University of Southampton

Faculty of Human and Social Sciences

Southampton Education School

Young children's cognitive representations of number and their number line estimations

Joanna Williamson

Thesis for the degree of Doctor of Philosophy

December 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Southampton Education School

Doctor of Philosophy

YOUNG CHILDREN'S COGNITIVE REPRESENTATIONS OF NUMBER AND THEIR NUMBER LINE ESTIMATIONS

By Joanna Williamson

This doctoral research examines children's cognitive representations of number during their first year of formal schooling (in England, from age five to six). Extensive previous research on this topic has been carried out in education and also cognitive psychology, including an important strand focusing on children's number line estimations. This doctoral research is theoretically underpinned by an inclusive understanding of representation in mathematics, as set out by Raymond Duval, and uses this theoretical perspective to provide an original analysis of number line estimation tasks and to make original connections with children's imagistic representations of number, as previously studied separately in education research.

In a longitudinal multiple case-study design, thirteen children took part in five video-recorded interviews each, at 6-8 week intervals. In each interview children completed both number line estimation tasks and imagistic representation tasks, thus providing longitudinal and qualitative data not seen in previous research. The theoretical framework necessitated multimodal analysis of representations, and quantitative analyses from existing research were also carried out for comparability with previous work.

As found by previous research, children's number line estimations more closely resembled linear distributions with time. Changes in children's estimations were convincingly linked to the representations of number structure that they made during the estimations. This connection provided a better explanation of the observed changes than either a proportional reasoning or log-linear shift account of number line estimation. Children's strategies and representations varied with the particular context of estimation trials, indicating adaptability and further weakening the case for inferring from estimation trials a representation of entire number ranges. The key educational implication of this research is that all children (not only the 'high-attaining') represented structural aspects of number and developing connections, which should be harnessed and encouraged in order to support their developing concepts of number.

Contents

Chapter 1 Introduction 1

1.1 Overview 1

1.2 Statement of the research problem 1

1.3 Key theoretical concepts 1

1.4 Relation of research to existing literature 5

1.5 Justification of research topic 7

1.6 Research questions 8

Chapter 2 Literature review 11

2.1 Overview 11

2.2 Representation 11

2.3 Conceptions of number 20

2.4 Theories of number development 24

2.5 What is represented by a representation of number? 39

2.6 Number and space 40

2.7 Representational changes 42

2.8 Strategy in estimation and its relation to cognitive representation 47

2.9 Imagistic representations of number 52

2.10 Numerical representations and mathematics 54

2.11 Literature review postscript 58

Chapter 3 Theoretical Framework 67

3.1 Overview 67

3.2 Duval's theory of cognitive representations and visualisation 68

3.3 Cognitive integration 79

3.4 Duval's theory and the research problem 81

3.5 The analysis of imagistic representations 82

Chapter 4 Methodology 89

4.1 Overview of research design 89

4.2 Task-based interviews 95

4.3 Research design: specific description 105

4.4 Theory for analysis 112

4.5 Plan of analysis 120

Chapter 5 Pilot study 127

5.1 Setting and sample 127

5.2 Study design 127

5.3 Task 2: quantitative analysis 128

5.4 Task 3 130

5.5 Task 4: quantitative analysis 131

5.6 Case study: Imogen 134

5.7 Changes to research design based on the pilot study 140

Chapter 6 Findings from the group 143

6.1 Abbreviations 143

6.2 Participants 143

6.3 Task 3 144

6.4 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research? 147

6.5 RQ2: What strategies can be identified in children's interactions with number line estimation tasks? 157

6.6 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling? 159

Chapter 7 Case study: Patrick 171

7.1 What does this case say in response to important ideas in the literature? 171

7.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research? 173

7.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks? 192

7.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling? 205

Chapter 8 Case study: Marta 217

8.1 What does this case say in response to important ideas in the literature? 217

8.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research? 218

8.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks? 232

8.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling? 260

Chapter 9 Case study: Catharina 267

9.1 What does this case say in response to important ideas in the literature? 267

9.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research? 268

9.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks? 279

9.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling? 302

Chapter 10 Summary and conclusions 311

10.1 Overview 311

10.2 Conclusions 312

10.3 Significance of findings 320

10.4 Theoretical and methodological reflections 322

10.5 Implications for practice 323

10.6 Considerations for future research 325

References 327

Appendix 1 336

Appendix 2 337

Appendix 3 339

Appendix 4 340

Appendix 5 341

Appendix 6 342

Appendix 7 346

Academic Thesis: Declaration Of Authorship

I, Joanna Williamson, declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

Title: Young children's cognitive representations of number and their number line estimations

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as:
    a. Williamson, J. (2013). Young Children's Cognitive Representations of Number and Their Number Line Estimations. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 401-408). Kiel, Germany: IPN–Leibniz Institute for Science and Mathematics Education at the University of Kiel.
    b. Williamson, J. (2014). Investigating young children's number line estimations using multimodal video analysis. In S. Pope (Ed.), *Proceedings of the 8th British Congress of Mathematics Education 2014* (pp. 367-374). British Society for Research into Learning Mathematics.

Signed:

Date:

Acknowledgements

I would like to thank Dr Julie-Ann Edwards and Dr Charis Voutsina for their unerring guidance, for challenging me when I needed challenging and for unwavering support when I needed support. I have enjoyed every minute of supervision and they have made the past four years brilliant.

I would like to thank all at Southampton Education School, but particularly comrades Eva Nedbalova, Gisela Cebrian-Bernat, Becca Ridley and Selar Othman, for all their friendship and for the CoP.

I would like to thank my parents, and Ben, and Sam, for more than I could ever list.

Chapter 1 Introduction

1.1 Overview

This chapter begins with a brief statement of the research problem and its situation relative to the existing research literature. Next, it introduces the key theoretical concepts necessary for situating and understanding the research. It then discusses the specific niche in the literature which the research is designed to address, and argues for the value of addressing this particular research problem. The chapter concludes with a full statement of the research questions.

1.2 Statement of the research problem

This research examines children's cognitive representations of number during their first year of formal schooling (in England, from age five to six). During this important time, children experience significantly increased exposure to the norms, artefacts and strategies of the mathematics classroom, and their mathematical knowledge typically increases considerably. An aspect of this mathematical knowledge that is particularly relevant for this research is children's conceptions of natural number, which – in ways that will be discussed in the following sections – are still developing during this period. The unique contribution aimed for by this research is to provide depth of understanding of representations of number during this particularly significant year of children's education. This aim is to be achieved by combining methods of investigation from rather different previous approaches to the research problem: the research will qualitatively analyse children's responses to the most commonly-used clinical-interview task, and will complement the majority of existing studies by adopting a longitudinal approach.

Research into cognitive representations of number has origins in two important and well-established fields of research. The first field is research into the development of children's conceptions of number. Significant research into this question has been carried out in education, cognitive psychology, and cognitive neuroscience. The second major research area that this topic is rooted in is research into the important role of representation in mathematical thinking and development. Included within this is the specific research area of imagery in mathematical thinking.

1.3 Key theoretical concepts

1.3.1 Representation

The lens of *representations* can be seen as a natural fit for mathematics education research. Representations are themselves central to mathematics practice, mathematical communication, and mathematics education (Duval, 1999; Vergnaud, 1987). A perspective that researches mathematics education through examining representations thus takes as its starting point an already existing and central construct, rather than imposing a construct from outside of the domain of research.

1.3.1.1 Understanding representation

The fundamental meaning of representation is clear: “one thing’s ‘standing for’, ‘being about’, ‘referring to or denoting’ something else” (Schwartz, 2005, p. 536). Theoretical approaches to representation commonly make a primary distinction between internal (or mental) representations and external representations: mental images, for example, are classified as internal representations, whilst inscriptions on paper are external representations (e.g. Goldin, 2008; Larkin and Simon, 1987).

The precise relation between a representation and that which it represents is a problematic point for both mental and external representations (Schwartz, 2005, p. 536), but in the case of mental representations, this difficulty is only one of many complications. The notion of mental representation arises in an innocuous way from a common way of describing thoughts: “When we think about the Eiffel Tower … we can be said to possess a mental representation of the Eiffel Tower” (Guttenplan, 2005, p. 441). However, “deep and vexing problems arise” when discussing mental representations (p. 441). These include the question “Are thoughts somehow made up of mental representations?” (Guttenplan, 2005, p. 441) and ontological concerns about what mental representations really ‘are’ (Schwartz, 2005, p. 540). These problems are among the most difficult and fundamental questions known, and in fact “set the agenda for a large part of contemporary philosophy of mind” (Guttenplan, 2005, p. 441).

1.3.1.2 Key concept: cognitive representation

The concept of representation utilised in this research is an inclusive one, as expressed by Raymond Duval (1999). Duval sets out a framework of all cognitive representations – i.e. representations involved in cognitive processes. This theoretical framing of cognitive representations acknowledges both *intentional* representations (deliberately constructed semiotic representations, which can be internal or external) and *automaticised* representations (including perceptual representations). It also acknowledges potential relations between these categories. This framing is particularly necessary in the context of representations of number, since empirical evidence suggests strong and as yet not fully understood connections between intentional and automaticised representations. An important feature of Duval’s framework is that “mental images” exist in both categories.

