasoning and reflective categories by 44.7 and 17.7 percentage points respectively.

In terms of the percentage of deeper questions asked in the available categories per teacher, in the reasoning classification, all teachers asked a higher percentage of deeper reasoning questions following the intervention than surface reasoning questions. This was the same for the reflective category, with the exception of Teacher R, who took a deeper approach to reflective questioning 34.8% of the time; this rises to 46.2% if the interim monitoring lesson is omitted. Both these figures are over double the baseline percentage of 17.1% deeper reflective questioning for Teacher R.

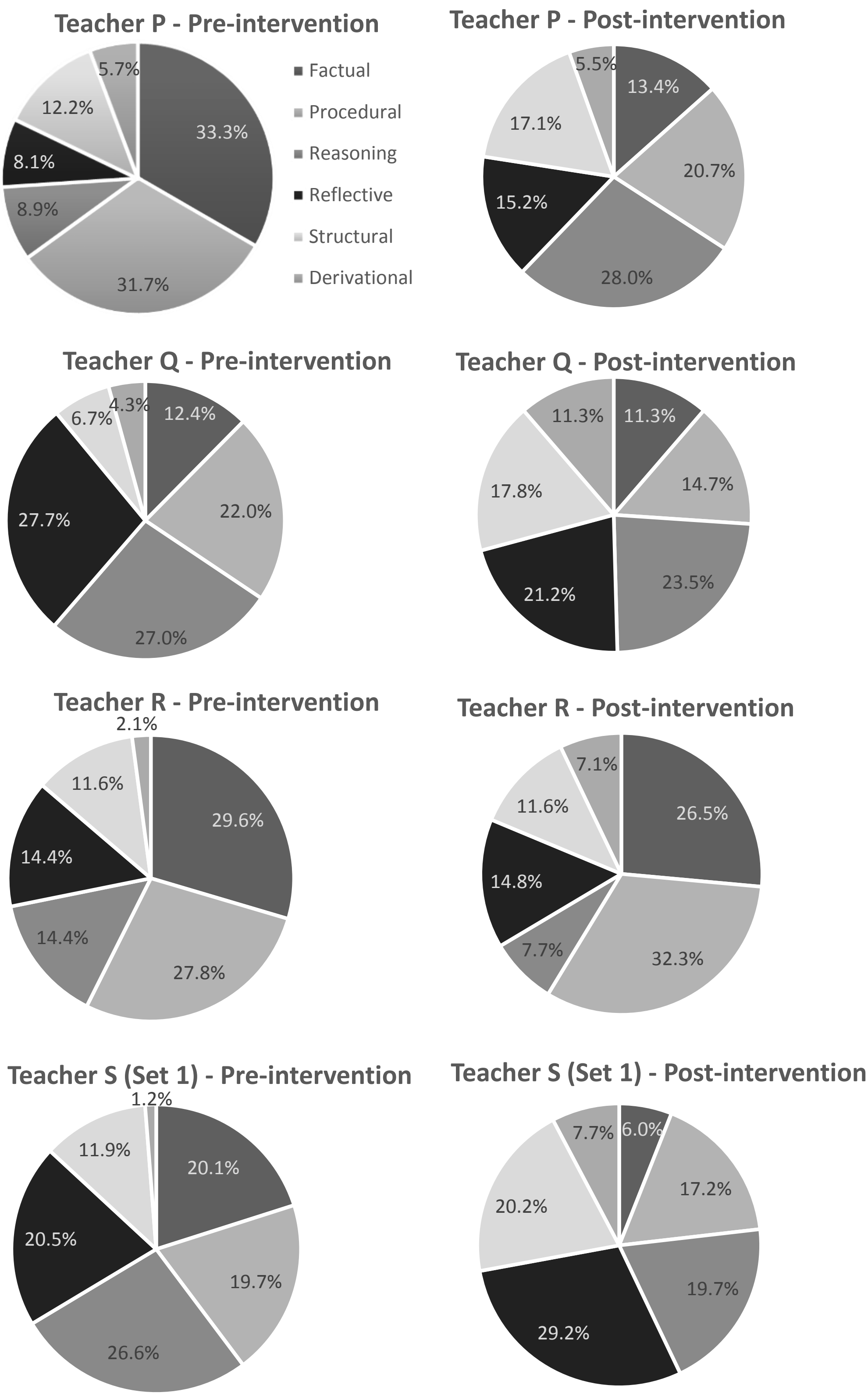


Figure 7.5. Percentages of questions type in the post-intervention observations.

The Lower Attaining Group

As discussed previously, lower attaining class for Teacher S only provided one lesson of post-intervention data, that of the monitoring lesson from February 2016. As a result of this more limited data, the findings have been analysed separately with a degree of caution to making generalisations due to the small sample of questions available for analysis.

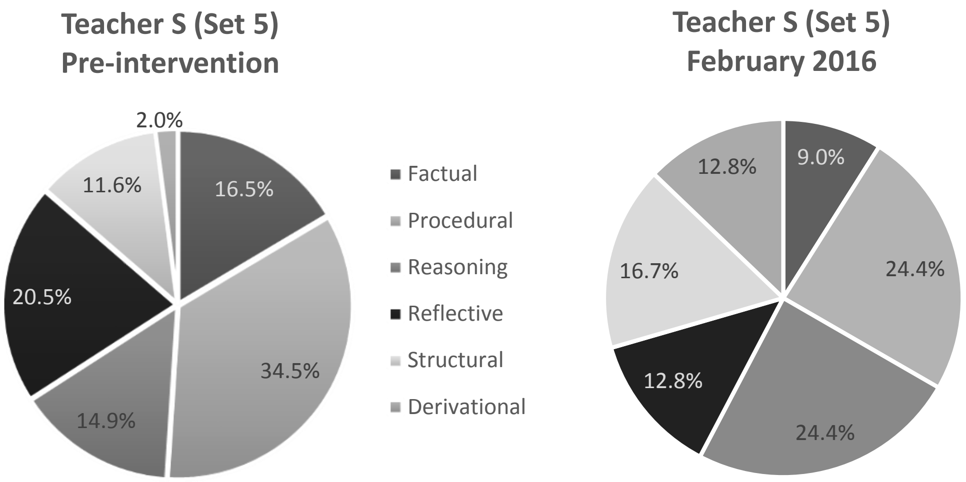


Figure 7.6. Change in percentages of question type for Teacher S (Set 5).

Figure 7.6 shows the increase in the variety of questions posed with the lower attaining class compared to the baseline observations. The biggest percentage increase can be seen in the proportion of opportunities for derivational thinking for the learners and a substantial decrease in the proportions of factual and procedural questioning which allowed for this. The *z*-test on these differences, indicates that the proportions in the February observation are significantly greater than in baseline ($p<0.001$), implying that, despite the smaller sample of questions to analyse compared to the other participant classes, the impact of the intervention was statistically significant with this class too.

Effect of Topic on Depth and Type of Questioning

In both the pilot study (Denton, 2013a) and during the training of the participant teachers in the main study (see Chapter 6), the teachers expressed that their questioning was dictated by the topic they were teaching. To analyse this in greater depth each lesson was coded according to one of the following areas: Geometry, Algebra, Number, Statistics, Problem-Solving, Mixed Topics, and Exam Skills. Table 7.5 shows these topics for each teacher ranked in terms of the proportion of deeper questioning observed in each lesson both pre- and post-intervention.

Teacher	Pre-intervention topic ranked in terms of depth	Post intervention topic ranked in terms of depth
P	1. Geometry 2. Number 3. Exam Skills	1. Statistics 2. Algebra 3. Number
Q	1. Geometry 2. Geometry 3. Geometry	1. Algebra 2. Geometry 3. Algebra
R	1. Exam Skills 2. Algebra 3. Algebra	1. Geometry 2. Mixed Topics 3. Algebra
S (Set 1)	1. Geometry 2. Geometry 3. Mixed Topics	1. Algebra 2. Algebra 3. Geometry

Table 7.5. Topics taught ranked from highest to lowest with respect to the percentage of deeper questioning.

In the pre-intervention observations, all three of the lessons with the lower attaining group with Teacher S were on algebra topics and the interim monitoring observation was on number, specifically percentages.

Pre-intervention, the topic appears to have a big influence on the depth of questioning, with all of the lessons on geometry featuring as the highest percentage of deeper level questioning for every teacher with a geometry lesson in the sample. From Table 7.5 it can be seen however that Teacher Q was only teaching geometry in the baseline observations, which prevents any conclusions being drawn from this for Teacher Q.

Post-intervention the spread of the topics with the highest deeper questioning was much more varied, with an algebra topic coming above a geometry topic for two teachers. The analysis of these lessons were grouped according to the topic area and the mean average found for each percentage of both depth and type of questioning. These results can be seen in Figure 7.7. Pre-intervention, the geometry lessons had the greatest percentage of deeper questions, 12.5 percentage points higher than the second highest of algebra; a percentage difference of 64.5% between algebra and geometry. Post-intervention, a lesson on statistics, specifically Spearman's Rank Correlation Coefficient, achieved the highest percentage of deeper questioning, however as this was a single lesson and there was no data in the baseline observations to make a comparison. Discounting the lesson on statistics, geometry remained the highest percentage of deep questioning post-intervention, however only 7.9 percentage points higher than the second highest of algebra, a percentage difference of just 15.9% between algebra and geometry percentages.

In terms of the question type, although reasoning came a very close second, procedural questioning was actually the highest proportion of question type posed in geometry topics pre-intervention. Procedural was also the highest proportion of question type in the algebra topics pre-intervention. In the number topic pre-intervention the highest proportion was factual questions, however this is based on only one lesson. Post-intervention, reasoning questions had the highest proportion in both the geometry and number topics, and reflective questioning had the highest proportion in the algebra topics.

Geometry – Pre-intervention (based on 6 lessons)

Question Type	Surface	Deeper	Total
Factual	17.8%		17.8%
Procedural	24.1%		24.1%
Reasoning	13.1%	10.5%	23.6%
Reflective	13.3%	6.4%	19.7%
Structural		12.1%	12.1%
Derivational		2.7%	2.7%
Total	68.2%	31.8%	100.0%

Geometry – Post-intervention (based on 3 lessons)

Question Type	Surface	Deeper	Total
Factual	10.6%		10.6%
Procedural	18.2%		18.2%
Reasoning	9.0%	16.0%	25.0%
Reflective	4.7%	13.5%	18.2%
Structural		17.0%	17.0%
Derivational		11.0%	11.0%
Total	42.5%	57.5%	100.0%

Algebra – Pre-intervention (based on 5 lessons)

Question Type	Surface	Deeper	Total
Factual	27.7%		27.7%
Procedural	29.7%		29.7%
Reasoning	11.4%	3.6%	15.0%
Reflective	12.0%	2.2%	14.2%
Structural		11.0%	11.0%
Derivational		2.4%	2.4%
Total	80.7%	19.3%	100.0%

Algebra – Post-intervention (based on 6 lessons)

Question Type	Surface	Deeper	Total
Factual	11.9%		11.9%
Procedural	21.5%		21.5%
Reasoning	5.5%	9.3%	14.9%
Reflective	11.4%	12.3%	23.7%
Structural		19.1%	19.1%
Derivational		9.0%	9.0%
Total	50.4%	49.6%	100.0%

Number – Pre-intervention (based on 1 lesson)

Question Type	Surface	Deeper	Total
Factual	46.2%		46.2%
Procedural	15.4%		15.4%
Reasoning	15.4%	3.8%	19.2%
Reflective	3.8%	0.0%	3.8%
Structural		0.0%	0.0%
Derivational		15.4%	15.4%
Total	80.8%	19.2%	100.0%

Number – Post-intervention (based on 2 lessons)

Question Type	Surface	Deeper	Total
Factual	10.5%		10.5%
Procedural	24.3%		24.3%
Reasoning	15.9%	12.9%	28.8%
Reflective	7.6%	7.1%	14.7%
Structural		12.1%	12.1%
Derivational		9.4%	9.4%
Total	58.4%	41.6%	100.0%

Statistics – Pre-intervention

No data

Statistics – Post-intervention (based on 1 lesson)

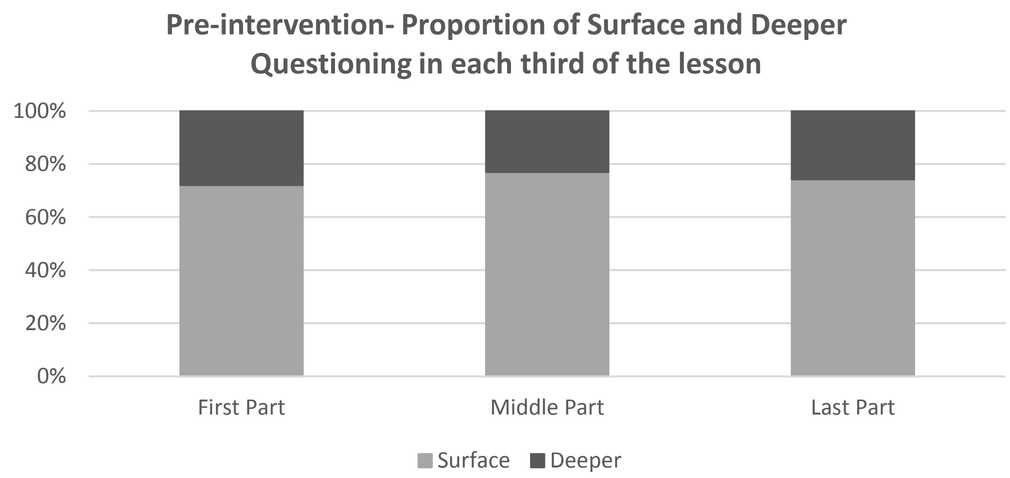
Question Type	Surface	Deeper	Total
Factual	13.0%		13.0%
Procedural	13.0%		13.0%
Reasoning	6.5%	17.4%	23.9%
Reflective	4.3%	8.7%	13.0%
Structural		32.6%	32.6%
Derivational		4.3%	4.3%
Total	37.0%	63.0%	100.0%

Figure 7.7. Percentages of type and depth of questioning with respect to topic.

Four of the remaining lessons were either revising a mixture of topics across attainment targets, focusing on exam technique or problem-solving involving multiple topics. Due to the diverse nature of these lessons, I decided that any comparisons would not be considered valid, so have not included them in this analysis.

Post-Intervention Question Depth with respect to Stage

In the same way as the baseline observation transcripts, to address the research question on how type and depth of questioning is affected by the stage of the lesson, the questions were analysed in relation to which third of the lesson they were posed.



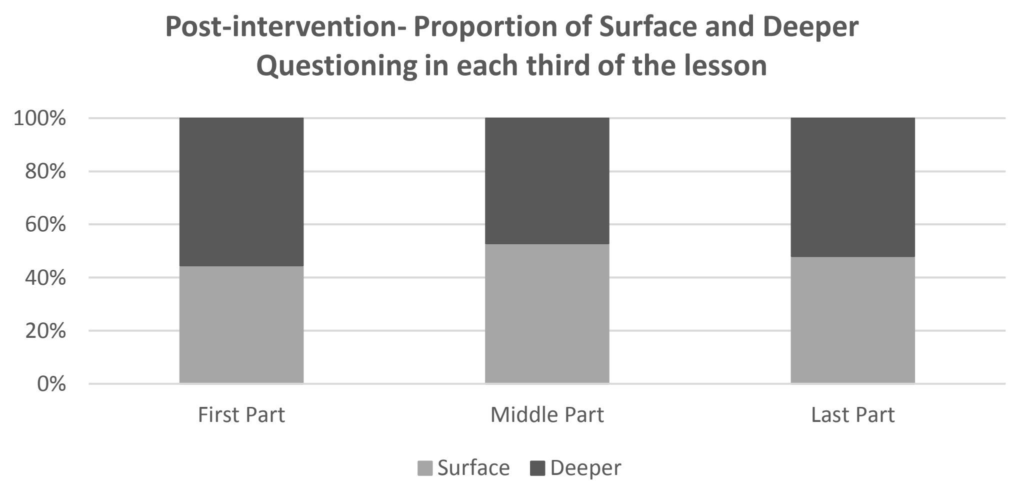


Figure 7.8. Comparison of proportions of surface and deeper questioning with respect to the stage of the lesson pre- and post-intervention.

A higher proportion of deeper questioning was observed in every stage of the lesson compared to the baseline observations. All three stages of the lesson had very similar proportions of surface and deeper questioning in their respective stages (see Figure 7.8). Again, this did vary per teacher as can be seen in Figure 7.9.

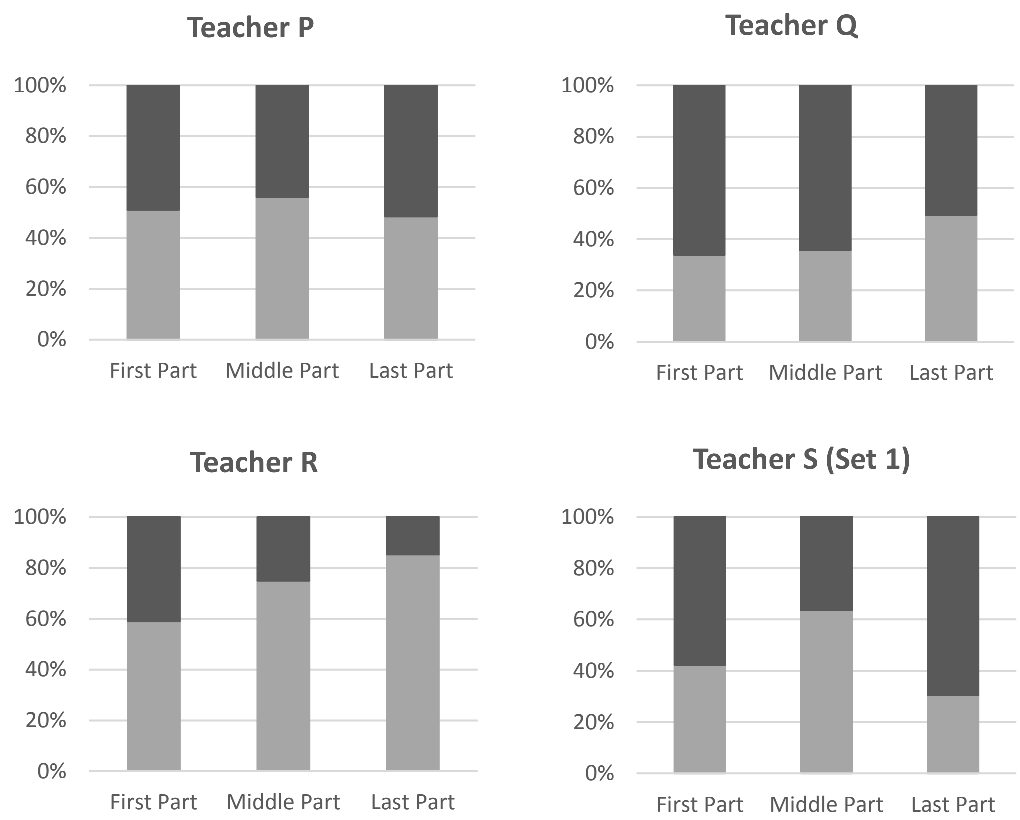


Figure 7.9. Proportions of surface and deeper questioning per teacher at each stage of the lessons post-intervention.