The distinction between intentional and automaticised representations in Duval's framework is taken as the primary distinction. In this respect, Duval's framework differs significantly from alternative theories, which, as noted, commonly prioritise the distinction between internal and external representations.

The distinction between intentional and automaticised representations has explanatory power with regard to the source of cognitive representations, and clarifies the extent of Duval's conception of cognitive representation. The distinction does **not** mean that representations must be considered only within these categories; and this thesis does not set out to categorise representations into intentional or automaticised. The aim is to recognise the relevance of both of Duval's types of cognitive representation, and adopt methods that are equipped to research both.

1.3.1.3 Relation of cognitive representation to internal/external representation

Whilst the term *cognitive representation* can appear in the literature as a synonym for *mental* or *internal representation*, Duval explicitly includes external representations within the scheme of cognitive representations. This has important implications. The terminological implication is that *cognitive* is not to be used as a synonym for *mental* or *internal*. The word *cognitive* is reserved to describe – precisely – phenomena involved in *cognition*, and these phenomena may or may not be internal; what is relevant is not their 'location' but rather their relation to *cognition*.

Once *cognitive* is distinguished from *internal*, a wider range of phenomena than traditionally considered may be admitted within the category of *cognitive* activity. As Duval's examples and comments indicate, cognitive representations in fact occur using a plurality of what are commonly termed *modes*, for example speech, inscriptions on paper, and mental imagery. Since cognitive representations may occur using any of these multiple modes, this theoretical perspective necessitates a multimodal research approach.

1.3.2 Concepts and conceptions of number

As the literature review discusses, there are many difficulties associated with the philosophical construct of *concepts*, and some further potential difficulties in the specific application *concept of number*. However, the construct is retained and used since in both informal reflection and philosophy of mind, the notion of *concept* plays a vital role. Concepts are essential to "the familiar form of explanation (so-called 'intentional explanation') by which we explain the behaviour and states of people" (Rey, 2005, p. 185).

Children's *conceptions* or *concepts* of number are a vital aspect of mathematical knowledge. In this thesis, I research children's cognitive representations and the features of conceptions of number that they express. The conceptions incorporate children's differing and maturing ideas about number. The thesis does not seek to *judge* the maturity of conceptions of number, but instead is seeking evidence of connections into the number concept as evidenced in children's cognitive representations. In order to precisely analyse the aspects of number represented by children, the thesis will make extensive use of Resnick's (1983) microstage theory of number development.

Both *concept* and *conception* appear in the relevant literature, and in terms of their strict definition, there is little difference. The Oxford English Dictionary offers "a concept" in defining *conception* and offers "the product of the faculty of conception" in defining *concept.* Both are used to describe ideas and notions, but *concept* is more often applied to a general or abstracted idea, whereas *conception* emphasises that the idea has been conceived by an individual, and is possibly more idiosyncratic and less general (especially as relates to "apprehension, imagination" (Oxford English Dictionary online, 2012)).

This thesis chooses to discuss children's *conceptions* of number, but it should be noted that in much of the literature, authors may use *concept* of number with the identical intended meaning. The important related phrases *number understanding* and *number sense* will be discussed fully in later chapters.

Despite the ubiquity of the notion *concept*, there exists no single definition of what a concept is. They are a form of thought capable of being shared, and the "Classical View" of concepts is that they can be analysed into their components: the canonical example given is the concept [bachelor], which is analysed into [eligible unmarried male] (Rey, 2005, p. 187).

This 'Classical View' does not apply easily to the case of *concept of number*. In fact "although the whole edifice of modern mathematics is built upon the concept of the natural number, this concept remains something of a mystery" (Lovell, 1965, p. 26), a paradox explored at length by Badiou (2008).

A particular difficulty is the ambiguous construction 'concept of X', seen frequently in research literature on education and developmental psychology– for example 'the child's concept of causality' (Rey, 2005). The phrase '*the* concept of number' is commonly used to refer to the general shared notion of number, e.g. "Such people believe that the concept of the natural numbers is the result of a primitive intuition" (Lovell, 1965, p. 27). In contrast, '*a* concept of number' is frequently used in developmental research: "difficulties in assessing the cardinality of a set imply lack of a concept of natural numbers" (Rips, Bloomfield and Asmuth, 2008, p. 630). In this example, "a" suggests that there is more than one concept of natural numbers; a child does not simply have or not have "the" universal concept *number* but develops, individually, "a" concept of number (which could then be expected to develop further still). In summary, *concept of number* is used to refer to both 'the' concept of number – the general shared notion, whatever that may be, and a person's (individual) conception of number.

Research in recent years has found increasing evidence of cognitive activity in babies and young children relating to quantities (see Mix, Huttenlocher and Levine, 2002). This has further complicated the challenge of pinning down *concept of number,* although most researchers agree that there exists a clear gap between the properties of the natural numbers, and the properties of the numerical conceptions studied in very young children (Rips, Bloomfield, et al., 2008, p. 623). The essence of concepts or conceptions of number, as opposed to numerical conceptions, is that they integrate knowledge: knowledge of representations of quantity, together with a system of counting. In this view, the concept of natural number is the integration of conceptions of cardinality and seriation (Lovell, 1965, p. 51; Nunes and Bryant, 2009, p. 12; Piaget, 1969a, p. viii; Resnick, 1983, pp. 146-147). The process of integrating these conceptions into a concept of number takes years. This thesis will use Resnick's (1983) microstage theory of number concept development to analyse at a finer level of detail the aspects of numerical and number conceptions that children represent. Built upon empirical research into children's mathematical activity, Resnick's theory identifies aspects of number structure that children typically represent as they gradually develop a concept of number that fully integrates their knowledge of counting and of quantities. The broad stages of number structure identified (within which the theory specifies microstages) are sequence structure, the relative numerosity or magnitude of numbers, and finally the part-whole structure of numbers.

This thesis accepts that concepts and conceptions are unavailable to research directly, and chooses to seek evidence about them through the phenomenon of cognitive representations. Cognitive representations are not taken to be concepts, nor proxies for concepts, but phenomena in their own right.

1.4 Relation of research to existing literature

Since this thesis addresses a problem that relates to multiple areas of existing research, the problem should be situated and understood in relation to the research literature in each of these areas.

A great deal of recent work on the cognitive representation of number has been carried out within cognitive psychology. A key conclusion has been that people commonly represent number on a **mental number line**, a left-to-right oriented number line 'in the mind' which is variously understood as a metaphor (Fias and Fischer, 2005), as mental imagery, or as synonymous with the

Analogue Magnitude System (Dehaene, 2001) - a cognitive system “in which the cardinal value of a set is represented by an analog symbol that is a linear or logarithmic function of the number of elements in the set” (Le Corre and Carey, 2008, p. 651).

A growing body of research concludes that there is a ‘shift’ in the structure of children’s mental number lines, which occurs with age. This log-linear hypothesis holds that younger children typically represent numbers on a mental number line with logarithmically positioned numbers, whilst older children and adults represent numbers on a line with linearly positioned numbers (see for example Berteletti, Lucangeli, Piazza, Dehaene and Zorzi, 2010; Dehaene, Izard, Spelke and Pica, 2008; Opfer and DeVries, 2008). The logarithmic to linear shift has been observed to occur “broadly and abruptly” (Opfer and Siegler, 2007, p. 189) but no work so far has examined the process of this representational change longitudinally.

The existing research literature presents limitations and gaps to be addressed. First of these is a reliance on number line estimation tasks; a specific problem type in which children are asked to indicate the position of a given number on an empty number line (with only the endpoints labelled). These tasks account for a very large proportion of research carried out into the development of numerical magnitude representations (Siegler, Thompson and Schneider, 2011, p. 5) and the logarithmic to linear shift has not been conclusively documented in other contexts.

A second and well-documented limitation is the reliance on cross-sectional studies for developmental hypotheses: differences have repeatedly been recorded between children of different ages, and changes in task responses have been stimulated in short-term experimental designs (Thompson and Opfer, 2010), but there is an absence of longitudinal research into the developmental trajectories of individuals. A number of studies explicitly recommend longitudinal work (Holloway and Ansari, 2009; Moeller, Pixner, Zuber, Kaufmann and Nuerk, 2011; Thomas, Mulligan and Goldin, 2002).

A significant gap is the lack of qualitative analysis of children’s representations, particularly in their interactions with number line estimation tasks. Recent research notes this “could greatly increase the understanding of developing mental representations” (White and Szucs, 2012, p. 11). Qualitative investigation may offer valuable insights, particularly into children’s strategic application of arithmetical knowledge, which will impact on the inferences we draw from their estimations. Previous research investigating children’s *interactions* with number line estimation tasks – either qualitative or quantitative - has been limited (White and Szucs, 2012).

Another significant gap that this thesis aims to address is the potential connection between the structure of children's imagistic representations (Thomas et al., 2002) and their number line estimations, a connection that has not so far been investigated.

This thesis aims to address these weaknesses using novel combinations of methods. The key aspects of the current study which address the points outlined above are:

- Longitudinal and case study methodology that will contrast existing work by focusing on the developmental trajectories of individuals in detail
- Collection and analysis of qualitative data in order to better capture the cognitive representations demonstrated by children in all tasks
- Explicit analysis of children's strategies in number line estimation tasks, and comparison with measures of linear accuracy
- Comparison of cognitive representations in an imagistic task and in number line estimation tasks, with particular focus on the structures of number that children represent
- Not only contrasting number line estimation tasks with alternative tasks, but exploring the possible variations within the task: examining multiple ranges, and both "number to position" and "position to number" versions of the task

1.5 Justification of research topic

The study of cognitive representations of number, and specifically numerical magnitude, is important to mathematics education research. Cognitive representations of number are implicated in the understanding of some of the most fundamental concepts in mathematics: representations of numerical magnitudes are "central to understanding the meaning of number symbols (e.g., knowing that ''6'' denotes six objects), to comparing the magnitudes of numbers (e.g., knowing that six is more than four), and to estimating quantities (e.g., knowing whether there are 6, 60, or 600 candies in a jar)" (Siegler, Thompson and Opfer, 2009, p. 144). Whilst children gain knowledge about quantities and the counting number words from an early age, the development of an understanding of number – a concept of what number 'is' and how it works – requires that children form robust connections between these areas of knowledge.

The development of numerical magnitude representations is "an important educational problem" (Siegler et al., 2009, p. 144) with "important educational consequences" (Thompson and Opfer, 2010, p. 6). Many students experience difficulty developing these representations, and there is evidence that "immature numerical magnitude representations [in this case, with low levels of linear accuracy] hinder these students' learning of mathematics" (Siegler et al., 2009, p. 144). For example, studies have found that the development of increasingly *linear* representations of number is an important factor in how quickly children are able to compare magnitudes (Laski and Siegler, 2007), and their ability to learn solutions to new problems (Booth and Siegler, 2008).

As a scientific problem, the study of cognitive representations of number is an “interesting” one (Thompson and Opfer, 2010, p. 6). Despite a large number of studies in the field, there remain disagreements over key characteristics of children’s cognitive representations, and the interpretation of existing data (e.g. Cohen, 2009; Ebersbach, Luwel, Frick, Onghena and Verschaffel, 2008; Rips, Asmuth and Bloomfield, 2006; Santens and Gevers, 2008). There also remains a lack of consensus on how children’s cognitive representations of number develop and change (Berteletti et al., 2010; White and Szucs, 2012).

It is widely agreed that humans develop a concept of number from either one or both of two proto-number representational systems which are present from birth. There exist multiple views on how these systems develop into a concept of number, but there is good evidence that they influence numerical estimation even in adults, that is, even once a mature or conventional concept of number has been formed. For this reason, research into cognitive representation in estimation has high relevance for research into the origins of the number concept, a debate with not only educational implications but strong epistemological implications for the foundations of mathematics (Longo and Viarouge, 2010).

1.6 Research questions

The specific questions that the research addresses are the following:

1. In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research?
    a. What are the modes and component signs used in the representations?
    b. What aspects of number structure are represented?
    c. What are the notable between-task and within-task connections between representations?
2. What strategies can be identified in children’s interactions with number line estimation tasks?
    a. What patterns can be detected in the way children use or do not use these strategies?
    b. How do the strategies used relate to children’s estimation results?
3. How do young children’s cognitive representations of number change during their first year of formal schooling?

a. In what ways does evidence support or not support the log-linear hypothesis?
b. What is the intra-child variability of children's numerical magnitude representations in estimation tasks at different times?
c. Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?

This introductory chapter has given an overview of the research problem. The research questions posed have been situated in relation to the existing literature and key theoretical concepts. The next chapter reviews the relevant research literature in depth. This enables deeper understanding of the research problem and the unique contribution which the current research aims to make.

Chapter 2 Literature review

2.1 Overview

The literature view falls into two parts. This first part (sections 2.2 to 2.4) addresses literature relating to the theoretical aspects of the research problem. The second part (sections 2.5 to 2.10) goes on to review the relevant literature on empirical findings relating to the problem.

The theoretical aspects considered in this first section of the literature review are firstly representation, and then concepts of number and their development.

Part 1: Theoretical aspects

2.2 Representation

2.2.1 Arguing for a representation viewpoint in mathematics education research

Many authors have argued that the perspective of representation is a 'natural fit' with mathematics education due to the central role of representation in mathematics. Vergnaud (1987), for example, argues that the idea of representation is "crucial" to a theory of mathematics education precisely because representation is ever-present in mathematics itself. Kaput notes that "Most of the results that mathematicians regard as truly fundamental are easily classifiable as representational" (Kaput, 1987, p. 25), and in fact goes further: "It should be apparent that *the idea of representation is continuous with mathematics itself*" (p. 25). Representation is also what allows mathematical communication: it is the internal representations encoded "in the brains of millions of people who have studied mathematics" that enable people to "interact coherently with each other" about mathematical matters (Goldin, 2008, p. 179).

Shifting attention slightly from mathematical practice to mathematics education does not diminish the centrality of representation. Goldin asserts that representation is essential to accurately conceptualising the psychology of mathematics education: "In the context of the psychology of mathematical learning and problem solving, we must be able to consider internal configurations and structures, external configurations and structures, possible representing relations, socially shared configurations and structures, and so forth" (Goldin, 2008, p. 197). This echoes an earlier explanation from Vergnaud, who also connects the necessity of addressing representation with the unique character of mathematical knowledge and concludes that "it is impossible to do without a developmental approach to the concept of representation in mathematics education" (Vergnaud, 1987, p. 232).

Kaput notes that if anything, there is a tendency to *underestimate* the role representation plays in standard mathematics education practice. For example, it is usually assumed that "the mathematics curriculum in the first 8 years of school is about numbers, whereas the actual school work is mainly about a particular representation system for numbers – the base 10 placeholder system – and *its* properties" (Kaput, 1987, p. 20). In fact, Kaput notes, the curriculum "ignores" the distinction between properties of number sensitive to the representation system and properties that are relatively independent of it (p. 21).

Another way in which the concept of representation pervades mathematics education is in the practice of teachers. Since it is mental or internal representations that "largely determine" the usefulness of external representational systems "accordingly to how the individual understands and interacts with them", Goldin points out that in everyday practice, "effective teachers *continuously make inferences* about students' internal representations" (Goldin, 2008, p. 182).

Goldin asserts that existing literature on representation attests to "the demonstrated value of the analysis of representation as contributing on many levels to mathematics education theory and practice" (Goldin, 2008, p. 197). Working within a theoretical model in which "the mathematical development of the individual takes place through the construction of internal representational systems" (p. 184), vital work in the field of mathematics education has been done based on the idea that "children's mathematical ability can be developed *through appropriate interactions with well-designed, carefully structured task representations embodying the desired patterns*" (Goldin, 2008, p. 183).

It should be noted that a representation perspective in mathematics education can be broad. Although the position "is not a common one", Goldin's model of representation for example is comprehensive enough to include *affect* as a system of internal representation (Goldin, 2008, p. 188).

2.2.2 Defining representation

Whereas good reasons for a representation viewpoint can be clearly delineated, defining *representation* is less straightforward.

In the first place, *representation* is difficult to define in mathematics education because it inherits the ambiguities of the term in general use. The term is technically abstract, yet used and associated with a wide variety of particular meanings. Restricting the scope of the question to a cognitive perspective makes things no simpler, since it is still the case that "Representation refers to a large range of meaning activities: steady and holistic beliefs about something, various ways to evoke and to denote objects, how information is coded" (Duval, 1999, p. 3).

At the same time as being too broad in scope, the term *representation* manages to carry overly specific associations with particular research approaches. Presmeg for example, refers to *inscriptions* where other authors typically use *external representations* because of the fact that the term “*representations* … became imbued with various meanings and connotations in the changing paradigms of the last two decades” (Presmeg, 2006, p. 207).

A more fundamental reason for the difficulties associated with *representation* is that it leads directly to consideration of epistemological beliefs, and theories of mind: concepts, visualisation, perception, mathematical intuition and more. These are not matters about which there exists consensus and secure understanding; the difficulty of *representation* lies not just in terminology, but in our limited knowledge of thought-objects.

2.2.2.1 Representation, epistemology and ontology

The position of “virtually all schools in the modern philosophy of science” is a weakly constructivist epistemology coupled with belief in some external reality. The solution to the potential tension is to “acknowledge that there is a pre-given world of persons, objects, and conventional knowledge … but to adopt an agnostic, tentative position about our knowledge of this world” (Ernest, 1996, p. 340). Despite the contemporary dominance of this position, of ‘agnosticism’, it should not be denied that *representation* relates to fundamental epistemological questions. Vociferous objections have been levelled at use of the construct of representation.

Radical constructivists, who reject “on a priori grounds all that is external to the worlds of experience of human individuals” (Goldin, 2008, p. 194) cannot reconcile their position with the construct of *internal representations*. This position is well explained by von Glaserfeld, who argues that what are commonly termed *internal representations* are in fact “conceptions” (German: *vorstellungen*). These are in no sense “replicas of external originals, simply because no cognitive organism can have access to ‘things-in-themselves’ and thus there are no models to be copied” (von Glaserfeld, 1987, p. 219). Since knowledge about the external world is a priori impossible, internal conceptions can in no way be *re-presentations* of anything. The conception of mathematical structures “as abstractions apart from individual knowers” is similarly incompatible with the radical constructivist epistemology (Goldin, 2008, p. 194).

Aspects of representation are also rejected by other theoretical positions. Cognitive theorists working within a strict mind-as-computer model insist on propositional representations of all cognitive encodings, “thus rejecting any kind of internal imagistic representation” (Goldin, 2008, p. 195).