Teachers P, Q and R asked the greatest proportion of deeper level questions at the start of the lesson as they had done in the baseline analysis. Also mirroring the baseline observations, Teacher S continued to ask the highest proportion of deeper level questions at the end of each lesson. All teachers asked a higher proportion of deeper questions at every stage of the lesson compared to the baseline observations.

Post-Intervention Question Type with respect to Stage

The percentage of factual and procedural questioning decreased most noticeably at the start and middle of the lessons. The starts of lessons saw a big increase in the percentage of structural thinking opportunities (see Figure 7.10) and more derivational questions in the middle sections of the lessons. The biggest change in the last section of the lesson was a reduction in the percentage of procedural questions posed.

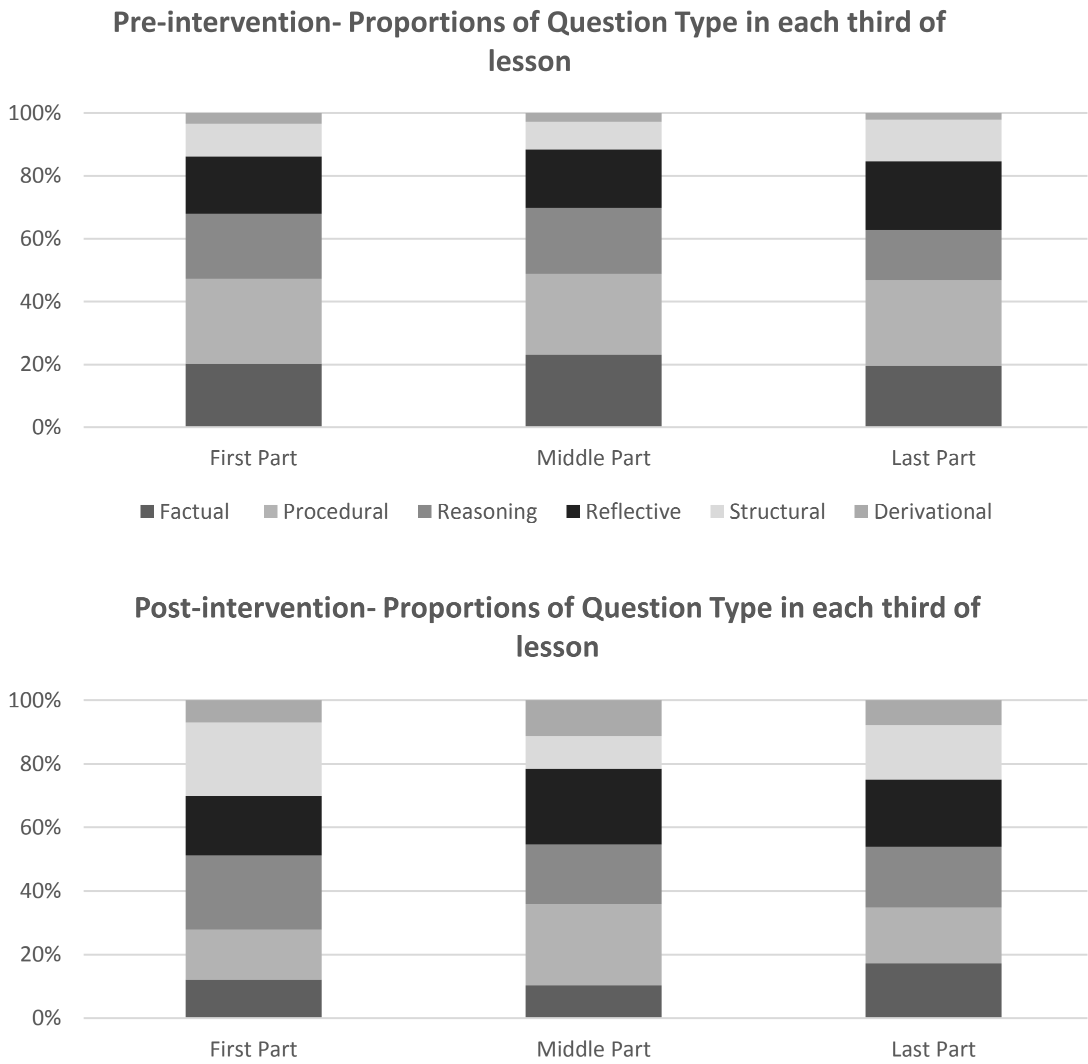


Figure 7.10. Comparison of proportions of question type with respect to the stage of the lesson pre- and post-intervention.

The spread of question type in each third of the lesson varied for each teacher, and the spread also changed post-intervention as shown in Figure 7.11.

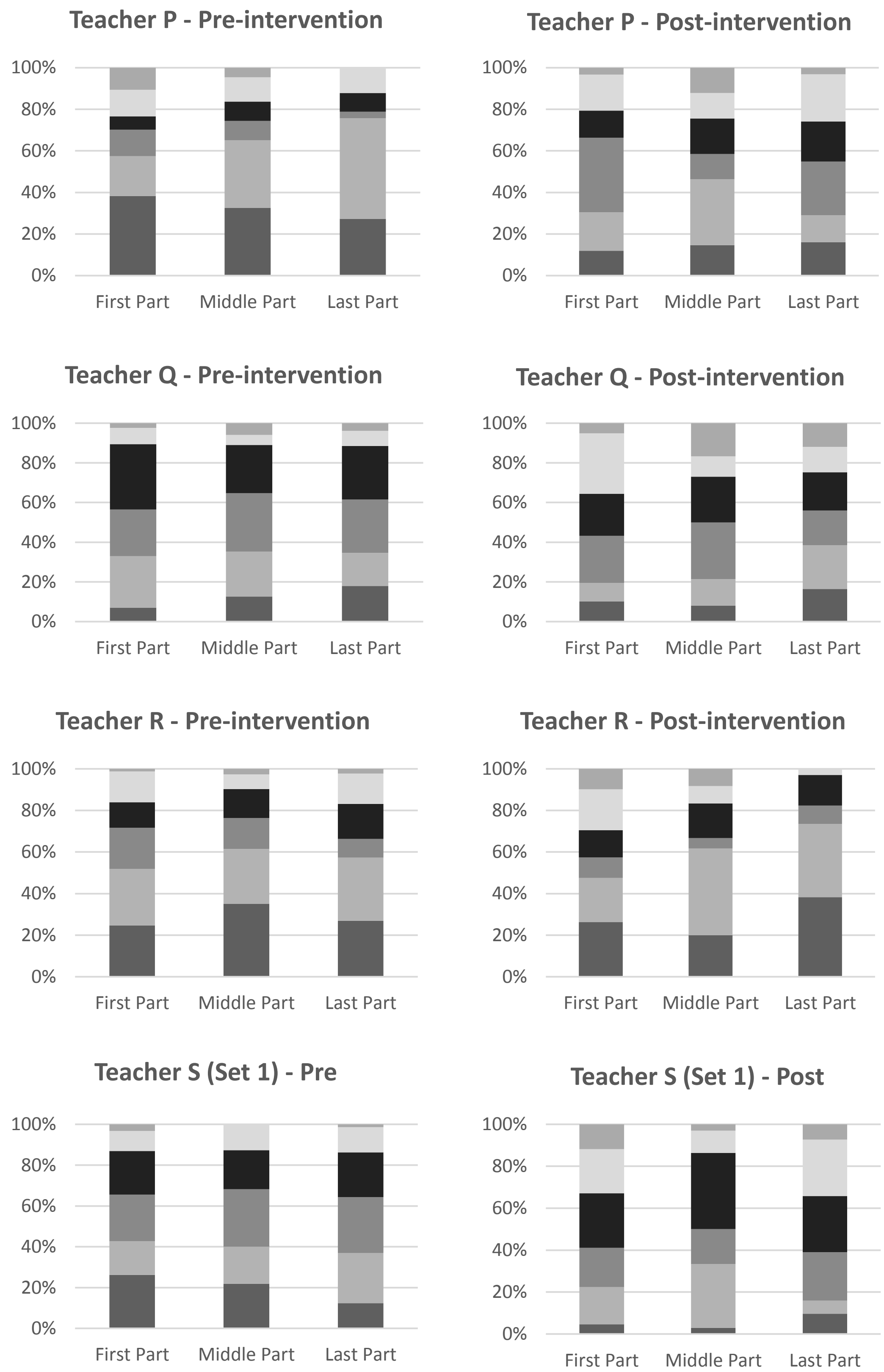


Figure 7.11. Comparison of proportions of question type with respect to the stage of the lesson pre- and post-intervention for individual teachers.

All the teachers increased the proportion of reflective and derivational questions posed during the middle part of the lessons. Similarly, all teachers reduced the proportion of factual questions in the middle parts of lessons, although the change in proportion of procedural questions in the middle part of lessons was less consistent amongst the participant teachers, with Teacher Q reducing the proportion, Teacher P staying approximately the same proportion, and Teachers R and S increasing the proportion. Even with the interim monitoring observation removed from this analysis, Teacher R still increased the proportion of procedural questions used in the middle section of the lessons (see Figure 7.12).

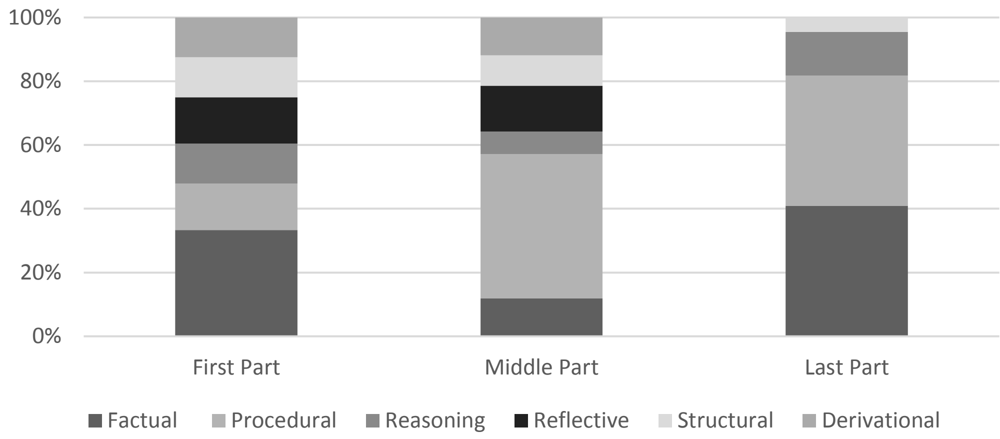


Figure 7.12. Proportions of question type with respect to the stage of the lesson pre- and post-intervention for Teacher R (Interim monitoring observation omitted).

All teachers increased the proportion of structural questions posed at the start of lessons and all except Teacher R also increased the proportion of structural intended thinking in the last part of the lesson too. Teachers P and Q were most consistent in maintaining the variety of question type across the three stages of the lesson.

Post-Intervention Depth of Approach to Reasoning and Reflective Questioning

As can be seen in Figure 7.13, the percentage of deeper questioning employed in the reasoning category increased in every stage of the lesson, with the percentages approximately doubling in the middle and final stages. The largest proportion were still asked at the start and end of the lessons.

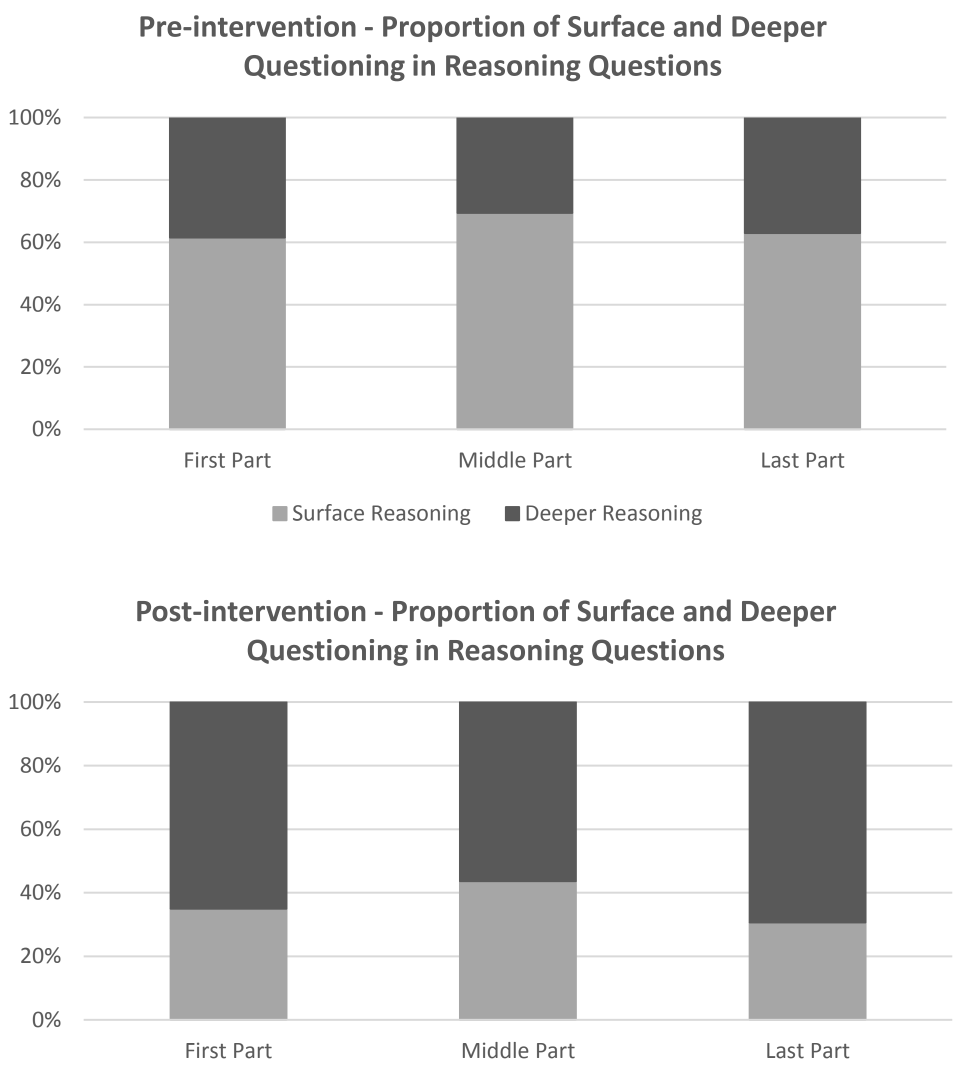
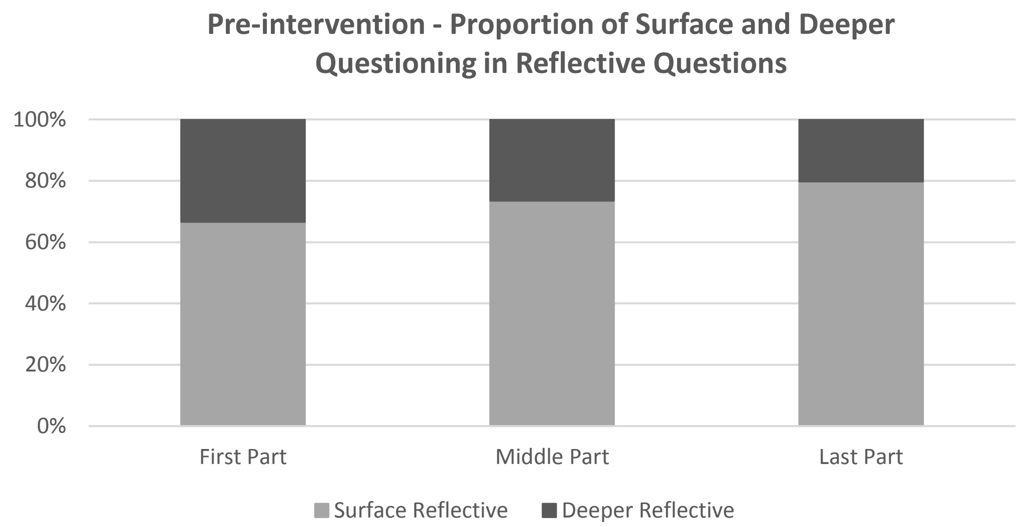


Figure 7.13. Comparison of proportion of surface and deeper approaches in reasoning questioning in each third of the lesson post-intervention.

In the baseline observations, the depth of reflective thinking intended through the teachers' questioning decreased as the lessons progressed, however in the post-intervention observations, the percentage of deeper reflective questioning increased as the lessons progressed, with the largest percentage change in the last stage of the lessons (see Figure 7.14).



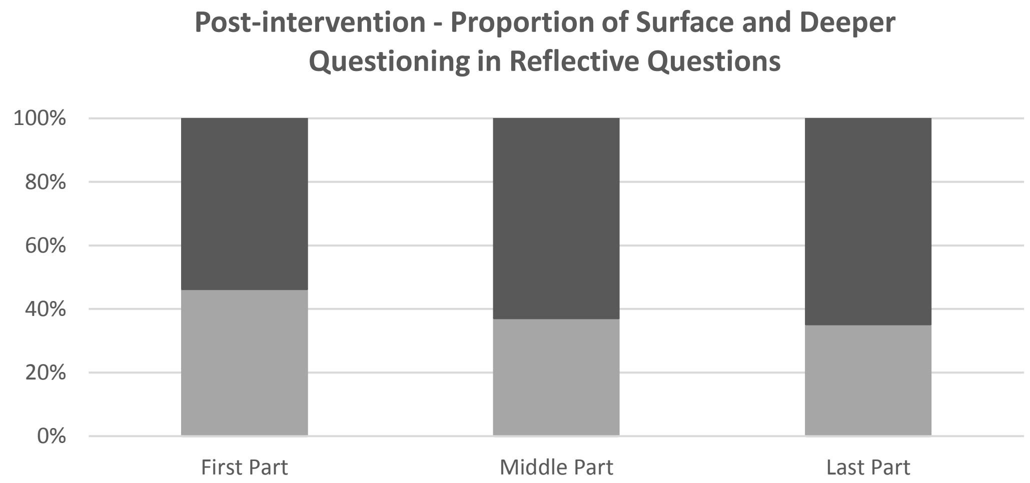


Figure 7.14. Comparison of proportion of surface and deeper approaches in reflective questioning in each third of the lesson post-intervention.

Sociomathematical Norms Post-Intervention

There was evidence of sociomathematical norms being established post-intervention. For example in an interim monitoring lesson, Teacher P asked the question: "Why do you think that one and not that one?" The teacher then proceeded to take alternative views from different individuals to develop an understanding of mathematical difference. This was further evidenced in a lesson observed in April 2016, where she compared and contrasted two approaches taken by different individuals to calculate the area of a triangle.

In the monitoring lesson, Teacher Q asked "I want to know *why* do you think that", stressing "that's what I'm interested in". Whether this is the classroom social norm of justification being established or the sociomathematical norm of mathematical justification could be argued here, however throughout the lesson, Teacher Q asked questions which continued to encourage deep mathematical thought through reminding learners that she was not interested in the solution yet. Furthermore, Teacher Q took the time to collect different methods from different groups of students in the class, asking the rest of the class to analyse them, e.g.