Goldin terms positions such as the above 'ideological', and makes a powerful case for the judgment that such ideologies cannot deal with all that we want to deal with in mathematics education.

> *Extreme educational ideologies often draw, tacitly or overtly, on radical theoretical or epistemological 'paradigms' whose exponents have achieved prominence in part by dismissing – often on a priori grounds – the most important constructs of other frameworks. To be clear, the frameworks I am terming 'ideological' or 'dismissive' are those where the system is closed to falsification either by empirical evidence or by rational enquiry, and/or where the fundamental tenets exclude by fiat consideration of the theoretical or empirical constructs of nonadherents.*
>
> (Goldin, 2008, p. 192)

Such ideological frameworks have value, for what they *highlight*, "focusing attention and study on particular domains of empirical phenomena, or particular sets of theoretical constructs" (Goldin, 2008, p. 196). The history of science, however, suggests strongly that "denial on first principles of the admissibility of one or other kind of construct is rarely fruitful" (p. 196). An exclusive paradigm in the end fails not because falsified but because it "leads to built-in, unnecessary limitations" (Goldin, 2008, p. 196).

Von Glaserfeld for one claims that "there can be no viable theory of representation without an explicit theory of knowledge" (von Glaserfeld, 1987, p. 215), and it cannot be denied that to use the construct of a representation entails that something is represented. However, it is a fallacy to deduce from this that we must come to a final and explicit agreement about 'what is really real'. A solution is found in pragmatic realism, as developed by Putnam (1987). What matters is that the mathematical objects being represented are real to those representing them. As Vergnaud points out, mathematical concepts, "once fully recognized and expressed, are just as real as a staircase: A function and a vector space are real objects for a mathematician!" (Vergnaud, 1987, p. 232)

2.2.3 Approaches to representation

It is vital to clarify assumptions that are made, and to distinguish between metaphor, scientific construct, and proposed accounts of 'reality'.

2.2.3.1 Goldin and Kaput

Goldin and Kaput (1996) acknowledge at the outset of their comprehensive account of representation that there is a cost in committing to any theoretical approach and that in this particular case "even the use of a term such as *representation* … may presuppose a perspective and set of commitments that some researchers are not willing to make" (Goldin and Kaput, 1996,

p. 398). For this reason, care is taken to emphasise the utility of thinking in terms of representation, and the compatibility of this theoretical approach with constructivism.

Goldin and Kaput's theory begins from a definition of representation in the most general sense. Goldin, expanding his position later, writes that a *representation* is simply "a configuration that can *represent* something else" (Goldin, 2008, p. 178). The nature of the configuration and its relation to that which it represents – which is often "bidirectional" (p. 179) – are left open.

A separation is immediately made between "internal representation" (or "mental configuration") and "external representation" (or "physically embodied configuration"). However, the authors are keen to downplay the philosophical significance of this, and state that they intend "no 'profound dualism'" between mind and matter (Goldin and Kaput, 1996, p. 402), with Goldin later describing the idea of internal representations as "an *explanatory theory* framed at a certain level of description" (Goldin, 2008, p. 181). The division is justified from a pragmatic viewpoint: external configurations are those accessible to direct observation; whilst internal configurations are "those characteristics of the reasoning individual that are encoded in the human brain and nervous system" (p. 402). Internal representations are available to neither direct observation nor introspection, at least not reliably, and observers "infer such representation from what individuals do, or are able to do, under varying conditions – i.e., from their observable behaviour, which may include interactions with observable external representations in their environments" (Goldin, 2008, p. 181)

The internal representation is presented as a valid theoretical element because it is **not** the "direct object of introspective activity" (Goldin and Kaput, 1996, p. 399). Instead, in Goldin and Kaput's theory an "internal representation" is a scientific construct, akin to 'intelligence' for example, arrived at by observations of behaviour such as verbal and gestural descriptions. The internal representations are held to be "possible mental configurations" of individuals (p. 398) but they are by definition "inferred from observations" in order to explain observable behaviour. In this way, the authors claim to have circumvented "ontological assumptions about 'the mind'" and the problems associated with this (p. 399).

Goldin and Kaput distinguish five kinds of mature, internal cognitive representational systems (Goldin and Kaput, 1996, p. 417). These, together with their most important characteristics, are summarised below:

1. **Verbal/syntactic system of representation**: this describes a person's capabilities for processing natural language. The system can represent configurations in other representational systems, and also has self-referential capability. The verbal/syntactic

system is partially formal, rather than imagistic or analogic. It is dynamic, and it is culturally provided (though Chomsky (1965), for example, would say built on an innate universal 'deep language structure').

2. **Imagistic**: there are several varieties of imagistic representation system. Important for mathematics are: visual/spatial, auditory/rhythmic, and tactile/kinaesthetic. Imagistic capabilities are necessary for meaningful interpretation of verbal statements, and encode students' nonverbal, non-quantitative (mis)conceptions. Characters from verbal and formal systems can also be treated as "objects" and processed imagistically. Internal imagistic representations are generally highly non-formal, action representations (418). They are highly individualistic, although with some apparently universal elements of structure (e.g. 'objects').
3. **Formal notational**: may be static or dynamic, and may have imagistic features to them.
4. **A system of planning, monitoring and executive control**: represents acts. In a sense this system operates metacognitively relative to the other kinds of representational system, but since all systems can represent each other and themselves, so no one system is taken to be uniquely metacognitive. It is neither imagistic nor formal. It is dynamic; partly cultural and partly individually generated.
5. **System of affective representation**: this system is neither formal nor imagistic. It seems to occur universally, and is highly dynamic.

All the above kinds of system are psychologically 'fundamental'. With the possible exception of formal notational systems, they occur universally, not only in mathematical problem solvers but in all humans (Goldin and Kaput, 1996, p. 417).

Goldin and Kaput's theory also covers representational development. Every kind of representational system develops through three stages (p. 424):

1. The inventive-semiotic stage (as in Piaget, 1969b): new characters are created or learned, and used to symbolise aspects of a prior representational system. A common problem in mathematics learning at this stage is that new characters are taken to "be" rather than symbolise the aspects of the previous system – leading to cognitive obstacles (confusion), a problem discussed at length by Duval (2006).
2. The structural-developmental stage: development or construction is driven principally by structural features of the earlier system.
3. The autonomous stage: the new system of representation, now mature, separates from the old.

The development of internal representational systems through such stages requires interaction with external representational structures (e.g. spoken language and mathematical constructions) (Goldin and Kaput, 1996, p. 424).

2.2.3.2 Contrasting account by Duval

Duval offers a definition by examples, stating that the term *representation* "refers to a large range of meaning activities: steady and holistic beliefs about something, various ways to evoke and denote objects, how information is coded" (Duval, 1999, p. 3). The term "mental representation" is used where necessary to distinguish from material or external signs. However, Duval argues that the customary distinction between mental and external representations is a "misleading division" (Duval, 1999, p. 5), since this distinction addresses only the "mode of production" of representations and not their "nature" or "form". For Duval, the more meaningful categorisation is based on precisely nature and form; he classifies cognitive representations as follows:

> *There are two kinds of cognitive representation. Those that are intentionally produced by using any semiotic system: sentences, graphs, diagrams, drawings … Their production can be either mental or external. And there are those which are causally and automatically produced either by an organic system (dream or memory visual images) or by a physical device (reflections, photographs).*
>
> (Duval, 1999, p. 5)

In summary, Duval's theory of representations holds that "the basic division is not the one between mental representation and external representation, which is often used in cognitive sciences as though it was evident and primary, but the one between semiotic representation and physical/organic representation" (Duval, 1999, p. 5).

Presmeg (2008) similarly rejects the idea that the internal/external distinction is an important dichotomy. Although key works that Presmeg draws upon make central use of this distinction (e.g. Goldin, 1992; Marcou and Gagatsis, 2003), Presmeg does not develop it in the construction of her overarching theory of visualisation. Explaining this decision, she states: "The reason for omitting this distinction is that I prefer to follow Piaget and Inhelder's (1971) claim that visual imagery (internal representation) underlies the creation of a drawing or spatial arrangement (external representation). Thus it does not seem fruitful to separate these modes of representation" (Presmeg, 2008, p. 2). This claim does not mean that we should prioritise the internal visual imagery nor see the external as merely a window onto it; it means that we should acknowledge that internal representation is involved in the process of creating the external representation, and thus that to insist on an internal/external dichotomy in cognitive representations is nonsensical.

2.2.4 Relation of representations to images

Imagery is a subject with an independent history of investigation and theorisation, with many researchers supporting the idea that mathematical reasoning is "at all levels firmly grounded in imagery" (Thompson, 1996, p. 267). For researchers such as Thompson, 'image' means "much more than a mental picture", instead something "constituted by experiential fragments from kinesthesis, proprioception, smell, touch, taste, vision, or hearing … [and] past affective experiences" (Thompson, 1996, pp. 267-268).

Thompson (1996, pp. 269-270) discusses several important theorisations of imagery, ranging from Piaget's conception of the image and its relation to mental operations (a dynamic understanding) (Piaget and Inhelder, 1967) to Kosslyn's (1980) conception of image as a representation of objective reality and eventually fixed "data structure". From Vinner comes the idea of a *concept image* (e.g. Vinner and Dreyfus, 1989), which "comprises the visual representations, mental pictures, experiences and impressions evoked by the concept name" (Thompson, 1996, p. 271). Each one of these is a broad and powerful conception of 'image'.