> What is Student X trying to do there?
> Why has he created that first line?
> How does Student X know to say that 3x + 5 is the same as 5 – x?
> What about this (method) from Student Y and Student Z? What are they trying to do?

Here, by discussing the different approaches taken by different learners, Teacher Q is establishing the sociomathematical norms of both mathematical explanation and mathematical justification and cements the importance of discovery over solution:

> **Teacher Q:** I'm not saying if you're right or wrong yet…we'll find out!

Teacher Q also laid the foundations for establishing the norm of mathematical difference:

> Do you get the same if you draw it?
> If we look at the four cases here, is there something similar about all of them?

Post-intervention, Teacher Q also used sorting and classification tasks on two occasions, which allowed for the exploration of mathematical difference:

> What have these go in common?
> That might be the case. How do you know?

Why does that work and not that?

And, indeed, mathematical efficiency:

Is it easy to see your solutions from that graph?

Which method do you prefer and why?

Tell me why this method might be more useful?

In this lesson, Teacher Q also asked "Is that your most efficient method?" and "What would be a really efficient method?", establishing the sociomathematical norm of mathematical efficiency (Yackel & Cobb, 1996). The sociomathematical norms of mathematical elegance and mathematical sophistication were also evident in Teacher Q's final recorded lesson with the simple question:

Could you express it in a nice way with integer values?

For Teacher R, in the middle part of the lesson, there tended to be more questions from the learners than the teacher. While this is not necessarily negative, Teacher R often responded by showing the learner how to do it rather than encouraging learners to think it through themselves. Teacher Q's "you tell me" could be useful here to get learners to talk through their difficulties before relying on the teacher for the answer.

Further examples of establishing mathematical efficiency, and indeed elegance, was evidenced in Teacher Q's final recorded lesson with the simple question: "Did you need to do that?" and by not accepting a monosyllabic reply, it made it a deeper level question. Teacher S also touched on mathematical efficiency and elegance in his post-intervention lesson on vectors:

Teacher S: What paths can we take? Which route is the easiest?

Teacher P encountered a learner giving a particularly procedural response post-intervention:

Teacher P: How do I do 500 divided by 0.2?

Student Z: Put the zero from the bottom on the top

Teacher P dealt with this by praising the quick method, but stressing the mathematical justification of mathematical equivalence from multiplying top and bottom by the same value. This is in contrast to Teacher X, in the pilot study, who walked away from a procedural response. Teacher P on the other hand made clear to the class what constituted an acceptable mathematical explanation and justification.

AfL Post-Intervention

In her interim monitoring observation, Teacher Q used mini-whiteboards for learners to investigate the following question:

> **Teacher Q:** How many points of intersection would there be with two straight lines crossing; a quadratic and a straight line; a cubic and a straight line; or two circles.

Using the whiteboards for investigation, and allowing the learners to discuss their ideas in pairs or small groups, generated a depth of discussion that was not observed pre-intervention. Of the 21 questions which followed, 20 were classified as deep, with a high proportion of structural questions as learners were asked to consider different scenarios, leading to the minimum and maximum number of points of intersection for each problem.

Teacher S also used mini-whiteboards in his interim monitoring lesson with his lower attaining class. Of the 46 questions asked using the mini-whiteboards, only 16 were classified as deeper level questioning, that is 35%. Although this is higher than the percentage of deeper questions asked by Teacher S pre-intervention, it is less than Teacher S for the whole interim monitoring lesson with this class and also less than the average for all his lessons post-intervention, suggesting that simply using mini-whiteboards may not automatically increase the depth of questioning employed by teachers.

As in the baseline observations, Teacher S gave his lower attaining class thinking time before taking feedback. Allowing the learners time to think in this way allowed for derivational thought and deeper reasoning when asked to explain why.

In his first observation in May 2016, Teacher R used mini-whiteboards with the class on visualising vectors. He only asked five questions in conjunction with the whiteboards, four of which took a deeper approach to learners' mathematical thinking. Although based on very limited data, this 80% is far higher than the 32.1% for the final two observed lessons. The high proportion of deeper questioning could be as a result of using questioning in conjunction with the mini-whiteboards, allowing the learners to think structurally about vectors and use this structural thought to lead onto more derivational thought.

Reliability and Validity Testing

As discussed in Chapter 5, it was decided that the inter-coder reliability testing was needed more to test for bias in the codings before and after the intervention, as opposed to being vital that two different coders would code a lesson in exactly the same way. The IMPaCT taxonomy is designed to be a tool for teachers to reflect on their use of questioning, with the aim of increasing the variety and depth of questioning employed, as opposed to being an evaluation tool to make judgements, as there is a degree of subjectivity in applying the taxonomy.

The issue with the crossover between deeper reasoning and reflective categories with the structural and derivational categories arose again. The second observer post-intervention felt that, in the context of the lesson, 10 out of the 140 questions could be placed in more than one category. As a result, a range of percentage agreement was calculated to reflect the smallest and largest percentage agreements between the two coders. The percentage agreement of whether the questions were coded as surface or deeper level was particularly high between 92.9% and 96.4%, an exact match of both question type and depth was between 81.3% and 90.2%. The agreements of surface and deeper classifications were very similar pre- and post-intervention and the both ends of the range of the one-to-one agreement between codings is higher than the pre-intervention reliability testing, suggesting minimal bias on the coding, implying the results are reliable in this respect.

Teachers' Voice Post-Intervention

After the final lesson observations and recordings, the participant teachers were interviewed again (see Appendix 7). The interviews were audio recorded and transcribed and the transcripts were analysed using a simple descriptive coding framework which developed through the analysis. The codes were then grouped into themes as shown in Table 7.6. Five themes emerged from the post-intervention interviews.

The first theme which emerged was that the participant teachers found the IMPaCT taxonomy easy to use. They liked the Venn diagram format to help distinguish between surface and deeper approaches to questioning:

> **Teacher Q**: It's easy to read.
>
> **Teacher R:** It's really clear, the Venn diagram really helps.
>
> **Teachers S**: Very straightforward…it's clearly labelled.

Teacher P admitted to still struggling with differentiating between the different categories in terms of the meaning of the words:

> **Teacher P**: I find some of the actual words difficult to actually understand. I mean but the actual process itself is not that difficult.

Topic in Interview	Code	Theme
Clear, easy to read format	CLEAR	Teachers find the IMPaCT Taxonomy easy to use
Venn diagram supports the process	VENN	
More relevant to maths than Bloom's Taxonomy	BLOOM	
More time spent planning for questioning	TIME	Effect of using the IMPaCT Taxonomy on teachers' questioning
More thought to depth of questioning	DEPTH	
More conscious of questioning used	CONS.	
Analysis of their questioning useful	ANALY	
Questioning more important than the resource	RESOURCE	Change in understanding importance of questioning
Deeper questioning linked to a deeper learner understanding	DEEP	
Questioning used to unpick understanding	QfL	
Learners expected to justify their reasoning	JUSTIFY	Understanding of benefits to learners' thinking
Learners think more for themselves	THINK	
Perceived change in questions learners were asking	LEARNER Qs	
Lack of mini-whiteboards and working pens to use in mathematics departments	LACK	Limited resources to support use of AfL
Use of scrap paper as an alternative to using mini-whiteboards	MWB	

Table 7.6. Thematic analysis of the post-intervention teacher interviews

Nevertheless, Teacher P found the IMPaCT Taxonomy easier to apply to her questioning in mathematics than Bloom's Taxonomy, of which she was previously critical and remained so in the post-intervention interviews:

> **Teacher P:** [The IMPaCT Taxonomy is] much more relevant to maths to be honest. I've always struggled with Bloom's Taxonomy.

Teacher S also preferred using the IMPaCT Taxonomy to Bloom's Taxonomy:

> **Teacher S:** It [the IMPaCT Taxonomy] applies to maths much more than Bloom's Taxonomy [...] it's far easier to go that way through than through the Bloom's Taxonomy.

Teacher R felt he had not used Bloom's Taxonomy enough in the past to compare using it to using the IMPaCT Taxonomy. Teacher Q on the other hand, liked how the IMPaCT Taxonomy was not specifically a hierarchy of levelling questioning:

> **Teacher Q:** I quite like the IMPaCT Taxonomy from the fact that the questions actually do overlap, but you can actually see how you can take a particular question into the deeper understanding rather than just thinking it's like a laddering thing.

The second theme which emerged from the post-intervention interviews was the positive effect that teachers felt working with the IMPaCT Taxonomy had had on their on their questioning:

> **Teacher Q:** It has made me focus a little bit more on this, the actual making sure that you were asking a question that relates to the deeper part of the understanding not just a superficial one.
>
> **Teacher R:** It made me consciously think about what questions I would have to ask.
>
> **Teacher S:** It makes you think more about the questioning you use.

For Teacher Q, the analysis of the lesson breakdown in the baseline and interim monitoring observations gave her reassurance of how her questioning was improving as a result of working with the IMPaCT Taxonomy:

> **Teacher Q:** [The IMPaCT Taxonomy interim analysis] made me think ah, I'm definitely on the right track.

All the teachers indicated in the post-intervention interviews that they dedicated more time to planning for questioning than pre-intervention and their comments, compared to the pre-intervention interviews, suggest that their attitudes to the importance of planning for questioning had changed. Pre-intervention the planning was much more focused on the activity rather than the questioning, however post-intervention comments suggested that the participant teachers had a better understanding of the importance of planning for questioning:

> **Teacher S:** It's all to do with questioning. It's not to do with the resource. You can get the most whizz bang resource but if you can't question students correctly then they don't understand.
>
> **Teacher P:** I've been looking at using through the questioning rather than sort of, worksheets and things like that. As a way of pushing them forward, so it has helped with that.

Furthermore, the participant teachers appeared to have a better understanding of the impact of their questioning in terms of the benefits to the learners:

> **Teacher P:** It gets them more to explain their findings and justify their answers [...] It gives them a bit of a deeper understanding
>
> **Teacher Q:** They get to a point where they're really understanding the answer they're giving.
>
> **Teacher R:** [Helps learners to] develop the process themselves as opposed to just giving it to them.
>
> **Teacher S:** Can they answer questions or can they justify the answers?

In addition, the Teacher S felt he had been able to use the IMPaCT Taxonomy outside his two participant groups:

> **Teacher S:** It's relevant to all classes. A Level, down to the weakest ability (sic) Year 9.

Teacher P also commented on how the IMPaCT Taxonomy had helped her develop her questioning with A Level classes as a means of extending higher attaining learners, although she had not used it specifically with lower attaining classes.

Teachers felt that their questioning would have had an effect on how their learners understand mathematics:

> **Teacher S:** It's helping me to think more about my questioning, so therefore it will reflect onto them [the learners].

However, it was only Teachers P and R who could evidence this impact:

Teacher P: Some of the questions they [the learners] were asking me were quite impressive, so that could be as an outcome of the questions I was asking them.

Teacher R: I did a lesson where [...] we were doing surface area of cuboids and prisms and instead of giving them the formula or giving them and example, I gave them the 3-D shapes and asked them to calculate the surface area and through discussions with each other, they were able to develop it. So they had a much more deeper understanding as to if I'd just given them the formula.

The limited availability of mini-whiteboards and pens, which was a theme from the pre-intervention interviews, remained a theme in the post intervention interviews, although Teacher S had actively tried to get around this issue since the pre-intervention interviews:

Teacher S: I try to use mini-whiteboards if I have got the resource. If not then I'll use scrap paper on the same kind of lines.

Teacher S also spoke about 'bouncing' questions around the room, which he did not mention in the pre-intervention interviews, and made the link about how this helped to establish the sociomathematical norm of mathematical difference:

Teacher S: You ask a question and then I'll try to bounce that question on to other students to expand their reasonings behind. So you kind of get the understanding throughout three or four [learners] then how they explain their question and answer it in different kind of ways.

In the post-intervention interview, Teacher P noted that the learners in the participant class had started to ask deeper questions themselves:

Teacher P: It kind of worked both ways [both the learners and the teacher developing their questioning], so that's something I found particularly interesting.

I had also observed this development of learners' questioning in two different lessons with Teacher P post-intervention, where the learners asked deeper questions themselves than had been noted in the pre-intervention observations, for example:

What's the difference between methods?
How can I tell which to use and when?
What's the easiest way to do this?

This chapter has presented the results from the interim monitoring observations, and the post-intervention observations and interviews. It has made comparisons with the pre-intervention observations and interviews from Chapter 5. Chapter 8 discusses these results and comparisons in more depth, and uses the findings to address the three research questions posed at the start of this thesis.

8. Discussion

This chapter discusses the findings presented in Chapter 7 in relation to the literature review, the pilot study and the findings from the action research cycle 1 analysis and discusses how these findings relate to each of the research questions for this thesis. The limitations of the research are also discussed.

Factors Affecting Questioning

What factors affect the type and depth of questioning used by mathematics teachers?

As discussed in the literature review, analysis of teachers' questions in isolation from the context (Yang, 2006) or how the teachers deal with the response (Kawanaka & Stigler, 2000) will not reveal whether a question, irrespective of being open or closed, surface or deeper, elicited higher-order thinking skills. For example, an open question such as 'the answer is 3 what is the question?' may not illicit any more higher-order thinking than a closed question such as '1+2=3, 2+1=3, 3+0=3, 4+x=3. What is x?' The latter, although closed, requires a conceptual understanding of directed number through pattern spotting and generalisation, as opposed to the open question which would only demand a higher level of mathematical thinking if the teacher probes the connection between the learners' responses and how one answer can be developed to produce an infinite number of possible solutions. In both these examples, it is how the teacher uses the responses from the learners in further discussion (Kawanaka & Stigler, 2000; Holster, 2006), and how they treat an open task (Chapman, 2013), that will develop conceptual understanding, as opposed to the initial open or closed nature of the question or activity. The same applies for both categorising the depth and type of questioning using the IMPaCT Taxonomy, as found in the pilot school when, during teacher training, the mathematics staff struggled to classify a set of questions without seeing in which context the questions were asked, how the learner responded, and how the teacher accepted that response.

The questions collected in the main study were therefore analysed according to several contextual factors, to see which factors had the greatest influence on the type and depth of questioning observed. The five factors considered were:

i. The participant teacher;
ii. the level of attainment of the class;
iii. the stage in the lesson;
iv. questions which were asked in conjunction with one of the following AfL questioning techniques: mini whiteboards; 'no hands up' and 'wait time'; discussion in pairs or small groups before taking feedback (Hodgen & Wiliam, 2006);
v. the topic being taught.

i. The Participant Teacher

There was a clear difference at the start of the intervention between the questions asked by the more experienced and less experienced teachers, and between the two different attainment classes with the same teacher. However, the findings in Figures 5.3-5.5 suggest that regardless of the level of experience or competence of the teacher, structural and derivational questioning were equally lacking in eliciting this type of thinking from learners in mathematics lessons at the start of the intervention. Post-intervention, the gap between the more experienced and less experienced teachers closed to some extent, particularly for Teacher P in terms of the variety of questions she was asking. This implies that, with the same training, the effect of the experience of the teacher on the depth and variety of questioning is reduced. This agrees with Chapman (2013), that it is down to the teacher to "optimize the learning potential" (p. 1)

It should also be noted that Teacher Q, with the highest proportion of deeper questioning throughout the research, was also the teacher with the highest grade from her last formal lesson observation, so further research would be needed to consider to what extent the depth of questioning was as a result of experience or her calibre of teaching in general.

ii. Level of Attainment

There was a notable difference between the depth and variety of questions posed by Teacher S with his higher and lower attaining classes. The number of procedural questions pre-intervention in the lower attaining group was nearly double that of the higher attaining class. Teacher S expressed, in the post-intervention interviews, that the IMPaCT Taxonomy was relevant for all levels of attainment, however, although the depth and type of questioning increased in the interim monitoring observation for the lower attaining class, the difference between the two classes widened. In her interview, Teacher P discussed how she was using the IMPaCT Taxonomy as a means of extending the learning of "higher ability" students. This implies that these teachers view learner attainment as having an impact on their questioning.

In the pre-intervention results in Chapter 5, data from the higher attaining class of Teacher S was similar to that of Teacher Q, but more like Teachers P and R for his lower attaining class in terms of spread of type of questioning. This suggests that an experienced teacher, capable of asking a range of question types, will revert to asking more factual and procedural questions with lower attaining learners (see Figure 5.10). Teacher S encouraged his higher attaining group to give mathematical reasons whilst accepting more procedural responses from the lower attaining group. This suggests that teachers may attach a lesser importance to establishing concrete sociomathematical norms with lower attaining learners, by accepting an explanation as opposed to a mathematical explanation and justification. It should be noted that these questions are coded in the context of the lesson and the learners, so while I would expect the questions to be of a higher-order for the top set class, I believe the spread of surface and deeper questioning and the range of question types should not be affected by the level of attainment of the learners. This stance is based on Whitenack and Yackel (2002), who state that all learners can benefit from classroom discourse involving explaining and justifying thinking, and also Watson et al. (2003), who found that lower attaining learners can experience *deep progress* when teachers model "the kinds of deep questioning which students can do for themselves in future" (p.43).

Despite the perception of the participant teachers in the pilot study that "ability" dictates the type and depth of planned questioning, and the results from the two different attainment classes discussed above, I believe that each of the questioning categories in the IMPaCT Taxonomy can be accessed at different levels of mathematical attainment. For example the following questions on place value allow learners to consider the structure of the mathematics and develop deeper reasoning skills:

- Show me a number between 0.35 and 0.39
- Convince me that that 0.35 is greater than 0.035
- Show me a number [that,] when multiplied by 10 [,] gives an answer greater than 350 (DfE, 2013)

Teacher P in the main study viewed the IMPaCT Taxonomy as a tool to extend the thinking of her most able learners, however, as seen by Teacher X in the pilot study (Denton, 2013a), focusing on the structure of the mathematics as opposed to the procedure could help lower attaining learners to formulate a bigger picture of place value by combining this structural knowledge of place value in decimals and multiplication and division of whole numbers by powers of ten, in order to extend this understanding to multiplying and dividing decimals by powers of ten, for example:

- Convince me that 35 ÷ 10 and 350 ÷ 100 give the same answer (DfE, 2013)

This could help to address the common misconception that to multiply by 10, you just add a zero (Ball et al., 2001).

iii. The Stage of the Lesson

The stage of the lesson had an influence on the type of questioning posed in the observed lessons, with more reflective questions posed at the end of lessons, mirroring the findings of Kawanaka and Stigler (2000). Working with the IMPaCT Taxonomy, however, had an effect on this, resulting post-intervention in a higher proportion of reflective questions being asked throughout the lessons post-intervention.

In analysing the depth of questioning with respect to the stage of the lesson, I found that, post-intervention, although the proportions increased for each teacher, the depth of questioning mirrored the baseline in terms of ranking (see Figure 7.8). This suggests that it is a matter of questioning style of individual teachers as to when they pose questions which require learners to take a deeper approach to their thinking.