Galton, in the earliest known modern study concerned with imagistic representation of number in the mind, stated that he was investigating participants' capabilities for "seeing images in their mind's eye", describing his research topic as the "various ways numerals are visualised" (Galton, 1880, p. 252). Galton's language – "powers of mental imagery" and "capability" –aligns seamlessly with modern constructivist accounts of imagery, framing the participant as active generator of image, and his collected data bear clear resemblances to data collected by current researchers investigating "mental representation".

One such modern study on the imagistic representation of numbers is that carried out by Thomas, Mulligan and Goldin (2002). The study is titled "Children's representation …" and the focus is described to be "internal imagistic representations" (p. 117). The study is theoretically framed by Goldin's model of representation, in which, as described in the previous section, imagistic representation includes the sub-systems of visual/spatial, auditory/rhythmic and tactile/kinaesthetic representation. The sub-system focused upon by Thomas et al., and the sub-system that corresponds most closely to the commonly understood meaning of the term *imagery*, is the visual/spatial system.

While interpreting results within the overall framework of Goldin's theory, the study is informed by literature specifically upon imagery, primarily the work of Presmeg. The contribution of this to Thomas et al. (2002) is the categorisation of the components or sub-units of visual images, as pictorial, iconic or symbolic. In terms of *types* of visual imagery, Presmeg's work has identified five forms or categories: concrete imagery (like a picture); pattern imagery (relationships without concrete detail); memory images of formulas; kinaesthetic imagery (involving physical movement); and dynamic imagery (the image itself is moved or transformed) (Presmeg, 2006, 2008).

According to Duval there exist within the overall scheme of cognitive representations “several registers for discursive representation and several systems for visualization” entailing “a complex cognitive interplay underlying any mathematical activity” (Duval, 1999, p. 6). For Duval, ‘images’ are certainly included within conscious cognitive representations (see section 3.2.2.3 for discussion of Duval's use of 'conscious'). There exist “two heterogeneous kinds of ‘mental images’: the ‘quasi-percepts’ which are an extension of perception … [belonging to the automatic/organic system of cognitive representations] … and the internalized semiotic visualizations [belonging to the intentional/semiotic system]” (Duval, 1999, p. 6).

2.2.5 Representation in cognitive psychology and cognitive neuroscience

A large body of work on the cognitive representation of number lies outside the field of education. Much research has been carried out in cognitive psychology and cognitive neuroscience, and makes claims about the origins, development and structure of cognitive representations of number. These claims have consequences for education, and there is a potentially fruitful interaction between the fields of education and cognitive neuroscience (De Smedt and Verschaffel, 2009). Care must be taken, however, over the meaning of *representation*, and particularly *mental representation*.

The concept of representation is the organising concept in modern cognitive psychology: indeed “The central hypothesis of cognitive science is that thinking can best be understood in terms of representational structures in the mind” (Thagard, 2011, online). In terms of what the representational structures are however, and what their nature is, there remains “much disagreement” (Thagard, 2011, online). Most work “assumes that the mind has mental representations analogous to computer data structures” and cognitive theorists have “proposed that the mind contains such mental representations as logical propositions, rules, concepts, images and analogies” (Thagard, 2011, online).

Cognitive neuroscience, as its name implies, attempts to explain cognitive processes at the level of neuron activity. The experimental methodologies adopted often involve fMRI brain scanning to identify active regions of the brain during specified tasks. Theoretical neuroscientists attempt to model the behaviour of “large numbers of realistic neurons” (Thagard, 2011, online). The relation between accounts at the level of neuron models and accounts at other levels is important to stress: Thagard points out the models are “not strictly an alternative to computational accounts in terms of logic, rules, concepts, analogies, images and connections, but should mesh with them and show how mental functioning can be performed at the neural level" (Thagard, 2011, online).

Cohen Kadosh and Walsh, in reviewing the literature upon numerical representation, give an explicit definition of representation in their field. They state that "we define representation only in the general sense that is most common in psychology and cognitive neuroscience. Here representation refers to patterns of activation within the brain that correspond to aspects of the external environment" (Cohen Kadosh and Walsh, 2009, p. 314). The narrowness of this definition contrasts markedly with, say, that of Goldin and Kaput.

The constructivist accounts of Goldin and Kaput describe an active individual – **constructing** their knowledge and **using** their representations. The language of "system" and "powerful" suggests something that an individual has access to for use as a tool, whether fully consciously (using their system of planning, monitoring and executive control) or less consciously – perhaps semi-automatically using a well-rehearsed representation. This perspective of the active individual contrasts strongly with that of research focused upon the automatic, "processing" sense of internal numerical representation.

2.3 Conceptions of number

Research into young children's representation of number overwhelmingly addresses representation of the positive integers, or natural numbers, and there are clear reasons for this. The natural numbers are a very basic aspect of human life in numerate societies. It is not only the case that difficulties with number "can lead to serious impairments in everyday life" (Cohen Kadosh and Walsh, 2009, p. 313); for the majority who have access to meaningful number words (at the minimum, the ability to count in order to quantify), life otherwise is almost unimaginable. In terms of evolutionary development, the ability to count pre-dates speech as well as writing (Box and Scott, 2004). In terms of individual children, conceptions of the natural numbers are developmentally prior to, and developmentally necessary for, conceptions of other kinds of number in mathematics.

In this section I will begin with a discussion of the complications surrounding conceptions of number, and then look at how researchers have characterised 'mature' concepts of number. This is in many cases difficult to separate from theoretical accounts of how number concepts develop, so I examine conceptions of number within accounts of the development of quantitative competencies more generally. I overview key theories and give particular attention to the theme of integrating earlier conceptions.

Beyond being some kind of notion or idea, concepts are characterised by the fact that they can be shared: "concepts as constituents of thought are shareable, both by different people, and by the same person at different times" (Rey, 2005, p. 186). As such, "they need to be distinguished from the *particular* ideas, images, sensations that, consciously or unconsciously, pass through our minds at a particular time" (p. 186). Rey summarises the three main philosophical approaches to answering the question of 'what' a concept consists of: "an extension in this world, possibly an intension that determines an extension in all possible worlds, and possibly a property that all objects in all such extensions have in common" (p. 192).

Concepts are often understood as a form of mental representation. For Carey and Sarnecka, for example, "Concepts are mental representations with conceptual content, as opposed to perceptual or sensory content. Mental representations are characterised by their extensions (the entities in the world they pick out) and by their computational role (the inferences they support, the rules of combination that yield new combinations, and so on)" (Carey and Sarnecka, 2006, p. 473).

2.3.1 Why is concept of number difficult?

Philosophically, there is a lack of an agreed concept of what number actually *is*, which also complicates efforts to form a clear definition of what it means to understand number. Alain Badiou argues compellingly that our thinking on what number *is* remains no clearer than in the late nineteenth century, making our modern relationship with number almost paradoxical: "we live in the era of number's despotism; thought yields to the law of denumerable multiplicities; and yet … we have at our disposal no recent, active idea of what number is" (Badiou, 2008, p. 1). Lovell makes a similar point: "It is not generally realized, however, that although the whole edifice of modern mathematics is built upon the concept of the natural number, this concept remains something of a mystery" (Lovell, 1965, p. 26). In turning to thinkers of the past, moreover, we still do not find answers. Nineteenth century thinkers were motivated to address the question of a concept of number due to the inability of the ancient Greek conceptualisation to encompass modern uses and types of number. However, speaking of attempts by Dedekind, Frege, Cantor and Peano to define a concept of number, Badiou writes unequivocally that they "failed", since none was able to produce a unifying concept: "It is as if, challenged to propose a concept of number … thinkers reserve the concept for one of its 'incarnations' (ordinal, cardinal, whole, real ...), without being able to account for the fact that the idea and the word 'number' are used for *all* of these cases" (Badiou, 2008, p. 12).

A second consideration is that the phrase *concept of number* is a case of the particular construction 'concept of X'. Rey notes that this usage presents particular difficulties to attend to

(Rey, 2005, p. 187), since the meaning of such a construction is ambiguous. There are four immediate possibilities:

> *This could mean the concept [causality], which the child has (as do most adults); or it could mean the child's ability to deploy the concept in reasoning and discrimination; or it could mean any of the extension, intension, or rule that children associate with the English word 'causality' and its related forms; or it could mean (as in fact it very often does mean) the representation and/or standard beliefs (what I prefer to call the conception) that children associate with the extension, intension, rule or ability [causality].* (Rey, 2005, p. 187)

The general shared notion of number, corresponding to the first possibility listed by Rey, is most often indicated by the phrase '*the* concept of number'. For example: "Such people believe that the concept of the natural numbers is the result of a primitive intuition" (Lovell, 1965, p. 27). In contrast, '*a* concept of number' is frequently used in developmental research: "difficulties in assessing the cardinality of a set imply lack of a concept of natural numbers" (Rips, Bloomfield, et al., 2008, p. 630). In this example, "a" suggests that there is more than one concept of natural numbers; a child does not simply (not) have "the" universal concept *number* but develops, individually, "a" concept of number (which could then be expected to develop further still). The meaning in this case seems to best correspond to the fourth possibility listed by Rey. A final example highlights the potential of this construction for ambiguity: "for Piaget the concept of number is not based on images or on mere ability to use symbols verbally" (Lovell, 1965, p. 51). In this case, it seems the author could be discussing the general shared concept, the concept in individuals, or both. An important point of difficulty which the above discussion hints at is disagreement over the extent to which a concept includes *understanding*, or the extent to which *understanding* a concept is somehow considered a separate accompaniment to the concept itself.