In the pilot study, Teacher W explained how she used quick-fire surface level questioning at the start of a lesson to 'get things going'. While I do not dispute that this is a valid strategy, the evidence from Teacher Q's lessons, particularly post-intervention, highlight that it is possible to begin a lesson with deeper structural questioning which allowed for deeper reasoning and reflection throughout the lesson rather than just in the summation phase of the lesson. This is not to say that there is no place for factual and procedural questioning (Fan & Bokhove, 2014) as previously discussed. However teachers should aim to give more focus throughout the lesson on structural and derivational questions which could in turn lead to deeper reasoning and reflection, as these deeper classifications in the IMPaCT Taxonomy focus on justification which could come from a structural understanding of mathematics, as opposed to the more surface level explanation which could rely more on a procedural understanding.

iv. AfL Questioning Techniques

The analysis of questions posed in conjunction with AfL questioning techniques supports the findings from the pilot study that the depth and type of questioning increases when using these techniques. However, particularly with the case of mini-whiteboards, this is not necessarily an automatic by-product of AfL techniques, as it is dependent on how the teacher uses mini-whiteboards to explore the mathematics rather than a means of assessing the whole class through quick-fire questioning. Mini-whiteboards were used in lessons by all the participant teachers, however they only produced a greater depth and variety of questioning when the mini-whiteboards were used to generate class discussion. Without the class discussion, their usage became a tool for testing learners' factual and procedural knowledge. Asking questions with multiple answers can support learner discussion as learners can look at others' responses and decide if they are correct or not, and can consider whether there is a finite or infinite number of solutions to the problem. Questions with a unique correct answer can also generate deeper thought if the teacher asks the learners to justify their solution verbally, or identify how an incorrect response might have been made. The results of this research suggest that AfL techniques should be used as a tool to encourage and probe deeper mathematical thinking.

In the pilot study, the participant teachers expressed reservations about using mini-whiteboards with older or more challenging learners. This was not seen as a limiting factor for the participant teachers in the main study, where a lack of resources was cited as the only reason they did not use mini-whiteboards as much as they would like to.

Wait time was used by Teacher S with his lower attaining group and the questions asked in conjunction with this AfL questioning technique had a higher percentage of deeper questioning than proportion of deeper questions asked overall with this class. However, randomly selecting learners to respond without allowing wait time, as observed happening with Teacher P, may have stopped the rest of the class from thinking through the question for themselves, thus not allowing them sufficient time to process the question (Rowe, 1986).

Black and Wiliam (1998) and Cobb and Yackel (1996) found that if learners feel they are simply trying to guess the answer in the teacher's head, their mathematical thinking will be inhibited, as discussed in Chapter 1. This seemed to be the case with Teacher Q pre-intervention, when she appeared to be looking for a pre-determined response when she asked what error the learner may have made, and then, when the learner made little attempt to respond, quickly told the learner that it did not go through origin. For Teacher S, the converse was true when he did not know the answer to a question about circle theorems and, as a result, a much richer conversation was had. This links to Wood's (1998) focusing rather than funnelling where, in the latter, learners believe they are 'determining a set of procedures that the teacher already knows and that it is their obligation to learn' (Wood, 1998, p. 175). Without Teacher S looking for a pre-determined response and funnelling the learners towards an expected answer, the learners may have felt less inhibited, hence enabling the learners to focus their own learning in the direction of their choosing.

v. The Topic

In the pilot study, the topic being taught was perceived by the teachers as being the biggest contributing factor to the type and depth of questioning used in lessons, however in the pilot observations it was found that the teacher had the biggest influence on the type and depth of questioning used; a functional lesson on Bowland mathematics was found to have a greater proportion of surface level questioning than another lesson on algebra. In the main study I investigated this phenomenon further and I found that, pre-intervention, geometry lessons had a higher proportion of deeper level questioning than other topics. One possible explanation for the higher proportion of reasoning in the higher attaining class of Teacher S could be the high proportion of questioning on circle theorems in this class' lessons, where two of the three lessons were on this topic. Circle theorems examination questions usually require learners to explain how they got their answer, that is, the reasoning they employed (see, for example OCR, 2011; Edexcel, 2012). The impact of topic on the depth and type of questioning, however, shifted post-intervention, with lessons on algebra topics producing the highest proportion of deeper questioning for two of the participant teachers and, overall, the gap between geometry and algebra lessons in terms of depth and variety of questioning narrowed considerably.

With individual teachers, the topic they were teaching had an impact on their questioning, with geometry topics having a more varied and deeper approach to questioning compared to algebraic and number topics. There was only one statistics topic in the sample, which was observed post-intervention; although it had a good selection of deep and varied questioning, generalisations cannot be made due to the small sample.

So, is it the teacher, class, stage of lesson, activity or topic which most dictates the type and depth of questioning? Circle theorems were observed being taught by three out of the four participant teachers. Teacher Q approached the topic in a different way to the other two teachers, presenting learners at the start of the lesson with a set of circle theorems problems without the circles drawn in, with the simple task to "find the value of the angles indicated". The discussion which ensued was rich and on the whole deep and reinforced the circle theorems as learners struggled with what information they knew and what they needed the circle for. This is a great example of Piaget's cognitive conflict (Wood, 1999) being used to move learners on in their thinking, as without the circles drawn in, the learners realised that they could not apply the theorems. This was followed by a sorting task in pairs which consolidated the learning. In general, the geometry lessons allowed for more conceptual thinking due to their visual nature, compared to the more abstract nature of algebra, however Teacher Q also achieved conceptual thinking with her lesson on simultaneous equations, where she made the task visual and investigative, allowing for structural thinking from the learners from the outset.

Furthermore, Teacher P achieved deeper and more varied questioning post-intervention, with both a lesson on approximations to calculations and a lesson on solving quadratics, than she achieved pre-intervention on a geometry lesson on transformations. This was also the case for Teacher S with two lessons post-intervention on the factor theorem and matrices respectively, which although had fewer reasoning questions than the two lessons pre-intervention on circle theorems, had a greater variety and a higher proportion of deeper questioning overall.

This implies once again that topic does play a major role in dictating the type and depth of questioning for individual teachers, however it is still more dependent on how the teacher approaches the questioning and suggests that working with the IMPaCT Taxonomy can close the gap between the differences caused by topic.

Working with the IMPaCT Taxonomy

Does working with the IMPaCT Taxonomy affect the type and depth of questioning used in mathematics lessons?

Much of the discussion for the first research question above suggests that working with the IMPaCT Taxonomy had a positive impact on the variety of the type and depth of questioning employed by teachers in the main study. Furthermore, the comparisons between the depth and type of intervention pre- and post-intervention show that the teachers increased the depth of their questioning by 30.5 percentage points. The literature review found that the changes in education as a result of Bloom's Taxonomy was at policy level rather than classroom level (Anderson & Sosniak, 1994). In both the pre- and post-intervention interviews, the teachers expressed difficulty applying Bloom's to their questioning in mathematics, however they found the IMPaCT Taxonomy much easier to implement, which suggests that the IMPaCT Taxonomy could produce a greater change at classroom level for mathematics than Bloom's Taxonomy.

As stated in the literature review, sociomathematical norms encourage mathematical justification (Gerson & Bateman, 2011), which, in turn, develop mathematical reasoning skills in learners (Whitenack & Yackel, 2002). The increase in the proportion of deeper reasoning questioning evidences how using the IMPaCT Taxonomy, and explicitly sharing Yackel's and Cobb's (1996) emergent perspective with the participant teachers (see Appendix 14), supported the participant teachers to establish sociomathematical norms in their classrooms. For example, Teacher P dealt with a very procedural response from a learner by first praising the quick method, but then stressing the importance of mathematical justification through demonstrating mathematical equivalence. This could have led to deeper thought for this learner if the question "Would that always work?" had been asked, encouraging the learner to justify mathematically for himself rather than the teacher modelling the justification. This could also have helped the learner to develop mathematical autonomy (Holster, 2006).

The learners in the pilot study questionnaire believed that their teachers gave greater importance to explanations than simply the final answer (Denton, 2013a). This was also found to be the case in the baseline observations in the main study. What was different though, was how the teachers established social norms for these explanations pre-intervention. However there was a noticeable difference post-intervention in the establishment of sociomathematical norms, in particular those of mathematical difference, efficiency, elegance, and sophistication, and what constitutes a mathematical explanation and justification. This implies that the classifications in the IMPaCT Taxonomy supported teachers in moving from questions which established social norms, for example:

> **Teacher P:** What could you do instead?
>
> **Teacher Q:** Could you do it a different way?

to questions which established sociomathematical norms, for example:

> **Teacher P:** Is that the same as the other suggestion?
>
> **Teacher P:** Why do you think that one and not that one?
>
> **Teacher Q:** Did you need to do that?
>
> **Teacher Q:** Is that your most efficient method? What would be a really efficient method?

These questions mirror those listed by Whitenack and Yackel (2002) which they believe, after sociomathematical norms are established, learners will begin to ask themselves when problem-solving in mathematics:

> Why might I use one approach over another? What information might I use to help me solve this problem? Can I solve the problem in more than one way? Are some approaches 'easier' or more efficient? (p.526)

As discussed in Chapter 7, such autonomy was evidenced post-intervention, where the learners started to ask deeper questions themselves:

> What's the difference between methods?
>
> How can I tell which to use and when?

What's the easiest way to do this?

The last question was actually answered by another student rather than the teacher, once again evidencing that the focus had moved away from the teacher explaining to the learners engaging in rich mathematical discourse.

The questions these students were asking are also very similar to those suggested by Whitenack and Yackel (2002) above, and suggest that the deeper questions the teachers were asking post intervention contributed to this learner autonomy (Holster, 2006), as the learners were asking more questions themselves to understand the mathematics rather than just asking what they should do. In the post-intervention interview, Teacher P noted that the learners in the participant class had started to ask deeper questions themselves, which coupled with these observed questions in the post-intervention lesson observations, gives some degree of triangulation to the data collection methods used in this thesis as there is an agreement in the findings. Furthermore, the findings agree with the quote from Black et al. (2006) in the literature review, that effective questioning is essential, so learners "can ask questions of each other and the focus can move from the teacher to the pupils" (p.128).

Pre-intervention, Teacher S described what makes a question easier or more difficult to the class without them having the time to formulate the distinction for themselves. The use of small group discussion here could give learners the opportunity to discuss for themselves what makes a problem simpler or more complicated without direction from the teacher first.

Andrews et al. (2005) found that mathematical efficiency was not given sufficient focus in the lessons they observed as part of their research. This was also the case in both the pilot study for this research and the baseline observations, however, after working with the IMPaCT Taxonomy, three of the four participant teachers supported learners to consider mathematical efficiency through their questioning. This evidences that, after working with the IMPaCT Taxonomy, the participant teachers supported learners to consider the advantages and disadvantages of using an alternative method rather than just asking for a different method as Yackel and Cobb (1996) found to be the case in their research. Yackel and Cobb (1996) found it less common for teachers to ask if learners had a more efficient or more sophisticated method. After working with the IMPaCT Taxonomy, teachers started to put less emphasis on accepting *what* learners did as an acceptable mathematical explanation, instead putting more emphasis on the *how* and *why* and indeed, by comparing approaches in this way, established the sociomathematical norms of efficiency and sophistication which Yackel and Cobb (1996) found lacking in their observations of teachers.

One lesson, post-intervention, evidenced the teacher asking learners to explain the thinking of others:

> **Teacher Q:** What about this one from Student G and Student H? What are they trying to do?

This was done in conjunction with the use of mini-whiteboards and the teacher used the learners' multiple representations to discuss mathematical difference, making connections with their prior learning. Another example of use mathematical difference in reflective questioning was used by Teacher P in her lesson on solving quadratics:

> **Teacher P:** "What else is different?"

This deeper reflective thinking should in turn lead to more derivational thinking on behalf of the learners, as the deeper reflective classification in the IMPaCT Taxonomy deals with associating ideas which could be considered the foundation to adapting procedures, hence derivational thought.

In terms of the sociomathematical norm of the acceptability of making mistakes in mathematics, the teachers used learners' mistakes to engage learners in the mathematics, as deemed important by Kazemi and Stipek (2001), particularly in terms of using errors made by learners formatively and creating an environment where learners see the value of learning from mistakes. For example, Teacher Q's emphasis in her lesson on simultaneous equations, where she stressed to learners several times that she was less interested in the answer than their way of working. Establishing this sociomathematical norm was also found to be important in the research of both Herschkowitz and Schwarz (1999) and Whitenack and Yackel (2002), in that engaging in mathematical activity can be as useful to learning as just getting the correct answer.

Of course, I have not measured the impact of developing autonomy on learner attainment, which is a potential area for further research; however, as Morgan and Saxton (2006) described Bloom's Taxonomy, this research was designed to "see the kind of thinking we can set into action through questions" (p. 19). Furthermore, what we have seen agrees with Mason (2014) that, by varying the types of question teachers pose, learners can begin to pose questions themselves.

The high proportion of surface level reasoning questioning pre-intervention agrees with Orrill's (2013) findings that few questions posed by teachers required learners to justify their thinking. Post-intervention, the percentage of deeper reasoning questions rose from 25.3% to 51.7%, which was shown to be statistically significant and I was able to reject the null hypothesis that this difference could have happened by chance ($p<0.01$). This strongly suggests that through engaging with the IMPaCT Taxonomy, teachers can begin to ask more questions which require deeper thinking.

Teachers' Understanding of the Role of Questioning

Does the IMPaCT Taxonomy affect mathematics teachers' understanding of how their questioning impacts on their learners' mathematical thinking?

As discussed in the literature review, many educational researchers place value on the learner developing autonomy (e.g. Cobb & Yackel, 1996; Holster, 2006). The responses from the pre- and post-intervention interviews suggest that as a result of taking part in working with the IMPaCT Taxonomy, teachers improved their understanding of how their questioning impacted on learners' mathematical thinking and provided some anecdotal evidence of how learners were either asking deeper questions themselves, as with Teacher P, or were more able to solve problems without being given procedural methods by the teacher, as with Teacher R.

As mentioned in Chapter 7, the teachers were critical of Bloom's Taxonomy, for example:

> **Teacher P:** What does it mean to understand something in maths? Just because you can solve an equation, doesn't necessarily mean you understand equations.

This agrees with Watson (2007) on how Bloom's Taxonomy plays down understanding in mathematics. Conversely, the post-intervention interviews found that teachers liked the surface and deeper approach, as developed by Smith et al. (1996) and Marton and Saljo (1976a; 1976b) for classifying questioning. Pre-intervention, Teacher R seemed to have a good understanding of how the depth of a question was potentially more important than whether it was open or closed (see Chapter 5), agreeing with Watson's (2003) criticism of the simplicity of open and closed classifications of questioning in mathematics. However in the post-intervention interviews he referred to open and closed questions more than surface and deeper.

Although it was stressed in the training on working with the IMPaCT Taxonomy with the participant teachers that there is a place for procedural questioning in mathematics, as discussed by Fan and Bokhove (2014), Teacher R suggested in his post-intervention interview that he believes all surface questioning has a negative impact on learners' thinking:

> **Teacher R:** Sometimes the surface ones would still crop up […] every time I asked a closed [surface] one, I was always kicking myself.

As stated in the literature review, process can play an important role in supporting the development of conceptual understanding (Sfard, 1991; Gray & Tall, 1992; Dubinsky & McDonald, 2001), particularly when focus is given to the structure of a procedure (Fan & Bokhove, 2014). It is therefore important that when using the IMPaCT Taxonomy, teachers do not deem all surface questioning as negative, the criticism of the surface-deeper perspective made by Howie and Bagnall (2013).

In the pre-intervention interviews, the themes emerged that the participant teachers felt that they either did not have the time to plan for questioning or that planning for questioning was more something that trainee and new teachers needed to do. Post-intervention, teachers appeared to understand the importance of planning for questioning regardless of the level of experience and found the IMPaCT Taxonomy a straightforward way of achieving such planning in limited time.

Limitations of the Research

There were several factors which caused problems for collecting the data in a timely manner. Despite parental consent letters first being sent home in February 2015, due to a lack of response an opt-out letter was sent instead. This needed to allow parents time to respond if necessary and so 27th April was given as the time to reply. At this point getting in to observe or video lessons was difficult as teachers were under a lot of pressure with examination preparation and then Year 10 had their own examination period. In the end a total of six lessons from two teachers were observed or videoed in the second half of the summer term and the remaining 9 lessons from the three remaining classes and two teachers were recorded in the first half of the Autumn term.

The change from using IRIS in the baseline observations to an alternate makeshift recording technique in subsequent lessons is likely to have had some impact on the results, however this is difficult to measure. In some lessons, where the background noise was particularly loud, some questions posed by the teacher were inaudible, despite wearing a microphone, as the video recorder itself was capturing the audio alongside the microphone and speaker. Therefore, if the learners close to the video camera were particularly loud, the teacher could not be heard even though the speaker linked to their microphone was next to the video camera.

As discussed in the pilot study in Chapter 2, as the sole researcher on this thesis, I observed the lessons, interviewed the teachers and analysed the results. This of course has the potential for bias in the findings; as I developed the IMPaCT Taxonomy I could be looking for reasons to support it rather than looking at the results objectively. To minimise this threat to validity and objectivity several steps were taken. Firstly a lesson was coded by another mathematics teacher from another school both in the pre-intervention and post-intervention observations to check that I wasn't using the IMPaCT Taxonomy differently before and after the intervention to show more progress than was actually evidenced. Similarly with the interviews, the same teacher looked through my coding and grouping into themes to check that I was not just picking out comments which supported the IMPaCT Taxonomy.

The interim monitoring observation for Teacher R undoubtedly affected the results for the empirical element of this thesis. From the video recording it could be seen that the questioning deteriorated following a significant proportion of the class returning from an examination during the lesson. As a researcher I could have justified removing this lesson from the sample due to the disruption to the lesson and given Teacher R the opportunity to record another lesson, however I made the decision not to do this on the grounds that it could be perceived as removing data which did not support the use of the IMPaCT Taxonomy.

As in the pilot study, it was found that the depth of questioning is easier to categorise than the type of questioning, evidenced by both the inter-observer reliability testing and the internal consistency testing where the lesson was coded similarly in terms of depth of questioning on the two occasions, but less so with the variety of questioning.

To address internal validity, more than one lesson for each teacher was observed before and after intervention and the questioning averaged over the number of lessons observed. This should have ensured that the data collected was more valid for analysing the questioning in general for each teacher. The only occasion where this was not possible was post-intervention with the lower attaining class with Teacher S, so although the interim monitoring lesson suggests that there was an increase in the depth and variety of questioning with this class post-intervention, there is insufficient evidence to suggest that this would be the case in any lesson observed with this class post-intervention.

9. Conclusions

This thesis set out to investigate the factors which affect the questioning posed by mathematics teachers, and whether working with the IMPaCT Taxonomy could affect the type and depth of questioning posed by mathematics teachers. This chapter summarises the findings from the literature review and the empirical element of this research with respect to each of the research questions. The implications arising for future research are also considered.

Research Question 1

What factors affect the type and depth of questioning used by mathematics teachers?

While the mathematical topic was found to have little impact on the type and depth of questioning in the pilot study, a more in-depth analysis in the main study suggests that the topic does have an impact, but only within the analysis for each teacher. It is the teachers themselves who have the biggest impact on the questioning, regardless of topic.