As noted, the *concept of X* construction is commonly used because researchers want to write about concept development. Whilst it may sometimes be possible to say that a person 'has' or does not have a particular concept, this is not clearly true for the case of concept of number. It would be expected, for example, that a secondary school student, numerate working adult, mathematics teacher, and mathematician would have differing conceptions of number, even though each would be said to 'have' a concept of number.

2.3.1.1.1 Challenge: early conceptions with numerical content

Mix et al. (2002) argue that "In many ways, the question of whether infants possess 'true number concepts' is more philosophical than empirical because it is not clear how one would define a true number concept" (p. 21). This difficulty has already been noted. The authors continue: "In fact, it is not obvious that such a point in development ever comes because humans' concepts are continually evolving" (p. 21). Whilst it is true that humans' concepts continually evolve, a practical solution to identifying children's development of conceptions of number does still seem possible.

A solution, used by many authors if not with the same terminology, is to situate conceptions with numerical content within what Mix et al. (2002, p. 5) term an individual's overall development of "quantitative competence". In this overall development, whilst they may not have become a 'finished' concept of number, some conceptions seem to exhibit all the features necessary to be called a conception of number, whilst others do not seem to do so.

The task here is to decide upon the point at which developing numerical conceptions qualify as conceptions of number. Some nativist theories insist that representations of quantity in infancy are *concepts of number*, and differentiate later conceptions of number from these using adjectives like "fully formed" (Gelman and Gallistel, 2004). However, many researchers, like Rips et al., specifically use *numerical* rather than *number* to describe the earlier conceptions with numerical content, a usage that is followed in this thesis.

The essence of conceptions of number as opposed to numerical conceptions is the integration of knowledge: knowledge and representations of quantity, together with a system of counting. The early numerical conceptions, for example infants' discrimination between sets based on their relative numerosity, are integrated with a system of counting. The system of counting need not be a conventional one, but it must follow certain principles. These have been identified as:

1. One-to-one principle: every item should be tagged exactly once
2. Stable order principle: counting tags should maintain a consistent order
3. Cardinality principle: the final tag in a counting sequence is the numerosity of the set
4. Abstraction principle: any combination of discrete items can be counted
5. Order irrelevance principle: the order in which items are counted is not significant
   (from Gelman and Gallistel, 1978; cited by Mix et al., 2002, pp. 101-102)

As Mix et al. explain, "As long as one follows these five principles, any counting system will work, no matter how unconventional it may seem on the surface" (Mix et al., 2002, p. 102).

This integration is a view of the concept of number which has been expressed by Piaget and many mathematics education researchers since. Lovell summarises Piaget's view as follows: "the concept of number is not based on images or on mere ability to use symbols verbally, but on the formation and systemization in the mind of two operations; classification and seriation. For the concept to form in the mind these two operations must blend …" (Lovell, 1965, p. 51). Resnick's (1983) account of number development (see Representation of number structure, section 4.4.2, p.114), developed on the basis of independent empirical data, similarly supports this precise view:

"ordinal (counting) and cardinal (class inclusion or part-whole) relationships must be combined in the course of constructing the concept of number" (pp. 146-147). In Piaget's own words: "the fusion of inclusion and seriation of the elements into a single operational totality takes place, and this totality constitutes the sequence of whole numbers, which are indissociably cardinal and ordinal" (Piaget, 1969a, p. viii).

Nunes and Bryant (2009) argue that of all current theories of the development of the concept of number, the Piagetian approach remains the most satisfactory. They argue that most alternative accounts of the concept of number in fact restrict themselves by focusing heavily on either counting or reasoning about quantities. For Nunes and Bryant, it is clear that "it is only when children establish a connection between what they know about relations between quantities and counting that they can be said to know the meaning of natural numbers" (Nunes and Bryant, 2009, p. 12). The authors emphasise heavily that "Quantities and numbers are not the same thing" (p. 4) and that consequently, "the most important task for a child who is learning about natural numbers is to connect these numbers to a good understanding of quantities and relations" (p. 7). More specifically, the connection must exist in three particular ways: firstly, as cardinal number; secondly, as ordinal number; and thirdly, as cardinality understood in relation to addition and subtraction (Nunes and Bryant, 2009, p. 8). When children have begun these connections – at first only on the range of numbers with which they are familiar, they can be said to have a concept of number.

2.4 Theories of number development

2.4.1 Early numerical representations

It is widely agreed that conceptions of natural number are connected to earlier systems of representation with numerical content, present even in infants. The two commonly investigated systems are the Analogue Magnitude System (AMS), also referred to as the Approximate Number System (ANS); and the small exact number representation system (SENS), or parallel individuation system (Le Corre and Carey, 2008). I will begin by giving an overview of each system, before in the next section discussing their hypothesised roles in number concept development.

2.4.1.1 Analogue Magnitude System

The AMS is defined succinctly as a cognitive system "in which the cardinal value of a set is represented by an analog symbol that is a linear or logarithmic function of the number of elements in the set" (Le Corre and Carey, 2008, p. 651). The AMS "handles relatively large numerosities" and produces "approximate rather than exact quantity representations" (Slaughter, Kamppi and Paynter, 2006, p. 33).

Since the representation is not precise, numbers are not always easily distinguished. In experimental data, the AMS demonstrates a "numerical distance effect, in which the speed and accuracy of judgment increase with the difference between numerical values" and a "numerical magnitude effect, wherein speed and accuracy decrease with number" (Cantlon, Safford and Brannon, 2010, p. 289). The result of these effects is that the discriminability of any two values represented by the AMS is a function of their ratio, as described by Weber's law (Cantlon et al., 2010, p. 289; Carey and Sarnecka, 2006, p. 477).

In terms of its representation, the AMS is considered to "operate like a mental number line, with numerical magnitudes represented by distance travelled along the line" (Slaughter et al., 2006, p. 33). For some researchers, the AMS and mental number line are synonymous. It is described by Dehaene for example as "this core analogical representation (the 'number line')" (Dehaene, 2001, p. 16) and similarly "analogue magnitude system (or mental number line)" (Krajcsi and Palatinus, 2004).

In summary, the key features of the AMS are that the representative symbol is proportional to the represented magnitude, that the representation is approximate, that it can handle large numerosities, and that it is limited by the ratio limit on discriminability (Condry and Spelke, 2008).

There exists both confusion and disagreement over the extent to which the AMS is a conception of natural number. In Gelman and Gallistel's 2004 viewpoint piece, "Language and the Origin of Numerical Concepts" for example (with the title referring to "numerical" rather than "number" concepts), the authors discuss both "a concept of number" and "a fully formed conception of number" and use these phrases to signify different phenomena. Gelman and Gallistel are forced to make this strong distinction between "a concept of number" and "a fully formed conception of number" precisely because they give the status of "a concept of number" to the AMS. They do this despite agreeing with other researchers that the AMS is an "imprecise nonverbal representation of number" which consists of "imprecise mental magnitudes" (2004, p. 441).

To what extent is it valid or helpful to classify the AMS as "a concept of number"? If "a concept of number", it is a limited one: "'ninety' does not mean 'approximately ninety' any more than 'eight' could mean 'approximately eight'" (Nunes and Bryant, 2009, p. 14). Nunes and Bryant, like Carey and Sarnecka, highlight the fact that "representations can have numerical content and still fall short of being representations of the integers" (Carey and Sarnecka, 2006, p. 476).

2.4.1.2 Small exact number system

Whilst the existence and features of AMS are well established by research, less agreement exists regarding the second system of early numerical representation, the small exact number system

(SENS). The system is strongly supported by Carey and researchers supporting the 'bootstrapping' theory of number concept development. The SENS is also referred to as an object-file system (Feigenson and Carey, 2003), "object-based attention" (Barner, Thalwitz, Wood, Yang and Carey, 2007), and the "parallel individuation of small sets" (Carey and Sarnecka, 2006).

The SENS produces exact representations of quantities, but can only represent quantities up to three, a limit established by experimental data (e.g. Feigenson and Carey, 2003, 2005; Le Corre and Carey, 2007). Representations in this system use a representational token for each individual, and number is only implicitly represented – there is no summary symbol for the cardinal value of the set (Carey and Sarnecka, 2006, p. 478). The system is limited by set size, rather than ratio discrimination.

Two bodies of experimental evidence support the existence of the SENS. One set is box-reach tasks, in which infants are shown objects being placed into a box, and then allowed to reach in to retrieve them one at a time. The child shows by their pattern of reaching how many they expect to find there, and whilst infants "succeed at ratios of 2:1 and 3:2" they fail at ratios of "4:2 and even 4:1" (Carey and Sarnecka, 2006, p. 479).

The second body of evidence for the SENS consists of experiments in which infants watch crackers being placed, one at a time, into two tall opaque boxes, and are then allowed to crawl towards a box of their choice. Carey and Sarnecka report that "when the choices are 1 vs. 2 or 2 vs. 3 crackers, infants overwhelmingly approach the box with more crackers. But when the choices are 3 vs. 6, 2 vs. 4 or even 1 vs. 4, performance falls to chance" (Carey and Sarnecka, 2006, p. 479). These authors note that 2 vs. 3 and 1 vs. 4 are particularly interesting cases: both involve the placement of five crackers overall, and "In terms of Weber ratios 1 vs. 4 is clearly easier to discriminate than 2 vs. 3" (p. 480). However, infants fail the task as soon as either set exceeds three items (Carey and Sarnecka, 2006).

Intervention studies in which participants have been given training on either exact or approximate number representation (Kucian et al., 2011; Obersteiner, Reiss and Ufer, 2013) support the theory that number processing involves these two kinds of numerical representation. Finding "no crossover effect" between improvement in exact and approximate number processing gives weight to the theory that approximate and exact number processing "rely on distinct cognitive systems" (Obersteiner et al., 2013, p. 132).

2.4.2 From early numerical representations to conceptions of number

The exact relationship between the early systems of numerical representations (AMS and SENS) and later conceptions of natural number is not clear.

A necessary component of a conception of natural number is a system of representation able to represent the natural numbers. Carey and Sarnecka (2006) state that a system of representation can express the integers if it represents (a) the cardinal value of sets; and (b) the successor relation among adjacent cardinal values.

Condry and Spelke give an excellent summary of the two main theoretical sides in the debate surrounding natural number concept acquisition (Condry and Spelke, 2008, p. 23). According to researchers including Condry and Spelke, and notably Carey, children initially have no understanding of the logic of the natural numbers. They must construct this understanding, by building upon the two core systems of AMS and SENS. Broadly speaking, Condry and Spelke argue that the concept of a 'numerically distinct individual' arises from the small exact number system, and that the concept of 'set' arises from the analog magnitude system, which supports the quantificational system of natural language. Their experimental evidence supports the idea that the natural number concept is acquired **with** or **after** language.

The principal opposing view to this position, held by Gelman and Gallistel (1978), Dehaene (1997), and Wynn (1992), is that children have innate understanding of the natural numbers, which is embodied within the AMS. According to this view, the AMS shows all the logical features of the natural numbers, and it is 'noise' within the AMS which prevents children using exact larger numbers. Whilst children are demonstrably not able to use larger numbers at an early age, they do nonetheless understand that each large set of objects has some unknown but determinate cardinal value. Specifically, they understand that that cardinal value will change if a single individual is added or removed.

2.4.2.1 Enriched parallel-individuation hypothesis

For researchers supporting this hypothesis, it is clear that, since natural numbers show neither of the limits of the two core numerical systems (AMS and SENS), something additional is required for children to reach an understanding of the natural numbers (Condry and Spelke, 2008). This echoes the point made by Carey and Sarnecka (2006, p. 482): "none of the three systems alone has the power to represent the positive integers".

A very basic concept that infants lack is the singular/plural distinction. Carey and Sarnecka (2006) observe that in order to successfully discriminate between one and four, "infants need not represent *exactly* 4 or even approximately 4; they need only represent the set of 4 as a plurality and hence as more than 1. In other words, all they need is a singular/plural distinction" – which they appear not to have (p. 480). The researchers describe this finding as intuitively "surprising" but no longer so when the AMS and SENS systems are considered: "Neither of the core systems with numerical content includes a computationally relevant break between single individuals, on the one hand, and sets of more than one individual, on the other" (p. 481).

Condry and Spelke (2008) are confident that language is the catalyst for number concept acquisition. In their experiments with 3-year-old children, every child in the sample had mastered the meaning of at least "one", in a reasonably abstract way. The children also appreciated that a single array of toy animals that undergoes no change cannot be both *five* sheep and *ten* sheep, and that this is specific to number words (*five* sheep and *hungry* sheep is acceptable). Condry and Spelke argue that if children also possessed a full set of natural number concepts, then the two capabilities just described should be sufficient to induce that each word in the count list indicates a specific unique number. Contrary to this prediction, the experimental results find that children who have learned the meaning of the first few words in the count list still fail to appreciate that later words in the list refer to cardinal numbers. For Condry and Spelke, "children's failure to make these inferences would truly be puzzling if children possessed the system of natural number concepts" (Condry and Spelke, 2008, p. 35). These results accord with recent findings from Sarnecka and Gelman (2004) and also Le Corre and Carey (2007).

The theory above has been criticised in a number of ways. Gelman and Butterworth (2005) for example, observe that groups with restricted language, such as particular indigenous communities, still understand quantity and are able to understand number. The 'bootstrapping' role of language that Carey's theory proposes has been criticised for pre-supposing the knowledge it is supposed to develop (Rips et al., 2006; Rips, Asmuth and Bloomfield, 2008).

2.4.2.2 The alternative view: AMS is privileged

The principal alternative view, which asserts that the AMS is privileged and forms the number concept, has been endorsed by Dehaene, Wynn, and Gelman and Gallistel (1992). A key source of evidence for these theorists is that the AMS alone can support several types of numerical computation, for example the ordering magnitudes, and to some extent addition and subtraction (Barth, Baron, Spelke and Carey, 2009; Slaughter et al., 2006).

Another source of evidence, cited by Gelman and Gallistel (2004) for example, is that the AMS underpins even numerate adults' intuitive mental representation of number, a claim based largely on response times in numerical tasks. Whether the AMS is used by adults or not in numerical tasks, the participating adults already possessed a concept of natural number. Hence with regard to how the concept of number formed, this evidence is not compelling.

The main criticism of a nativist AMS-centred theory is expressed nicely by Nunes and Bryant:

However important this basic system may be as a neurological basis for number processing, it is not clear how the link between an analog and imprecise system and a precise system based on counting can be forged: 'ninety' does not mean 'approximately ninety' any more than 'eight' could mean 'approximately eight'. In fact, as reported in the previous section, three- and four-year-olds know that if a set has 6 items and you add one item to it, it no longer has 5 [sic] *objects: they know that 'six' is not the same as 'approximately six'.*
(Nunes and Bryant, 2009, p. 14)

2.4.3 When is a numerical conception a concept of number?

Nunes and Bryant do not define concept of number, but consider only the question of *understanding natural number* , a decision which refers back to the debate over what should be considered part of a concept, and what merely accompanies it (Rey, 2005, p. 187).

To be said to understand number, a child must connect the ideas of quantity and (counting) number in three specific ways: as cardinal number, as ordinal number, and as cardinality understood with relation to addition and subtraction (Nunes and Bryant, 2009, p. 8). Most accounts of children's conceptions of number, according to these authors, are too restricted; they either "leave out the number system altogether and concentrate instead on children's ability to reason about quantities, or they are strictly confined to how well children count sets of objects" (Nunes and Bryant, 2009, p. 7).

Nunes and Bryant suggest there is no simple answer to the question of how children develop understanding of cardinal numbers (2009, p. 15). Their summaries of the three leading theories are as follows:

1. The Piagetian approach: Piaget argued that understanding numbers entailed making a connection "between numbers and the relations between quantities that are implied by numbers" (Nunes and Bryant, 2009, p. 13)
2. The AMS approach: theory privileging the role of the AMS
3. The 'bootstrapping' approach: theory based upon Carey's work on 'enriched parallel individuation' (Carey, 2004, p. 65)

Nunes and Bryant argue that, of these theories, the Piagetian approach remains the most satisfactory. Piaget argued that the connection between numbers and relations between quantities was "established by children as they reflected about the effect of their actions on quantities" (Nunes and Bryant, 2009, p. 13), with the additional help of counting and other social interactions. Their criticism of the other two theories centres on their incompleteness: "Both Gelman's and Carey's theory only address the question of how children give meaning to number words: neither entertains the idea that numbers represent quantities and relations between quantities, and that it is necessary for children to understand this system of relations as well as the fact that the word ‘five’ represents a set with 5 items in order to learn mathematics” (Nunes and Bryant, 2009, p. 15). For Nunes and Bryant then, to consider how children develop conception of number without consideration of their *understanding* of number is to miss the point.

2.4.4 Integration of earlier concepts

In all accounts of the development of the number concept a common theme is the integration of earlier number-related concepts. In many examples, the implicit or explicit target is a more **abstract** understanding of number.

Some researchers specify particular aspects that must be integrated before understanding of number can be said to have been reached. For Nunes and Bryant, following the Piagetian account, these aspects are counting numbers and relations between quantities. They emphasise heavily that “Quantities and numbers are not the same thing” (Nunes and Bryant, 2009, p. 4) and that consequently, “the most important task for a child who is learning about natural numbers is to connect these numbers to a good understanding of quantities and relations” (p. 7).

Siegler et al. (2011) present an alternative and AMS-centred model concerned with the integration of numerical magnitude representations. The authors describe the developmental trajectory of numerical understanding in terms of a single arc, whose unifying element is numerical magnitude: “numerical development is at its core a process of progressively broadening the class of numbers that are understood to possess magnitudes and of learning the functions that connect that increasingly broad and varied set of numbers to their magnitudes” (Siegler et al., 2011, p. 2). Siegler et al. relate their integrated model to a proposal from Case and Okamoto (1996) that “the central conceptual structure for whole numbers, a mental number line, is eventually extended to other types of numbers, including rational numbers” (Siegler et al., 2011, p. 2).

2.4.4.1 Integration– towards abstraction?

A further relation is to the integration of types of number, in the sense studied by Sophian and Wood (1996), that is, number as written numeral, spoken numeral, or quantity in a set. Their findings suggest that if ‘number’ is a single ontological concept, it would have to integrate the above-mentioned types, which both children and adults seemed to quite reliably view as ontological categories of number, and through which participants appeared to consider instances of number.