The stage of the lesson also affected the type and depth of questioning, although again this was different for individual teachers, suggesting once again that the teacher has the overriding influence. This appears to be down to the experience of the teacher, although it has been shown that working with the IMPaCT Taxonomy can close the gap between more and less experienced teachers.

The use of AfL techniques had an impact on the depth and type of questioning unless they were only used as a quick-fire method. In order for a technique such as mini-whiteboards to increase the depth and variety of questioning, they need to be used to generate discussion amongst the learners.

The only factor which had as much influence on the type and depth of questioning as the teacher was the level of attainment of the class, as evidenced by Teacher S following the profile of a more experienced teacher with his higher attaining class, but closer to the profile of a less experienced teacher for the lower attaining class. Although Teacher S still made significant improvements to the depth and variety of questioning with his lower attaining group in the interim monitoring observations, it was less significant than the difference made in his higher attaining group. This suggests that while the intervention on the IMPaCT Taxonomy can close the gap in the depth and variety of questioning between teachers, it may not close the gap between the depth of questioning used with learners of different attainment.

Research Question 2

Does working with the IMPaCT Taxonomy affect the type and depth of questioning used in mathematics lessons?

The comparison of the analysis of the three lessons pre-intervention with the three lessons post-intervention, suggests that working with the IMPaCT Taxonomy has a significant impact on the type and depth of questioning used by teachers in mathematics lessons, with most of these differences in the type and depth of questioning being very unlikely to be attributable to chance.

There was, however, variation on the impact of the training for individual teachers, as shown by the smaller amount of progress made by Teacher R compared to the other teachers, even with the interim monitoring observation removed from his analysis. The additional CPD given to Teacher R following the interim monitoring observation does appear to have made a difference, which, although it is not as statistically significant as the differences for other teachers ($p<0.05$ as opposed to $p<0.001$ in most cases), the impact of the additional intervention is still statistically significant, in that although the null hypothesis could not be rejected for the interim monitoring observation, it was rejected for the post-intervention observations, even with the interim monitoring observation included in the results. Nevertheless, this indicates that different teachers require different levels of support to develop their understanding of the IMPaCT Taxonomy and the concept of deep and surface questioning. Teacher R's interchangeable use of the terms closed and surface questioning in his post-intervention interview highlights this need for further clarification of the definition of the terms in the IMPaCT Taxonomy.

The interviews indicate that, on the whole, the participant teachers found the IMPaCT Taxonomy straightforward to use and would recommend that other teachers and trainee teachers use it to support their questioning. The ease with which the participant teachers claimed to find using the IMPaCT Taxonomy with minimal support, overcomes the issue of the time commitment associated with developing questioning through lesson study (Olson et al., 2011), especially as time constraints for planning for questioning was a theme in the pre-intervention interviews, but was not cited as an issue for the participant teachers post-intervention. As stated above, however, despite the teachers' perception that the IMPaCT Taxonomy was fairly straightforward to implement, the additional intervention required by Teacher R and the differing impact of the intervention on the two classes of Teacher S, suggest that some teachers may need more time and support than others to work with the IMPaCT Taxonomy effectively.

Research Question 3

Does the IMPaCT Taxonomy affect mathematics teachers' understanding of how their questioning impacts on their learners' mathematical thinking?

Pre-intervention, the teachers' perspective was that planning for questioning was for trainee teachers and newly qualified teachers, however working with the IMPaCT Taxonomy and seeing the impact on their own questioning through the analysis sheets provided during the intervention (see Appendix 15), teachers developed an understanding of the importance of their questioning on their learners' mathematical thinking. There is, however, the possibility that this could be attributable to the feedback the teachers received as part of the intervention.

The comparisons between the themes which emerged in the pre- and post-intervention teacher interviews evidence that the participant teachers changed their perspectives on the importance of planning for questioning as a result of this action research and developed an understanding of how their questioning can develop their learners' thinking and indeed their own questioning. There was also evidence that teachers developed an understanding of how their questioning can act as a model for learners to develop their own questioning. However, despite the intervention the participant teachers experienced, teacher perceptions in the post-intervention interviews suggest that deeper questioning is still considered more important for higher attaining learners.

Areas Requiring Further Research

There are two main areas which require further research. Firstly, how can we close the gap between the depth of questioning experienced by higher and lower attaining classes? Watson et al. (2003) found lower attaining learners benefit from opportunities for deep mathematical thinking, however Teacher S had the same intervention to apply to both types of class and yet a statistically more significant change in the depth of his questioning was found for the higher attaining class. What can be done to change teachers' perceptions that higher attaining learners require deeper questioning than lower attaining learners?

Secondly, when analysing the participant teachers' questioning both pre- and post-intervention, the following question arose: What is more important, the proportion of deeper questioning used in a lesson or the number of deeper questions asked? A similar question could be raised with respect to proportion and numbers of each question type. Are deeper questions watered down in a lesson when too many factual and procedural questions asked? For example Teacher P asked the highest proportion of derivational questions in the baseline observations, but the actual numbers of questions posed in this category was lower than Teacher Q. Similarly, the proportion of derivational questions had the largest percentage increase from pre- to post-intervention, however it remained the lowest, alongside factual, in terms of the number of questions asked post-intervention. This is an interesting question which in a sense is unanswerable, as the effects of both proportion and number of deep questions can only be understood in the context of the lesson and thus it would not be possible to collect comparable data to address this conundrum.

Some Final Thoughts

The reason I first embarked on research in questioning in mathematics four years ago was due to my own questioning in lessons being identified as an area for development by my head teacher at the time. As I write this concluding paragraph for this EdD thesis, I have beside me the feedback from a lesson observation carried out on me this week, where my questioning was identified as a strength, due to being 'rich and thoughtful'. The action research was planned to identify if working with the IMPaCT Taxonomy could affect the type and depth of questioning employed by mathematics teachers in my school as part of my role as Lead Practitioner for Mathematics; the findings suggest that it can, and I am delighted that I have improved my own practice in questioning along the way.

Black et al. (2006) wrote of the need for teachers to develop effective questioning strategies in order for the focus to shift away from the teachers and towards the learners. This research has shown that the IMPaCT

Taxonomy can support this process. Furthermore, Orrill (2013) stated that further research was required to “identify and characterize more effective questioning strategies” which are accessible to mathematics teachers. This research has shown that the IMPaCT Taxonomy could fill this gap through the impact it has had on developing teachers’ understanding of the importance of questioning and increasing the depth and variety of their questioning in lessons. It has shown that while some teachers may need some additional support to understand the IMPaCT Taxonomy, it is on the whole an accessible and visual tool to improve the depth and variety of questioning in the mathematics classroom.

References

Anderson, L. W. and Krathwohl, D. R. (Eds.), Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. and Wittrock, M. C. (2001) *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives*, New York, Allyn Bacon Longman.

Anderson, L. W. and Sosniak, L. A. (Eds.) (1994) *Bloom's taxonomy: A forty-year perspective. Ninety-third Yearbook of the National Society for the Study of Education*, Chicago, University of Chicago Press.

Andrews, P., Hatch, G. and Sayers, J. (2005) 'What do teachers of mathematics teach? An initial episodic analysis of four European traditions', in Hewitt, D. and Noyes, A. (Eds.) *Proceedings of the Sixth British Congress of Mathematics Education* held at the University of Warwick, pp. 9-16.

Ball, D. L., Lubienski, S. T. and Mewborn, D. S. (2001) 'Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge', in Richardson, V. (Ed.) *Handbook of research on teaching*, Fourth Edition, pp. 433-456.

Barwell, R. (2013) 'Discursive psychology as an alternative perspective on mathematics teacher knowledge', *ZDM*, Vol. 45, No. 4, pp. 595-606.

Bashir, M., Afzal, M. T. and Azeem, M. (2008) 'Reliability and validity of qualitative and operational research paradigm', *Pakistan Journal of Statistics and Operation Research*, Vol. 4, No. 1, pp. 35-45.

Bauersfeld, H., Krummheuer, G. and Voigt, J. (1988) 'Interactional theory of learning and teaching mathematics and related microethnographical studies', in Steiner, H. G. and Vermandel, A. (Eds.) *Foundations and methodology of the discipline of mathematics education*, Antwerp, Proceedings of the Theory of Mathematics Education Conference, pp. 174-188.

Bauersfeld, H. (1993) 'Teachers' pre and in-service education for mathematics teaching', *Seminaire sur la Representation*, No. 78, CIRADE, Universite du Quebec a Montreal, Canada.

BERA (2011) *Ethical Guidelines for Educational Research,* London, British Educational Research Association.

Bereiter, C. and Scardamalia, M. (1999) ‘Beyond Bloom's Taxonomy: Rethinking Knowledge for the Knowledge Age’, in Hargreaves, A., Libermann, A., Fullan, M. and Hopkins, D. (Eds.) *The International Handbook of Educational Change*, Dordrecht, Kluwer Academic Publishers, pp. 675–692.

Biggs, J. & Collis, K. F. (1980) SOLO taxonomy, *Education News*, Vol. 17, No.5, pp. 19-23.

Biggs, J. B. and Tang, C. (2007) *Teaching for quality learning at university*, Third Edition, Berkshire, Open University Press.

Black, P. and Wiliam, D. (1998) *Inside the black box: Raising standards through classroom assessment*, London, nferNelson.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003) *Assessment for learning - putting it into practice*, Maidenhead, Open University Press.

Black, P., McCormick, R., James, M. and Pedder, D. (2006) ‘Learning How to Learn and Assessment for Learning: A Theoretical Inquiry’, *Research Papers in Education*, Vol. 21, No.2, pp. 119-132.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. and Krathwohl, D. R. (1956) *Taxonomy of educational objectives: Handbook I: Cognitive domain*, New York, David McKay.

Britten, N. (1995) ‘Qualitative research: qualitative interviews in medical research’, *Bmj*, Vol. 311, pp. 251-253.

Burton, N., Brundrett, M. and Jones, M. (2014) *Doing your education research project*, London, SAGE.

Campbell, P. F., Nishio, M., Smith, T. M., Clark, L. M., Conant, D. L., Rust, A. H., DePiper, J. N., Frank, T. J., Griffin, M. J. and Choi, Y. (2014) ‘The relationship between teachers' mathematical content and pedagogical knowledge, teachers' perceptions, and student achievement’, *Journal for Research in Mathematics Education*, Vol. 45, No. 4, pp.419-459.

Chan, C. (2010) ‘Assessment: Blooms’ Taxonomy’, *Assessment Resources @HKU*, University of Hong Kong. Available at: http://ar.cetl.hku.hk/large_class.htm [Accessed: 24 October 2015]

Chapman, O. (2013) ‘Mathematical-task knowledge for teaching’, *Journal of Mathematics Teacher Education*, Vol. 16, No. 1, pp. 1-6.

Chick, H. (1998) 'Cognition in the formal modes: Research mathematics and the SOLO taxonomy', *Mathematics Education Research Journal*, Vol. 10, No. 2, pp. 4-26.

Clough, P. and Nutbrown, C. (2012) *A student's guide to methodology*, London, SAGE.

Cobb, P., & Yackel, E. (1996) 'Constructivist, emergent, and sociocultural perspectives in the context of developmental research', *Educational psychologist*, Vol. *31*, No. 3-4, pp. 175-190.

Coghlan, D. and Brannick, T. (2005) *Doing action research in your own organization*, London, SAGE.

Cohen, L., Manion, L. and Morrison, K. (2007) *Research methods in education*, New York, Routledge.

Creswell, J. W. (2003) *Research design: Qualitative, quantitative, and mixed methods approaches*, Second Edition, Thousand Oaks, CA, SAGE.

Creswell, J. W., Shope, R., Plano Clark, V. L. and Green, D. O. (2006) 'How interpretive qualitative research extends mixed methods research', *Research in the Schools*, Vol. 13, No. 1, pp. 1-11.

Delamont, S. (2002) *Fieldwork in Educational Settings: Methods, pitfalls and perspectives*, Abingdon, Routledge.

Denscombe, M. (2007) *The Good Research Guide for small-scale social research projects*, Third Edition, Maidenhead, Oxford University Press.

Denzin, N. K. (1989) *The Research Act,* Third Edition, Englewood Cliffs, NJ, Prentice Hall.

Denzin, N. and Lincoln, Y. (Eds.) (2005) *Handbook of Qualitative Research*, Third Edition, London, SAGE.

DfES (2004a) 'Unit 7: Questioning', *Pedagogy and Practice: Teaching and Learning in Secondary Schools,* Key Stage 3 National Strategy, Crown Copyright.

DfES (2004b) *Assessment for Learning (AfL): Whole School Training Materials,* Key Stage 3 National Strategy, Crown Copyright.

DfES (2007) *Teaching and learning functional mathematics*, 14-19 education and skills, Crown Copyright. Available at: http://archive.excellencegateway.org.uk/pdf/TandLMathematicsHT281107.pdf [Accessed 26 September 2013]

Drageset, O. G. (2014) 'Redirecting, progressing, and focusing actions - a framework for describing how teachers use students' comments to work with mathematics', *Educational Studies in Mathematics*, Vol. 85, No.2, pp. 281-304.

Dubinsky, E. and McDonald, M.A. (2001) 'APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research', in Hoton, D. (Ed.) *The Teaching and Learning of Mathematics at University Level, An ICMI Study*, New York, Kluwer Academic Publishers, pp. 275-282.

Edexcel (2012) *Examiners' Report November 2012 GCSE Mathematics (Linear) 1MA0 Foundation (Non-Calculator) Paper 1*, Harlow, Pearson. Available at: http://qualifications.pearson.com/content/dam/pdf/GCSE/Mathematics%20A/2010/Exam%20materials/1MA0_1F_pef_20131210.pdf [Accessed 11 September 2016)

Elliot, J. (1991) *Action research for educational change*, Milton Keynes, Open University Press.

Fan, L. and Bokhove, C. (2014) 'Rethinking the role of algorithms in school mathematics: a conceptual model with focus on cognitive development', *ZDM Mathematics Education*, Vol. 46, No. 3, pp. 481-492.

Fan, L., Ngai-Ying, W., Jinfa, C. and Shiqi, L. (Eds.) (2004) *How Chinese learn mathematics: Perspectives from insiders*, Singapore, World Scientific.

Fraivillig, J., Murphy, L. and Fuson, K. (1999) 'Advancing children's mathematical thinking in everyday mathematics classrooms', *Journal for Research in Mathematics Education,* Vol. 30, No. 2, pp. 148-170.

Furst, E. J. (1994) 'Bloom's taxonomy: Philosophical and educational issues' in Anderson, L. W. and Sosniak, L. A. (Eds.) *Bloom's taxonomy: A forty-year perspective. Ninety-third Yearbook of the National Society for the Study of Education*, Chicago, University of Chicago Press, pp. 28-40.

Gall, M. D. (1970) 'The use of questions in teaching', *Review of Educational Research*, Vol. 40, No. 5, pp. 707-721.

Gerson, H. and Bateman E. (2011) ‘Authority in the Negotiation of Sociomathematical Norms’, in Brown, S., Larsen, S., Marrongelle, K. and Oehrtman, M. (Eds.) *Proceedings of the 14th Annual Conference on Research in Undergraduate Mathematics Education*, Vol. 1, Portland, Oregon, p. 115-127.

Gorard, S. and Taylor, C. (2004) *Combining methods in educational and social research*, Maidenhead, McGraw-Hill International.

Greene, J. (2005) ‘Combining Qualitative and Quantitative Methods in Social Inquiry, Key Concepts’, in Somekh, B. and Lewin, C. (Eds.) *Research Methods in the Social Sciences*, London, SAGE, pp. 274-277.

Gray, E. and Tall, D. (1991) ‘Duality, Ambiguity and Flexibility in Successful Mathematical Thinking’, *Proceedings of PME 15*, Vol. 2, pp. 72–79.

Gray, E. and Tall, D. (1992) ‘Success and Failure in Mathematics: The Flexible Meaning of Symbols as Process and Concept’, *Mathematics Teaching 142*, pp. 6-10.

Griffiths, M. (1998) *Educational research for social justice: Getting off the fence*, Buckingham, Open University Press.

Gunasekare, T. P. (2015) ‘Mixed Research Method as the Third Research Paradigm: A Literature Review’, *International Journal of Science and Research (IJSR)*, Vol. 4, No. 8, pp. 361-367.

GURO21 (2012) Gearing Up Responsible and Outstanding Teachers in Southeast Asia for the 21st Century, Course 1, Module 2, SEAMEO INNOTECH. Available at: http://iflex.innotech.org/GURO21/Module2/l2_28.html [Accessed 9 July 2017]

Hargreaves, D. H. (1984) ‘Teachers’ questions: open, closed and half-open’, *Educational Research*, Vol. 26, No. 1, pp. 46-51.

Hartas, D. (Ed) (2010) *Educational research and inquiry: Qualitative and quantitative approaches*, London, Bloomsbury.

Hershkowitz, R. and Schwarz, B. (1999) ‘The emergent perspective in rich learning environments: Some roles of tools and activities in the construction of sociomathematical norms’, *Educational Studies in Mathematics*, Vol. 38, pp. 149-166.

Hodgen, J. and Wiliam, D. (2006) *Mathematics inside the black box: Assessment for learning in the mathematics classroom*, London, Granada Learning.

Hoepfl, M. C. (1997) 'Choosing qualitative research: A primer for technology education researchers', *Journal of Technology Education*, Vol. 9, No. 1. Available at: http://scholar.lib.vt.edu/ejournals/JTE/v9n1/hoepfl.html [Accessed 24 September 2016]

Hogg, R. V. and Tanis, E. (2014) *Probability and Statistical Inference*, Harlow, Pearson.

Hollingsworth, P. M. (1982) 'Questioning: The heart of teaching', *The Clearing House: A Journal of Educational Strategies, Issues and Ideas,* Vol. 55, No. 8, pp. 350-352.

Holster, T. (2006) 'Purposeful questioning in mathematics: A guiding framework', *ACE Papers: Science, Mathematics and Technology Education*, Issue 17, pp. 1-13. Available at: https://researchspace.auckland.ac.nz/handle/2292/25160 [Accessed 9 June 2016]

Howe, K. R. (1988) 'Against the quantitative-qualitative incompatibility thesis or dogmas die hard', *Educational researcher,* Vol. 17, No. 8, pp. 10-16.

Howie, P. and Bagnall, R. (2013) 'A critique of the deep and surface approaches to learning model', *Teaching in Higher Education*, Vol. 18, No. 4, pp. 389-400.

Jarvis, M. (2005) *The psychology of effective learning and teaching*, Cheltenham, Nelson Thornes.

Johnson, R. B. (1997) 'Examining the validity structure of qualitative research', *Education*, Vol. 118, No. 2, pp. 282-292.

Johnson, R. B. and Onwuegbuzie, A. J. (2004) 'Mixed methods research: A research paradigm whose time has come', *Educational Researcher*, Vol. 33, No. 7, pp. 14-26.

Kajornboon, A. B. (2005) 'Using interviews as research instruments', *E-Journal for Research Teachers*, Vol. 2, No. 1. Available at: http://hsmi.psu.ac.th/upload/forum/Annabelinterviewguide.pdf [Accessed 26 August 2016]

Kawanaka, T. and Stigler, J. W. (2000) 'Teachers' use of questions in eighth-grade mathematics classrooms in Germany, Japan and the United States', *Mathematical Thinking and Learning,* Vol. 1, No. 4, pp. 255-278.

Kazemi, E. and Stipek, D. (2001) 'Promoting conceptual understanding in four upper elementary mathematics classrooms', *The Elementary School Journal*, Vol. 102, No. 1, pp. 59-80.

Kerlinger, F. N. (1970) *Foundations of behavioural research*, New York, Holt, Rinehart & Winston.

Knapp, A., Bomer, M. and Moore, C. (2011) 'Lesson study as a learning environment for coaches of mathematics teachers', in Hart, L. C., Alston, A. and Murata, A. (Eds.) *Lesson study research and practice in mathematics education*, New York, Springer, pp. 153-164.

Krathwohl, D. R. (2002) 'A revision of Bloom's taxonomy: An overview', *Theory into practice*, Vol. 41, No. 4, pp. 212-218.

Levinson, E., Tirosh, D. and Tsamir, P. (2009) 'Students' perceived sociomathematical norms: the missing paradigm', *The Journal of Mathematical Behavior*, Vol. 28, pp. 171-187.

Lindsay, G. (2013) 'The benefits of combined (mixed) methods research', *Social Work and Social Sciences Review*, Vol. 16, No. 2, pp. 76-87.

Mackenzie, N. and Knipe, S. (2006) 'Research dilemmas: Paradigms, methods and methodology', *Issues in educational research*, Vol. 16, No. 2, pp. 193-205.

Martino, A. M. and Maher, C. A. (1999) 'Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us', *The Journal of Mathematical Behavior*, Vol. 18, No. 1, pp. 53-78.

Marton, F. and Saljo, R. (1976a) 'On qualitative differences in learning: I - Outcome and process', *British Journal of Educational Psychology,* Vol. 46, No. 1, pp. 4-11.

Marton, F. and Saljo, R. (1976b) 'On qualitative differences in learning: I – Outcome as a function of the learner's conception of the task', *British Journal of Educational Psychology* Vol. 46, No. 2, pp. 115-127.

Mason, J. (2014) 'Questioning in mathematics education', in Lerman, S. (Ed.) *Encyclopedia of mathematics education*, Netherlands, Springer, pp. 513-519.

Mason, J. (2000) 'Asking mathematical questions mathematically', *International Journal of Mathematical Education in Science and Technology*, Vol. 31, No. 1, pp. 97-111.

Mason, J., Stephens, M. and Watson, A. (2009) 'Appreciating mathematical structure for all', *Mathematics Education Research Journal*, Vol. 21, pp. 10-32.

Maxwell, J. A. and Loomis, D. M. (2003) 'Mixed methods design: An alternative approach', *Handbook of mixed methods in social and behavioral research*, Vol. 1, pp. 241-272.

Michener, E. R. (1978) 'Understanding Understanding Mathematics', *Cognitive Science 2*, pp. 361-383.

Miller, J. and Glassner, B. (2004) 'The "inside" and the "outside". Finding realities in interviews', in Silverman, D (Ed.) *Quantitative Research*, London, SAGE.

Miles, M. B. and Huberman, A. M. (1994) *Qualitative Data Analysis: An Expanded Source Book*, Second Edition, Thousand Oaks, SAGE.

Morgan, N. and Saxton, J. (2006) *Asking better questions*, Markham, Ontario, Pembroke Publishers.

Moyer, P. S. and Milewicz, E. (2002) 'Learning to question: Categories of questioning used by preservice teachers during diagnostic mathematics interviews', *Journal of Mathematics Teacher Education*, Vol. 5, No. 4, pp. 293-315.

Muijs, D. (2010) *Doing quantitative research in education with SPSS*, London, SAGE.

Muijs, D. and Reynolds, D. (2011) *Effective teaching: Evidence and practice*, London, SAGE.

OCR (2011) *GCSE Mathematics A, Report to Centres November 2011*, Nottingham, OCR. Available at: http://www.ocr.org.uk/Images/86904-examiners-reports-november.pdf [Accessed 11 September 2016]

Ofsted (2008) *Assessment for learning: the impact of National Strategy support*, Manchester, Crown Copyright. Available at: http://www.ofsted.gov.uk/resources/assessment-for-learning-impact-of-national-strategy-support [Accessed 27 May 2013]

Ofsted (2012) *Mathematics: made to measure*, Manchester, Crown Copyright. Available at: www.ofsted.gov.uk/resources/110159 [Accessed 25 May 2013]

Olson, J. C., White, P. and Sparrow, L. (2011) 'Influence of lesson study on teachers' mathematics pedagogy', in Hart, L. C., Alston, A. and Murata, A. (Eds.) *Lesson study research and practice in mathematics education*, New York, Springer, pp. 39-57.

Opie, C. (2004) *Doing educational research*, London, SAGE.

Ormell, C. P. (1974) 'Bloom's taxonomy and the objectives of education', *Educational Research*, Vol. 17, No. 1, pp. 3-18.

Orrill, C. H. (2013) 'Connection Making: Capitalizing on the Affordances of Dynamic Representations Through Mathematically Relevant Questioning', in Hegedus, S. J. and Roschelle, J. (Eds.) *The SimCalc Vision and Contributions*, Netherlands, Springer, pp. 285-298.

Pegg, J. and Tall, D. (2010) 'The Fundamental Cycle of Concept Construction Underlying Various Theoretical Frameworks', in Sriraman, B. and English, L. (Eds.) *Theories of Mathematics Education*, Berlin Heidelberg, Springer-Verlag, pp. 173-192.

Piaget, J. (1973) *To understand is to invent: The Future of education*, New York, Grossman.

Punch, K. (2009) *Introduction to Research Methods in Education*, London, SAGE.

Robson, C. (2002) *Real world research: A resource for social scientists and practitioner researchers*, Oxford, Blackwell.

Rossman, G. B. and Wilson, B. L. (1985) 'Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study', *Evaluation Review*, Vol. 9, pp. 627-643.

Rowe, M. B. (1986) 'Wait time: slowing down may be a way of speeding up!', *Journal of Teacher Education*, Vol. 37, No. 1, pp. 43-50.

Sapon-Shevin, M. and Schniedewind, N. (1991) 'Selling cooperative learning without selling it short', in Brandt, R. S. (Ed.) *Cooperative Learning and the Collaborative School: Readings from Educational Leadership*, Alexandria, Association for Supervision and Curriculum Development.

Seddon, G.M. (1978) 'The properties of Bloom's taxonomy of educational objectives for the cognitive domain', *Review of Educational Research*, Vol. 48, No. 2, pp. 303-323.

Sfard, A. (1991) 'On the Dual Nature of Mathematical Conceptions: Reflections on Processes and Objects as Different Sides of the Same Coin', *Educational Studies in Mathematics*, Vol. 22, No. 1, pp. 1-36.

Silverman, D. (2006) *Qualitative research: Theory, Method and Practice,* London, SAGE.

Simon, M. A. (1995) 'Reconstructing mathematics pedagogy from a constructivist perspective', *Journal for research in mathematics education*, pp. 114-145.

Skemp, R. R. (1976) 'Relational understanding and instrumental understanding', *Mathematics Teaching*, Vol. 77, pp .20-26.

Smith, G., Wood L., Coupland, M., Stephenson, B., Crawford, K. and Ball, G. (1996) 'Constructing mathematical examinations to assess a range of knowledge and skills', *International Journal of Mathematical Education in Science and Technology*, Vol. 27, No.1, pp. 65-77.

Smith, J. K. and Heshusius, L. (1986) 'Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers', *Educational researcher*, Vol. 15, No. 1, pp. 4-12.

Smithers, A. and Robinson, P. (2013) *The Science and Mathematics Teaching Workforce*. Report for The Royal Society, Centre for Education and Employment Research, University of Buckingham. Available at https://royalsociety.org/~/media/education/policy/vision/reports/ev-7-vision-research-report-20140624.pdf [Accessed 27/08/2016]

Stake, R. (2005) 'Qualitative case studies', in Denzin, N. and Lincoln, Y. (Eds.) *Handbook of Qualitative Research*, Third Edition, London, SAGE.

Steele, M. D. (2013) 'Exploring the mathematical knowledge for teaching geometry and measurement through the design and use of rich assessment tasks', *Journal of Mathematics Teacher Education*, Vol. 16, No. 4, pp. 245-268.

Swann, J. (1994) 'Observing and recording talk in educational settings', in Graddol, D., Maybin, J. and Stierer, B. (Eds.) *Researching language and literacy in social context*, Clevedon, Multilingual Matters Ltd, pp. 26-48.

Thornby, J. A., Mawson, K. and Haywood, W. (2016) Tackling the teacher shortage–near-peer mentoring of STEM students for outreach and engagement using trainee teachers. Online infographic available at: http://wrap.warwick.ac.uk/76140/7/WRAP_HEA%20Poster.pdf [Accessed 27 August 2016]

Trekles, A. M. and Sims, R. (2013) 'Designing Instruction for Speed: Qualitative Insights Into Instructional Design for Accelerated Online Graduate Coursework', *Online Journal of Distance Learning Administration*, Vol. XVI, No. IV, University of West Georgia, Distance Education Center. Available at https://www.westga.edu/~distance/ojdla/winter164/trekles_sims164.html [Accessed 24 October 2015]

Vygotsky, L. (1978) 'Interaction between learning and development', *Mind and Society*, pp. 79-91, Harvard University Press, Cambridge. Reprinted in Gauvin, M & Cole, M. (Eds.) (1997) Readings on the development of children, pp. 29-35.

Warner, R. (2016) *Optimizing the Display and Interpretation of Data*, Amsterdam, Elsevier.

Watson, A. and Mason, J. (1998) *Questions and Prompts for Mathematical Thinking*, Derby, ATM.

Watson, A. (2003) 'Opportunities to learn mathematics', *Mathematics education research: Innovation, networking, opportunity*, pp. 29-38.

Watson, A., De Geest, E. & Prestage, S. (2003) *Deep Progress in Mathematics-The Improving Attainment in Mathematics Project*, University of Oxford, Department of Educational Studies.

Watson, A. (2007) 'The nature of participation afforded by tasks, questions and prompts in mathematics classrooms', *Research in Mathematics Education,* Vol. 9, No. 1, pp. 111-126.

Watson, A. and Barton, B. (2011) 'Teaching mathematics as the contextual application of mathematical modes of enquiry', in Rowland, T. and Ruthven, K. (Eds.) *Mathematical knowledge in teaching*, Netherlands, Springer, pp. 65-82.

Whitenack, J. and Yackel, E. (2002) 'Making mathematical arguments in the primary grades: the importance of explaining and justifying ideas', *Teaching Children Mathematics,* Vol. 8, No. 9, pp. 524-529.

Wood, T. (1998) ‘Alternative patterns of communication in mathematics classes: Funneling or focusing?’, in Steinbring, H., Bartolini Bussi, M. G. & Sierpinska, A. (Eds.) *Language and communication in the mathematics classroom,* Reston, Virginia, National council of teachers of mathematics, pp. 167-178.

Wood, T. (1999) ‘Creating a context for argument in mathematics class’, *Journal for research in mathematics education*, Vol. 30, No. 2, pp. 171-191.

Yackel, E. and Cobb, P. (1996) ‘Sociomathematical norms, argumentation, and autonomy in mathematics’, *Journal for Research in Mathematics Education*, Vol. 27, No. 4, pp. 458-477.

Yang, M. (2006) ‘A critical review of research on questioning in education: limitations of its positivistic basis’, *Asia Pacific Education Review*, Vol. 7, No. 2, pp. 195-204.

Zhang, Y. and Wildemuth, B. M. (2009) ‘Qualitative analysis of content’, in Wildemuth, B. M. (Ed.) *Applications of social research methods to questions in information and library science*, Westport, CT, Libraries Unlimited, pp. 308-319.

Appendices

Appendix 1 – Early Version of the IMPaCT Taxonomy

QUESTION TYPE *Adapted from Smith et al. (1996) and Andrews et al. (2005)*	**PROMPTS** *Adapted from Watson's analytical instrument(2007, p.119)*	**FORMATIVE QUESTION STEMS** *From Hodgen & Wiliam (2006)*	**SURFACE APPROACH QUESTION CODING**	**DEEPER APPROACH QUESTION CODING**
Factual	• Name • Recall facts • Give definitions • Define terms		SUR-FAC	DP-FAC
Procedural	• Imitate method • Copy object • Follow routine procedure • Find answer using procedure • Give answer		SUR-PRO	DP-PRO
Structural	• Show me… • Analyse • Compare • Classify • Conjecture • Generalise • Identify variables • Explore variation • Look for patterns • Identify relationships	Tell me about the problem. What do you know about the problem? Can you describe the problem to someone else? What is similar/different. . . ? Do you have a hunch? . . . a conjecture? What would happen if . . . ? Is it always true that . . . ? Have you found all the solutions?	SUR-STR	DP-STR
Reasoning	• Justify • Interpret • Visualise • Explain • Exemplify • Informal induction/deduction	Can you explain/ improve/add to that explanation? How do you know that . . . ? Can you justify . . .?	SUR-REA	DP-REA
Reflective	• Summarise • Express in own words • Evaluate • Consider advantages/ disadvantages	What was easy/difficult about this problem/this mathematics? What have you found out? What advice would you give to someone else about . . . ?	SUR-REF	DP-REF
Derivational	• Prove • Create/Design • Associate ideas • Apply prior knowledge (in new situations) • Adapt procedures • Find answer without known procedure	Have you seen a problem like this before? What mathematics do you think you will use? Can you find a different method? Can you prove that . . . ?	SUR-DER	DP-DER

Appendix 2 - Pilot Study - Learner Questionnaire

The purpose of this questionnaire is to find out about what types of questioning in mathematics lessons help you to really think and take an active role in the lesson. The findings will be used to support a study on the use of questioning in mathematics. Completing the questionnaire is voluntary and the information you are asked to give does not identify you or your teacher. Please fill out this form based on your mathematics lessons this year.

1. What is your gender? *(Please tick)*
   Male ☐ Female ☐

2. Which year group are you in? *(Please tick)*
   Year 7 ☐ Year 8 ☐ Year 9 ☐ Year 10 ☐ Year 11 ☐

3. What is your current level in maths? *(Please write your level or tick 'do not know')*
   Level/Grade _______ Do not know ☐
4. How confident do you feel to answer questions in mathematics lessons? *(Circle the number that represents your view:* 1=Very confident, 2=Somewhat confident, 3=Only Slightly confident, 4=Not at all confident, 5=Do not know)
   1 2 3 4 5

5. How often do you do the following in your maths lessons? *(Tick one box in each row)*

	Every Lesson	Most Lessons	Sometimes	Never	Don't Know
Teacher uses random methods to choose a student to answer (e.g. names from hat)					
Hands up to answer a question					
No hands up (with 'wait time' to think)					
Discuss answer in pairs/groups first					
Use mini-whiteboards to write answers					
Choose from a few answers (e.g. Using fans)					
Teacher writes up selection of student answers on board then class discuss					
Explain an answer you've given					
Explain an answer someone else has given					
Asked if you agree with another student					
Asked questions with lots of possible answers					
Always/Sometimes/Never True Problems					
Spot the odd one out					
Spot the mistake					
Any other questioning technique? If so, please state ___________________________________					

6. Rank the following from 1-5 of the likelihood that you would try to work out an answer to a question (1=most likely; 5=least likely)

Hands up to answer a question	
No hands up and longer 'wait time' to think	
Discuss answer in pairs/groups first	
Use mini-whiteboards to write answers	
Choose from a few answers (e.g. Using fans)	

7. Rank the following from 1-5 of the likelihood that you would volunteer an answer (1=most likely; 5=least likely)

Hands up to answer a question	
No hands up and longer 'wait time' to think	
Discuss answer in pairs/groups first	
Use mini-whiteboards to write answers	
Choose from a few answers (e.g. Using fans)	

8. Which seems most important in answering questions in your mathematics lessons? *(Please tick one box only)*

The final answer ☐ How I got the answer ☐ Both equally important ☐ Don't know ☐

9. Can you think of any examples of questions that have been asked in mathematics in the last week which have really made you think and explain your thinking?

____

____

____

____

Thank you for taking the time to fill out this questionnaire. If you have any questions please send me an email (jdenton@********.org.uk) or find me from Monday to Wednesday at ******** School. **Please return this form to either your form tutor, the admin office or directly to me in the envelope provided by 14 July 2013.**

Mrs J Denton, Advanced Skills Teacher of Mathematics

Appendix 3 – Pilot Study – Revised Learner Questionnaire

The purpose of this questionnaire is to find out about what types of questioning in mathematics lessons help you to really think and take an active role in the lesson. The findings will be used to support a study on the use of questioning in mathematics.

Completing the questionnaire is voluntary and the information you are asked to give does not identify you or your teacher. If you have any questions please send me an email (*****@*******.org.uk) or find me from Monday to Wednesday at ******* School.

Please fill out this form based on your mathematics lessons this year.

1. What is your gender? *(Please tick)*

   Male ☐ Female ☐

2. Which year group are you in? *(Please tick)*

   Year 7 ☐ Year 8 ☐ Year 9 ☐ Year 10 ☐ 'ear 11 ☐

3. What is your current level or grade in mathematics?
(Please write your level or tick 'do not know')

   Level/Grade _______ Don't know ☐

4. How confident do you feel to answer questions in mathematics lessons? *(Circle the number that represents your view:* 1=Very confident, 2=Somewhat confident, 3=Only Slightly confident, 4=Not at all confident, 5=Do not know)

   1 2 3 4 5

5. Which seems most important in answering questions in your mathematics lessons?
(Please tick one box only)

   The final answer ☐ How I got the answer ☐ Both equally important ☐ Don't know ☐

6. How often do you use the following techniques to answer questions in your maths lessons?

(Please tick one box in each row)	Every Lesson	Most Lessons	Sometimes	Never	Don't Know
Teacher uses random methods to choose a student to answer a question (e.g. pulls names from hat)					
Put hands up to answer a question					
No hands up rule (where teacher gives you 'wait time' to think then chooses a student to answer)					
The class is given time to discuss an answer in pairs or groups first					
Use mini-whiteboards to write answers to the teacher's questions					
Voting on a few multiple choice answers (e.g. holding up A, B, C, D fans to vote)					
Teacher writes up selection of student answers on the board then the class discusses them					
Teacher asks you to explain an answer you've given					
Teacher asks you to explain an answer someone else has given					
Teacher asks you if you agree with another student's answer to a question					
Teacher asks questions with lots of possible answers rather than just one answer					
Teacher asks if statements are Always/Sometimes or Never True					
Spot the odd one out from a list of numbers, shapes etc.					
Spot the mistake!					

7. Rank the following from 1-5 of the likelihood that you would try to work out an answer to a question (1=most likely; 2=second most likely, …, 5=least likely)

Hands up to answer a question	
No hands up and longer 'wait time' to think	
Discuss answer in pairs/groups first	
Use mini-whiteboards to write answers	
Voting on a few possible answers (e.g. holding up A, B, C, D fans to vote)	

8. Can you think of any other questioning technique that your maths teacher uses?

____

____

Thank you for taking the time to fill out this questionnaire. **Please return this form to either your form tutor, the admin office or directly to me in the envelope provided by 14 July 2013.**

Appendix 4 - Questioning Technique Coding Table

QUESTIONING TECHNIQUE	CODE
Use random methods to choose a student to answer (e.g. names from hat)	RAN
Hands up	HU
No hands up	NHU
'Wait time'	WT
Discuss answer for a set time in pairs/groups first	DPG
Use mini-whiteboards to write answers	MWB
Generate discussion from mini-whiteboards	MWB-D
Choose from a few answers (e.g. Using voting fans)	VOT
Ask a student to explain their answer	EXP
Ask a student to explain another student's answer	EXPA
Ask if a student agrees with another	AG
Identify the error	ERR
Writing up selection of responses on board then discuss	SRB-D
Odd one out	OOO
Always/Sometimes/Never True (or equivalent)	ASN
Problems with more (or less) than one correct solution	PMA

Appendix 5 – Pilot Study - Teacher Semi-Structured Interview Schedule

"Thank you for letting me join your lesson to look at the questioning you're using with your students. I'd like now to talk about a selection of the questions and techniques you used and what your motives and intentions behind those were."

Question Types
"By *question type* I mean the mathematical thinking you intended for your students by asking that question" (Explain the Question Type coding table to the interviewee)

- "Can give me an example(s) of any questions that you asked which you believe needed a surface approach?" (Clarify this term if required, or make suggestions if required)
- "Which questions that you asked would you say needed a deeper approach?"
- "What proportion of your questions this lesson would you estimate required a deeper approach?"
- "Would this be a similar proportion in a 'typical' lesson of yours?"
- "You asked the following question…what type of mathematical thinking did you hope your students would experience from this?"
  Prompts: "Did the students react/respond as expected?"
  "Which category from the coding table do you think it is?"
  (Repeat for other key questions/interesting responses from the lesson)

Questioning Techniques
"Q*uestioning technique* refers to the strategies that you put in place for your students to think about and respond to your questions" (Explain the Questioning Technique coding table with the interviewee).

- "You used the following technique….Why did you choose this technique for this question(s)?"
  Prompts: "In your opinion, what are the benefits of this technique?"
  "Are there any drawbacks to this technique?"
- "Can you think of any other technique which might support a deeper approach from the students?"
- "How do you ensure your questioning assesses what the students know and understand?"
  Prompts: "Do you think this could be improved upon?"
  If yes: "How?"

Thank you very much for your time. Just to remind you that your comments are confidential and you will not be identified in the write up, but some of the dialogue we have had now may be anonymously quoted.

Appendix 6 - Teacher Semi-Structured Interview Schedule (Pre-Intervention)

"Thank you for letting me observe/video your lessons to look at the questioning you're using with your students. I'd like now to talk about your approach to planning for questioning and the techniques you use to support questioning. Just to remind you that your comments are confidential and you will not be identified in the write up, but some of the dialogue we have may be anonymously quoted. You also have the right to withdraw from taking part at any time. Could you please sign this form to show you have understood" (Give participant Teacher Consent Form to sign)

Do you spend time planning the specific questions you will ask in lessons?

- How do you do this? (e.g. written down, in your head, past experience)
- Do you plan for follow up questions depending on students' responses? If so, how?

The following diagram is Bloom's Taxonomy (show diagram) and can be used to classify the types of questioning used in lessons. Do you use (or have you ever used) Bloom's Taxonomy (or similar) to support this process?

- If yes, in what way?
- Have you ever used a taxonomy to classify the types of questions you ask?

Approximately, what proportion of the time you spend on planning a lesson is spent on planning your questioning?

- Do you think it would make any difference to the lesson if you spent more or less time on this?

How do you ensure your questioning assesses what the students know and understand?

- Do you think this could be improved upon?
- If yes, how?

Assessment for Learning techniques are formative techniques used to assess learner's progress (e.g. no hands up, odd one out, mini-whiteboards etc). Do you use any AfL techniques to support your questioning?

- Can you think of any other technique which might support a deep approach from the students?

Can you give me an example(s) of an open question you might use (either in the observed lessons or otherwise)?

- Clarify this term if required, or make suggestions if required.

What proportion of your questions would you estimate are open?

- A separate estimate for each lesson if possible, but overall or in general is also useful.
- Would this be a similar proportion in a 'typical' lesson of yours?"

Can you give me an example(s) of any questions (either in the observed lessons or otherwise) which you believe take a surface approach in terms of the thought process for students?

- Clarify this term if required, or make suggestions if required.

Can you give me an example(s) of any questions (either in the observed lessons or otherwise) which you believe need a deeper approach?

- Clarify this term if required, or make suggestions if required.

What proportion of your questions in these lessons would you estimate required a deeper approach?

- A separate estimate for each lesson if possible, but overall or in general is also useful.
- Would this be a similar proportion in a 'typical' lesson of yours?"

What is your impression of the responses to open questions on the deep/surface spectrum?

- Is this always the case?

"Thank you very much for your time."

Appendix 7 - Teacher Semi-Structured Interview Schedule (Post-Intervention)

"Thank you for letting me observe/video your lessons to look at the questioning you're using with your students. I'd like now to talk about if participating in this research has changed how you consider questioning. Just to remind you that your comments are confidential and you will not be identified in the write up, but some of the dialogue we have may be anonymously quoted. You also have the right to withdraw from taking part at any time."

(Handouts for participants – Bloom's Taxonomy and IMPaCT Taxonomy)

How straightforward do you find using the IMPaCT taxonomy?

Has working with the IMPaCT taxonomy changed the way you plan for questioning with the participant class?

- If so, how?
- Has it had an impact on how you consider questioning for your other classes?

How do you think your questioning affects how your students understand mathematics?

Do you think that working with the IMPaCT Taxonomy has had any effect on how your students understand mathematics?

- If yes, can you give any examples which support this?
- If no, why do you think this is?

How would you compare using the IMPaCT to using Bloom's Taxonomy?

How self-explanatory do you think the IMPaCT taxonomy is for teachers to engage with it without training?

"The following questions are ones which I asked you at the start of the research. I will analyse your responses before and after to see if your thoughts have changed over the course of this research."

How do you ensure your questioning assesses what the students know and understand?

Assessment for Learning techniques are formative techniques used to assess learner's progress (e.g. no hands up, odd one out, mini-whiteboards etc.). Do you use any AfL techniques to support your questioning?

What is your understanding of the difference between a surface and a deeper approach to questioning?

Can you give me an example(s) of any questions (either in the observed lessons or otherwise) which you believe take a surface approach in terms of the thought process for students?

Can you give me an example(s) of any questions (either in the observed lessons or otherwise) which you believe need a deeper approach?

"Thank you very much for your time."

Appendix 8 – Ethical Approval

Application for Ethical Approval for Research Degrees
(MA by research, MPHIL/PhD, EdD)

Name of student: Jo Denton

MA By research	EdD	PhD
	√	

Project title: Encouraging Deeper Conceptual Questioning in the Mathematics Classroom: Working with the Intended Mathematical Processes and Cognitive Thought (IMPaCT) Taxonomy

Supervisor: Sue Johnston-Wilder

Funding Body (if relevant): n/a

Please ensure you have read the Guidance for the Ethical Conduct of Research available in the handbook.

Methodology

Please outline the methodology e.g. observation, individual interviews, focus groups, group testing etc.

- Non-participant lesson observations and video recordings.
- Questioning types and techniques, and responses, will be scribed and coded after the lesson from the lesson observation or video recording.
- Individual interviews with teachers.

Participants

Please specify all participants in the research including ages of children and young people where appropriate. Also specify if any participants are vulnerable e.g. children; as a result of learning disability.

- Nine mathematics teachers from [redacted] School, [redacted] to be observed and/or video recorded with focus on questioning in lessons (question types and techniques to be coded) and interviewed (semi-structured).
- Year 10 students (aged 14-15) taught by the nine mathematics teachers.

Respect for participants' rights and dignity

How will the fundamental rights and dignity of participants be respected, e.g. confidentiality, respect of cultural and religious values?

- No names or identifiable features will be given in write up.

Privacy and confidentiality

How will confidentiality be assured? Please address all aspects of research including protection of data records, thesis, reports/papers that might arise from the study.

- Teachers will be referred to as Teacher X (TX) etc. in the write up.
- Students will be referred to as Student A (SA) etc.
- Video recordings will be automatically stored electronically in accordance with school policy and teachers will be made aware of those who may have access to the recordings and informed of their right to withdraw from the process.

Consent - will prior informed consent be obtained? Yes

- from participants? Yes from others? Yes

- explain how this will be obtained. If prior informed consent is not to be obtained, give reason:

- Permission from the Head Teacher will be obtained to conduct the study and video lessons in accordance with school policy.
- Non-participant lesson observations and video recordings will be with prior consent of teachers.
- Teachers selected will have the purpose of the study explained and how observational data and information given by them in semi-structured interviews will be used. They will be assured of the voluntary nature and confidentiality of participation.
- Students and their parents will be informed by letter of the purpose of the observations and video recordings and how they will be used. Prior consent of students and their parents will be obtained through a rely slip on the letter. If a student does not return their slip then their responses in class will not be used in the write up.

- will participants be explicitly informed of the student's status? Yes

Competence

How will you ensure that all methods used are undertaken with the necessary competence?

All methods have been covered in FRM and ARM taught sessions. Both lesson observations and interviews were carried out in a similar manner for the FRM assignment. Semi-structured interview structure will be discussed with supervisor prior to commencing interview process. Support will be obtained from the school's IT team on the effective use of the IRIS (hardware to record lessons).

Protection of participants

How will participants' safety and well-being be safeguarded?

- Teachers will be assured that it is not a graded lesson observation of their teaching abilities, simply an information gathering process to identify how we can improve learning in the schools.
- Normal classroom health and safety procedures will apply.
- Teachers will be given the choice of videoing or being observed and reminded of the right to withdraw to protect the participants from embarrassment or stress.

Child protection

Will a DBS (Disclosure and Barring Service formerly CRB) check be needed?

No – have a current DBS already as work in the school used in this research.

Addressing dilemmas

Even well planned research can produce ethical dilemmas. How will you address any ethical dilemmas that may arise in your research?

I will in the first instance refer to the BERA guidelines and if I am still unsure I will seek advice from my supervisor.

Misuse of research

How will you seek to ensure that the research and the evidence resulting from it are not misused?

The coded lesson analysis and interview data obtained will not be identifiable to any individual. Original notes from lesson observations and scribed lessons will be kept to ensure misuse and misinterpretation is minimised. A copy of the video will be kept electronically which can only be accessed by me and the network administrators.

Support for research participants

What action is proposed if sensitive issues are raised or a participant becomes upset?

If a participant becomes upset they are reminded that they have the right to withdraw at any time, including stopping the recording during the lesson or if a student wishes to leave the room as a result of the research, provision will be made for them to work at the back of another classroom.

Integrity

How will you ensure that your research and its reporting are honest, fair and respectful to others?
Inter-reliability testing on the observation coding should help to get honest and fair data collected. Participants will be sent a copy of the research once written up and invited to discuss and make any corrections.

What agreement has been made for the attribution of authorship by yourself and your supervisor(s) of any reports or publications?
Supervisor role will be credited.

Other issues?

Please specify other issues not discussed above, if any, and how you will address them.

None

Signed

Research student: J L Denton Date: 18/01/2015

Supervisor: Sue Johnston-Wilder Date: 18/01/2015

Appendix 9 - Parent/Carer Consent Letter

Dear Parent/Guardian

I am doing some research into questioning in mathematics for my Doctorate in Education at the University of Warwick. This research will involve videoing and/or observing Year 10 mathematics lessons. I am primarily interested in the questions asked by the teachers, however students' responses may also be noted and included in my thesis. All participants will remain anonymous and no student will be identifiable in the write up as they will simply be referred to as Student A (SA), Student B (SB) etc. Any video recordings of lessons made will only be accessible to me and your son/daughter's mathematics teacher. If you would prefer for your son/daughter to either not be quoted in the thesis or not be visible in any video recordings please indicate this on the response slip below. You have the right to change your mind and withdraw your child at any time. If you have any queries please feel free to contact me at jdenton@********.sch.uk.

Yours faithfully

Mrs Jo Denton
Lead Practitioner for Mathematics

Please complete and return to your son/daughter's mathematics teacher.

Student's Name:____

Mathematics Class:__________ Mathematics teacher:_______________________

1. I do/do not* give my permission for my son/daughter to be visible in the video recording of their mathematics lesson.
2. I do/do not* give my permission for my son/daughter to be anonymously quoted in the thesis write up.

Signed:____ (Parent/Guardian)

Print name:_______________________________________ (Parent/Guardian)

Appendix 10 – Parent/Carer Consent (Opt-Out) Letter

Dear Parent/Guardian

I am doing some research into questioning in mathematics for my Doctorate in Education at the University of Warwick. This research will involve videoing and/or observing Year 10 mathematics lessons. I am primarily interested in the questions asked by the teachers, however students' responses may also be noted and included in my thesis. All participants will remain anonymous and no student will be identifiable in the write up as they will simply be referred to as Student A (SA), Student B (SB) etc. Any video recordings of lessons made will only be accessible to me and your son/daughter's mathematics teacher. If you would prefer for your son/daughter to either not be quoted in the thesis or not be visible in any video recordings please return the response slip below by Monday 27th April 2015. You have the right to change your mind and withdraw your child at any time. If you have any queries please feel free to contact me at jdenton@********.sch.uk.

Yours faithfully

Mrs Jo Denton
Lead Practitioner for Mathematics

Please complete and return to your son/daughter's mathematics teacher.

Student's Name:____

Mathematics Class:_________ Mathematics teacher:_______________________

1. I would not like my son/daughter to be visible in the video recording of their mathematics lesson.
2. I do not give my permission for my son/daughter to be anonymously quoted in the thesis write up.

Signed:_______________________________________ (Parent/Guardian)

Print name:_____________________________________ (Parent/Guardian)

Appendix 11 – Participant Teacher Consent Form

I agree to my lessons being observed and/or videoed to look at the questioning I am using with my students. I understand that my lessons will not be judged and any recordings will only be accessible to the researcher and will not be shared with other colleagues in the school. I understand that in the thesis my questions will be analysed and quoted where required but my identity will remain anonymous.

I also agree to participate in two interviews to talk about my approach to planning for questioning and the techniques I use to support questioning. I understand that the comments I make are confidential and I will not be identified in the write up, but some of the dialogue we have may be anonymously quoted.

I understand that I have the right to withdraw from taking part in this research at any time.

Signed ………………………………………………………………………………

Date………………………………………………

Appendix 12 – Questions Transcribed from 2 Methods

Q no.	Observed lesson	Code	Recorded Lesson	Code
1	Why isn't a funnel a complete cone?	REA-D	If we imagine a funnel as a complete cone, why isn't it?	REA-D
2	What shape cone would use the least amount of plastic to manufacture a funnel containing a set volume?	DER-D	What shape cone would be using the least amount of plastic to manufacture a funnel containing a set volume?	DER-D
3	Would that be a good idea?	REF-S	Would that be a good idea?	REF-S
4	What might you do then?	PRO-S	What might you do then?	PRO-S
5			Why (do you need to make a funnel)?	REA-S
6			Can it be any size?	REF-S
7	Why don't you try and create a cone, how would you do that?	DER-D	How are you going to design a cone?	DER-D
8	What do we need to consider?	REF-D	What do we need to consider?	REF-D
9	Is that a cone?	FAC-S	Is that a cone?	FAC-S
10	Can you make a cone without using so much paper…a more efficient cone?	REF-D	Can you make a cone without using so much paper? A more efficient cone?	REF-D
11	That's an overlap. How are you going to get over that?	REF-D	That's an overlap. How are you going to get around that?	REF-D
12	Why do you need a circle?	REA-S	Why do you need a circle?	REA-S
13	What's the problem with doing that?	REF-S	What's the problem with doing that?	REF-S
14			Why is it an issue?	REF-S
15			Why do we need a protractor?	PRO-S
16	What do we need to consider if we are trying to make a funnel with least amount of material?	REF-D	What do we need to consider if we are trying to make a funnel with the least amount of material for a set volume?	REF-D
17			What dimensions do we need to be in control of?	STR-D
18	We can change the radius and height, what else do we need to think about?	REF-S	We can change the radius and height, what else do we need to think about?	REF-S
19	How do we know we are going to meet this challenge?	REF-D	How do we know we are going to meet this challenge?	REF-D
20			What are we going to try?	REF-S
21	We're going to fix the volume to be a litre. What do we need to know about that?	PRO-S	We're going to fix the volume to be a litre. What do we need to know about that?	PRO-S
22	Does anyone know how to calculate the volume of a cone?	FAC-S	Does anyone know how to calculate the volume of a cone?	FAC-S

23			What are we fixing this at?	PRO-S
24			How are we going to measure the dimensions of this/What are we going to use to measure it?	PRO-S
25	Is that in litres?	REF-S	Is that in litres?	REF-S
26	What do we need to consider?	REF-D	So what problem do we need to consider?	REF-D
27	What do we need to know?	FAC-S	What do you need to know?	FAC-S
28	(Can you) go and find the volume of the one you've created?	PRO-S	How are you going to work out the volume of the one you've got? You've got the formula.	PRO-S
29			What do we need?	REF-S
30	Is the radius of the circle the same as that radius?	STR-D	Is that the same radius as the original circle you cut out?	STR-D
31	How are we going to find the radius of that?	DER-D	So how are we going to find the radius of that?	DER-D
32	Where's the radius (now)?	FAC-S	Where's the radius (now)?	FAC-S
33	99? How did you get that?	REA-S	99? How did you get that?	REA-S
34			Make it a cone…what happens?	REF-S
35			It is less, so what do you need to find out now?	REF-S
36			What might be the easiest thing?	REF-D
37	What have you got?	FAC-S	What have you got?	FAC-S
38			What's your volume? Is it?	REA-S
39	There's mine. What did you measure to get the radius?	REA-S	There's mine. What did you measure to get the radius?	REA-S
40			All of this or did you do this?	PRO-S
41			There's your hole..now if I fold it up…is it the same radius?	STR-D
42	How are you going to measure the height of the cone?	REF-S	How are you going to work out the height of the cone?	REF-S
43			How are you going to measure it?	PRO-S
44	Is the height of the cone the same as this?	FAC-S	Is the height of that cone the same as this (radius)?	FAC-S
45			Are you sure?	REF-S
46	Can you measure how high my cone is?	PRO-S	There's my cone…can you measure how high the cone is?	PRO-S
47			Is that the height? If they've gone and done that for Mount Everest, did they drop a long rope down and said it's 8848m?!	REF-S
48	Have you found the radius?	PRO-S	Have we found the radius of that?	PRO-S
49	What did you notice when you went from that to that?	REF-S	What did you realise about this (3D cone) compared to	REF-S

	What did you have to realise?		that (circle)? What did you have to realise?	
50	How do we find the radius of something like this?	REF-S	How do we find the radius of something like this?	REF-S
51	Some people have said that's easy just measure the radius. Why doesn't that work?	REA-S	Now I need the height, some people said that's easy just measure the radius of the circle we cut out. Do we?	REF-S
52			Why not?	REA-D
53			When people were measuring the radius of the actual circle they cut out what were they finding?	STR-D
54	What am I trying to find (now)?	REF-S	What do we need to find?	REF-S
55	How can I do that?	DER-D	How are we going to do that?	DER-D
56	Can you rearrange the equation/formula?	PRO-S	Can you rearrange the formula?	PRO-S
57			Can you now find the height and your volume?	PRO-S
58	Using Pythagoras, can you find the volume?	PRO-S	Using Pythagoras, can we find the volume?	PRO-S
59	Have you used Pythagoras to find the height?	PRO-S	Can we find the height using Pythagoras?	PRO-S
60	That was 6 was it? Show me.	REA-S	That was 6 was it?	REA-S
61	There's the rule, (can you) find the volume?	PRO-S	What did you get (for the volume)...there's the rule, find the volume.	PRO-S
62			Why are you taking the larger one away from the smaller one?	REA-S
63			Which is the larger one out of those two?	FAC-S
64	Can you find out how much material you're using?	DER-D	Can you find out how much material you're using?	DER-D
65			Not sure about it? Why not?	REF-S
66			Doesn't look like 7.4 to me, that would be up here wouldn't it?	REF-S
67	The slant height is 4.2, the diameter is 6.4, so what is the radius?	PRO-S	The slant height is 4.2, the diameter is 6.4, so what is the radius?	PRO-S
68	How do you find how much paper used?	REF-D	Can you find out how much paper in your cone?	DER-D
69			Oh it's about is it?	REF-S
70			What do you need to consider?	REF-D
71	Mine isn't half the circle, how are you going to do mine?	DER-D	Mine isn't (half the circle), what are you going to do about mine?	DER-D
72			Is that one quarter?	FAC-S
73	So what is it if not 90 degrees?	FAC-S	So what is it if not 90 degrees?	FAC-S

74			How did that help you then work out (the sector)	REA-S
75			How much is your sector?	PRO-S
76	So what is the surface area?	PRO-S	Now can you start to think about that (the surface area)?	REF-S
77			3000? 3000 what?	FAC-S
78	Why have you got a radius of 1?	REA-S	Why have you got a radius of 1?	REA-S
79	How does that affect the height?	STR-D	How does that affect the height?	STR-D
80			How might you write down ideas to track what's most appropriate?	REF-D
81	What shape did we start with?	FAC-S	What shape did we start with?	FAC-S
82	What's the area of a circle?	FAC-S	What do you know about finding the area of a circle?	FAC-S
83	Is that a full circle?	FAC-S	Is that a full circle?	FAC-S
84	How much of a circle?	FAC-S	How much of a circle?	FAC-S
85			How did you find the fraction?	PRO-S
86	It looks a bit like a quarter. How could you make it more accurate?	REF-S	It looks a bit like a quarter. How could you make a more accurate measurement? What do you need to know?	REF-S
87	Why might this be useful?	REA-S	Why might this be useful?	REA-S
88	How do I find the area of the sector I've removed?	DER-D	How am I going to find out the area of this? How am I going to find out the area of the sector I've removed?	DER-D
89	What would make it a quarter?	REF-S	What would make it a quarter? What angle would I have here if I cut out a quarter?	REF-S
90	What angle if I made half?	FAC-S	What angle would I have here if I took off a half of it?	FAC-S
91			Well I haven't done that so what do I need to figure out?	REF-S
92			Can we measure it?	PRO-S
93			Are we sure?	REF-S
94	How does that help you find what's left?	REF-S	How does it help you find what's left?	REA-S
95	What's that?	PRO-S	What are you doing? What's that?	REA-S
96	(Can you) then square root it to give you the height?	PRO-S	(Can you) then square root it to give you the height?	PRO-S
97	What is 5.1^2-3.8^2	PRO-S	Can you do 5.1^2-3.8^2? What is the answer to that?	PRO-S
98	What do you do to that number?	PRO-S	What do you do with that (number)?	PRO-S
99			What have we got so far?	REF-S
100	Why is the radius less than 10?	REF-S	The radius less than 10, how do you know?	REF-S
101	What's the height when the radius is 10?	PRO-S	What's the height if the radius is 10?	PRO-S

102			Can you get any closer without integer measurements?	REF-S
103	If the radius was 1cm, 2cm, 3cm, what would the height be?	PRO-S	If the radius was 1cm, 2cm, etc, what would the height be?	PRO-S
104	Why is the size of the angle important?	REA-S	Why is the size of the angle important?	REA-D
105	That was the angle he had out of 360. Why S_{10}?	REA-D	That was the angle he had out of 360. Why S_{10}?	REA-D
106	Where does that come from S_{11}?	REA-S	Where does that come from S_{11}?	REA-S
107	Problem with cone don't know angle. The formula is…?	FAC-S	Problem with cone don't know angle. The formula is…?	FAC-S
108	J & K can you tell what you did?	REA-S	J & K can you tell what you did?	REA-S
109	We need to rearrange this for everyone else. What do I do first?	REA-S	We need to rearrange this for everyone else. What do I do first?	REA-S
110	What next?	REA-S	What next?	REA-S
111	What did you find that to be?	REA-S	What did you find that to be?	REA-S
112	Is that practical?	REF-D	Is that practical?	REF-D

Appendix 13 - Questions to sort using IMPaCT Taxonomy

What's 3 x 6? FAC

Can you describe the properties of a rectangle? FAC

Can you remember what product means? FAC

When we're dividing by fractions, what do we do? PRO

Have you added it up yet and divided by 3? PRO

0.2 x 0.5

In this question how many decimal places do we have? So how many in the answer? PRO

You know what x is. How can you use this information to find what y is equal to? SUR-STR

You know 4 x 5 = 20, so what's 400 x 5? DP-STR

How did you work that out? SUR-REA

Why has she chosen to do that? DP-REA

Are you happy that you are right? How did you check? DP-REA

What other ways could you possibly do the decrease? SUR-REF

Which average would be best to use here? DP-REF

(After practising adding and subtracting decimals): Billy goes to visit his Gran. She gives him £2.75. Granddad then gives him £4.60. How mush has he got altogether? SUR-DER

In groups discuss how you might solve these simultaneous equations with a 'substitution method'? (Substitution method not met before) DP-DER

Appendix 14 - Developing a Climate for Effective Classroom Discourse

Classroom Social Norms

Social norms are expectations established by the teacher in the classroom. Social norms relating to effective classroom discourse are characterised by developing the following:

- explanation
- justification
- argumentation
- a climate where it is acceptable to make mistakes

These characteristics are not specific to mathematics lessons, as learners should be expected to justify their own thinking and challenge the thinking of others across the curriculum, not just in mathematics.

Sociomathematical Norms

To develop learners' mathematical thinking, norms which are unique to the learning of mathematics need to be established. A sociomathematical norm is intended to "set an expectation in the classroom that encourages strong mathematical activity in the form of justification" (Gerson & Bateman, 2011, p. 115). e.g.

- understanding of what is constitutes an *acceptable mathematical explanation and justification*
- developing an understanding of *mathematical difference*
- developing an understanding of *mathematical sophistication*, *mathematical efficiency* and *mathematical elegance*

Appendix 15 – Example of Baseline Analysis Handout

Lesson 1

No. of Qs	S1	S2	S3	Total Surface	D1	D2	D3	Total Deeper	Total
Factual	6	4	1	11					11
Procedural	4	6	10	20					20
Reasoning	2	0	0	2	1	0	0	1	3
Reflective	1	1	1	3	0	1	0	1	4
Structural				0	5	1	2	8	8
Derivational				0	2	0	0	2	2
Total	13	11	12	36	8	2	2	12	48

% of Qs	% S1	% S2	% S3	% Surface	% D1	% D2	% D3	% Deeper	Total %
Factual	12.5	8.3	2.1	22.9%					22.9%
Procedural	8.3	12.5	20.8	41.7%					41.7%
Reasoning	4.2	0.0	0.0	4.2%	2.1	0.0	0.0	2.1%	6.3%
Reflective	2.1	2.1	2.1	6.3%	0.0	2.1	0.0	2.1%	8.3%
Structural					10.4	2.1	4.2	16.7%	16.7%
Derivational					4.2	0.0	0.0	4.2%	4.2%
Total	27.1	22.9	25.0	75.0%	16.7	4.2	4.2	25.0%	100.0

Lesson 2

No. of Qs	S1	S2	S3	Total Surface	D1	D2	D3	Total Deeper	Total
Factual	4	5	3	12					12
Procedural	0	2	2	4					4
Reasoning	1	3	0	4	0	0	1	1	5
Reflective	0	0	1	1	0	0	0	0	1
Structural				0	0	0	0	0	0
Derivational				0	2	2	0	4	4
Total	5	10	6	21	2	2	1	5	26

% of Qs	% S1	% S2	% S3	% Surface	% D1	% D2	% D3	% Deeper	Total %
Factual	15.4	19.2	11.5	46.2%					46.2%
Procedural	0.0	7.7	7.7	15.4%					15.4%
Reasoning	3.8	11.5	0.0	15.4%	0.0	0.0	3.8	3.8%	19.2%
Reflective	0.0	0.0	3.8	3.8%	0.0	0.0	0.0	0.0%	3.8%
Structural					0.0	0.0	0.0	0.0%	0.0%
Derivational					7.7	7.7	0.0	15.4%	15.4%
Total	19.2	38.5	23.1	80.8%	7.7	7.7	3.8	19.2%	100.0

Lesson 3

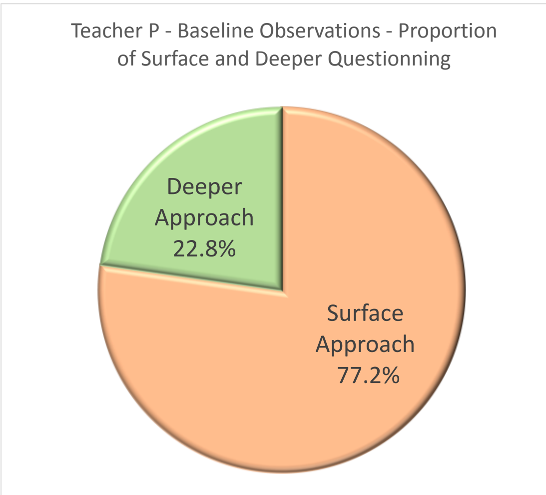
No. of Qs	S1	S2	S3	Total Surface	D1	D2	D3	Total Deeper	Total
Factual	8	5	5	18					18
Procedural	5	6	4	15					15
Reasoning	2	1	0	3	0	0	0	0	3
Reflective	1	1	0	2	1	1	1	3	5
Structural				0	1	4	2	7	7
Derivational				0	1	0	0	1	1
Total	16	13	9	38	3	5	3	11	49

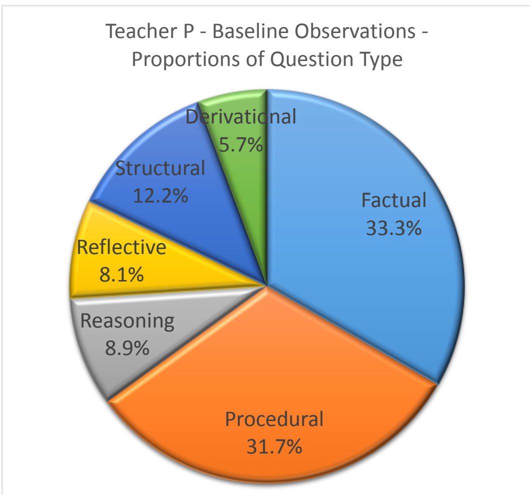
% of Qs	% S1	% S2	% S3	% Surface	% D1	% D2	% D3	% Deeper	Total %
Factual	16.3	10.2	10.2	36.7%					36.7%
Procedural	10.2	12.2	8.2	30.6%					30.6%
Reasoning	4.1	2.0	0.0	6.1%	0.0	0.0	0.0	0.0%	6.1%
Reflective	2.0	2.0	0.0	4.1%	2.0	2.0	2.0	6.1%	10.2%
Structural					2.0	8.2	4.1	14.3%	14.3%
Derivational					2.0	0.0	0.0	2.0%	2.0%
Total	32.7	26.5	18.4	77.6%	6.1	10.2	6.1	22.4%	100.0

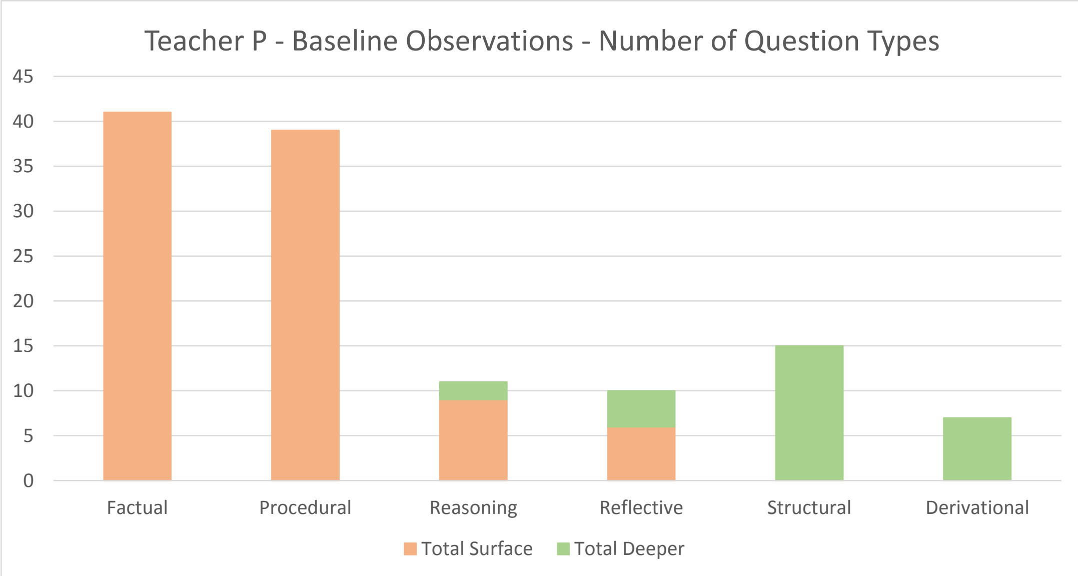
Overall

No. of Qs	S1	S2	S3	Total Surface	D1	D2	D3	Total Deeper	Total
Factual	18	14	9	41					41
Procedural	9	14	16	39					39
Reasoning	5	4	0	9	1	0	1	2	11
Reflective	2	2	2	6	1	2	1	4	10
Structural					6	5	4	15	15
Derivational					5	2	0	7	7
Total	34	34	27	95	13	9	6	28	123

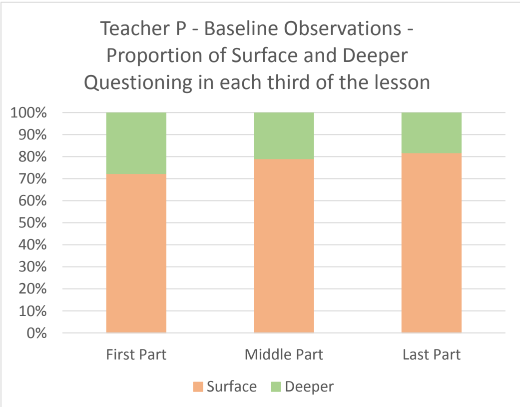
% of Qs	% S1	% S2	% S3	% Surface	% D1	% D2	% D3	% Deeper	Total %
Factual	14.6	11.4	7.3	33.3%					33.3%
Procedural	7.3	11.4	13.0	31.7%					31.7%
Reasoning	4.1	3.3	0.0	7.3%	0.8	0.0	0.8	1.6%	8.9%
Reflective	1.6	1.6	1.6	4.9%	0.8	1.6	0.8	3.3%	8.1%
Structural					4.9	4.1	3.3	12.2%	12.2%
Derivational					4.1	1.6	0.0	5.7%	5.7%
Total	27.6	27.6	22.0	77.2%	10.6	7.3	4.9	22.8%	100.0



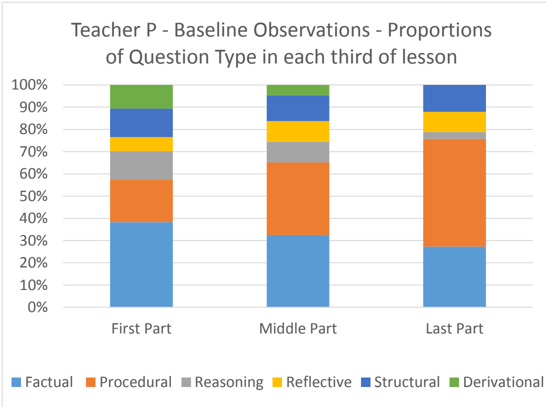




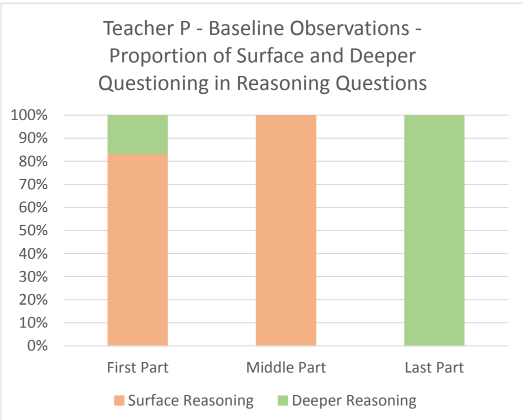
Teacher P - Baseline Observations - Number of Question Types
45
40
35
30
25
20
15
10
5
0
Factual
Procedural
Reasoning
Reflective
Structural
Derivational
Total Surface
Total Deeper



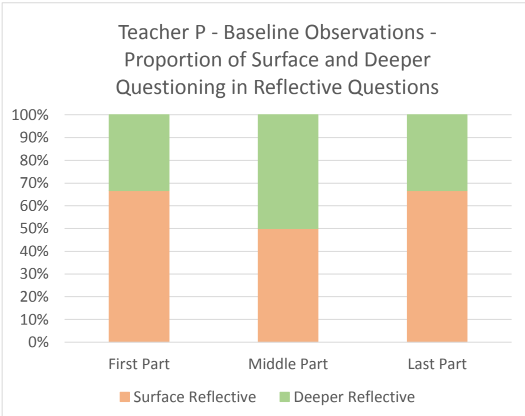
Teacher P - Baseline Observations - Proportion of Surface and Deeper Questioning in each third of the lesson
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
First Part
Middle Part
Last Part
Surface
Deeper



Teacher P - Baseline Observations - Proportions of Question Type in each third of lesson
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
First Part
Middle Part
Last Part
Factual
Procedural
Reasoning
Reflective
Structural
Derivational



Teacher P - Baseline Observations - Proportion of Surface and Deeper Questioning in Reasoning Questions
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
First Part
Middle Part
Last Part
Surface Reasoning
Deeper Reasoning



Teacher P - Baseline Observations - Proportion of Surface and Deeper Questioning in Reflective Questions
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
First Part
Middle Part
Last Part
Surface Reflective
Deeper Reflective