Badiou (2008) expresses the important emerging question as follows: “is there a concept of number capable of subsuming, under a single type of being answering to a uniform procedure, at least natural whole numbers, rational numbers, real numbers and ordinal numbers, whether finite or infinite?” (p. 13). This echoes Sophian and Wood, whose results with children and adults led them to ask “whether ‘number’ is a single ontological concept or several” (Sophian and Wood, 1996, p. 355). For some, including Badiou, the answer is yes: 'number' is a single ontological category. Overall, there remains uncertainty.

Kucian and Kaufmann (2009) propose a model in which children’s mental representations of whole number magnitudes undergo, with age, schooling and development, a shift from distinct (non-abstract) to shared (abstract). This proposes that a more ‘mature’ representation of number is one in which “three”, “3” and three objects have come to share in some abstracted representation of three-ness (Kucian and Kaufmann, 2009, p. 341).

Cohen Kadosh and Walsh (2009) suggest that the development of a truly abstract representation of number, if it happens at all in some circumstances, perhaps “occurs as a consequence of the intentional processing of numbers, which leads to explicit creation of connections between different notation-specific representations” (p. 326). This appears to tie in with Kucian and Kaufmann’s model, and addresses ‘integrated knowledge of number’ from the same perspective of uniting representations of whole number magnitude from different inputs/forms.

2.4.5 Number sense and cognitive representation

The phrase *number sense* is seen throughout both education and cognitive science literature in discussions on the origin of number concepts and mathematical understanding. For this reason, it is important to gain a sense of the phrase and what motivates its use.

There is no single definition of number sense; according to some authors “no two researchers have defined number sense in precisely the same fashion” (Gersten, Jordan and Flojo, 2005, p. 296). Laski and Siegler summarise number sense as “an ill-defined construct that nonetheless is widely viewed as crucial to success in mathematics” (Laski and Siegler, 2007, p. 1723). It is not only the construct itself, but also the way in which it relates to mathematical success, that is “not well understood” (Jordan, Kaplan, Olah and Locuniak, 2006, p. 154). Indicating the complexity of the construct and its use, James Greeno argues that, in fact, “*number sense* is a term that requires theoretical analysis, rather than a definition” (Greeno, 1991, p. 170).

2.4.5.1 Education literature on number sense

Howell and Kemp (2010) identify two differing uses of number sense within education literature, firstly to describe “the intuitive understanding of number that is prerequisite for success in school-based mathematics" and secondly "the informal understanding of number displayed by children prior to formal instruction in mathematics" (p. 412). This points to two important usages, but the boundary between the two is far from definite –number sense is also used as a term intended to include both of these meanings at once, for the very reason that the difference between intuitive understanding of number and informal pre-school understanding of number, and where these understandings come from, is a point of difficulty.

Definitions of number sense in education literature commonly take a componential form; authors specify a list of attributes an individual must possess, and accordingly, assessment of number sense usually involves a composite measure. In terms of components, "most agree that the ability to subitize small quantities, to discern number patterns, to compare numerical magnitudes and estimate quantities, to count, and to perform simple number transformations are key elements of number sense in young children" (Jordan et al., 2006, p. 154). Berch (2005) demonstrates a list of thirty "Alleged components of number sense" that have been claimed for the construct at some point or other (p. 334). Drawing upon factor analysis of children's kindergarten mathematics performance, Gersten et al. (2005) judge number sense to be structured on two basic components, "counting/simple computation" and "sense of quantity/use of mental number lines" (p. 297). These coincide with Nunes and Bryant's characterisation of "the meaning of natural numbers", that is to say the integration of *counting* with knowledge about *quantities* (Nunes and Bryant, 2009, p. 12).

Laski and Siegler offer a new perspective, with the suggestion that cognitive science findings on the linearity of numerical representations could be used to provide an "operational definition" of number sense (Laski and Siegler, 2007, p. 1740). The authors observe that one interpretation of number sense is "the ability to discriminate among numerical magnitudes … and use the discriminations to constrain and judge the plausibility of outcomes of mathematical operations". Linearity of numerical representations could be used as the operational definition of this interpretation of number sense since "… it allows differentiation among numerical magnitudes throughout the range" whereas reliance on logarithmic representations "results in numbers at the high end of the range being lumped together as 'all those big numbers'" (Laski and Siegler, 2007, p. 1740). A consequence of this definition of number sense is that "people may have good number sense within one range of numbers but not within other ranges" (p. 1740). This relates to the idea that individuals do not simply have or not have understanding of number in Nunes and Bryant's definition; the development of the number concept is an on-going process.

Less precise than the above views of number sense are characterisations which openly appeal to intuition or metaphor, for example "good intuition about numbers and their relationships"

(Howden, 1989, p. 11); “a way of thinking rather than a body of knowledge and skills” (Van de Walle and Watkins, 1993, p. 142); “a well-integrated mental map of a portion of the world of numbers and operations …” (Trafton, 1992, p. 79). Kaminski notes that characterisations such as these, though vague, successfully capture “an implied comfortability, a friendliness with numbers” (Kaminski, 1997, p. 225). Research into number sense has sought to add more specific details of one child’s fluency in the realm of number compared to another’s discomfort, but there is value in retaining characterisations which capture the affective aspect.

2.4.5.2 Cognitive psychology definitions of number sense

Berch (2005) describes the existence of a “major disparity” between understandings of number sense, on the one hand as a “biologically based ‘perceptual’ sense of quantity” and on the other a “‘higher order’ depiction as an acquired ‘conceptual sense-making’ of mathematics” (p. 334). Differences in understandings of number sense are overstated by imprecise language, and I argue that in practice the biologically based ‘perceptual’ sense of quantity is infrequently taken to be actually equivalent to number sense.

Within cognitive science, the number sense construct has been primarily used and popularised by Stanislas Dehaene, for example in the seminal *The Number Sense* (Dehaene, 1997). There can often be an unclear relation between Dehaene’s number sense, and the early representation system known as the Approximate Number System (ANS – also known as the Analog Magnitude System (AMS)). Halberda and Feigenson, for example, write a paper entitled “Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults” (Halberda and Feigenson, 2008). Though this title implies an equivalence between number sense and the ANS, the relationship is never made clear – the title and closing sentence are the only two mentions of the number sense in the whole paper.

An online news report by The Telegraph further illustrates the confusion with a misleading simplification. The news article states that number sense and the ANS are equivalent: “'number sense', also known as Approximate Number System …”(Telegraph, 2011). The research paper referred to, however, clearly states that ANS is a component rather than equivalent: “One central component of the number sense is the Approximate Number System (ANS)” (Libertus, Feigenson and Halberda, 2011, pp. 1292-1293).

With this in mind, it is worth examining closely exactly what Dehaene himself writes about number sense. Dehaene writes: “‘Number sense’ is a short-hand for our ability to quickly understand, approximate, and manipulate numerical quantities” (Dehaene, 2001, p. 16) and later, on the same page, “I collectively refer to those fundamental elementary abilities or intuitions about numbers as 'the number sense'". In terms of its character, or outcomes, Dehaene's number sense as characterised here does not differ at all from the education research definitions, even down to the association with "intuition" about number. A similar view again, in passing, " … knowledge of numbers and their relations ('number sense')" (Dehaene, Dehaene-Lambertz and Cohen, 1998, p. 355) also fits comfortably with education definitions.

Where Dehaene and education definitions do differ, is in their explanation for the origins of number sense. Berch (2005) summarises the two positions, and here highlights a real and significant disparity: "With respect to its origins, some consider number sense to be part of our genetic endowment, whereas others regard it as an acquired skill set that develops with experience" (p. 334). The education research definitions in general see outcomes – the 'sense' – as emerging from a set of acquired components, which together make a 'sense' of number. Dehaene and others, in contrast, see the 'sense' as rooted almost exclusively in our biological or perceptual capability to recognise magnitude.

Dehaene writes, for example, that "number sense **rests on** cerebral circuits" (Dehaene, 2001, p. 16, emphasis added). Contrary to common readings, it is not claimed that number sense is itself the cerebral circuit. However, it is the case that specific cerebral networks underpin it: "number sense constitutes a domain-specific, biologically-determined ability" (Dehaene, 2001, p. 16). This suggests a strongly unified, non-componential understanding of number sense, which contrasts with education definitions.

2.4.6 Representational change

2.4.6.1 Conceptual and procedural knowledge

The classification of conceptual and procedural knowledge, and the relationship between the two forms of knowledge, has been a subject of various theoretical approaches. Resnick (1983) is unequivocal on the importance of this lens for number concept development: "We do not yet have a full theory to propose about exactly how practice in counting and other arithmetic procedures interacts with existing schematic knowledge to produce new levels of understanding ... [but there is nevertheless clear] active interplay between schematic and procedural knowledge" (p. 149).

Here I will look particularly at competence theories, as a background to the development of overlapping waves theory, which is the framework for much recent research on numerical magnitude representation.

Competence approaches maintain a Piagetian assumption that, at each given age, there is an underlying essence to children's thinking and that the task of developmental research is to discover this essence. In contrast to Piagetian ideas, competence approaches hold that the essences to be domain-specific rather than domain-general, and emphasise early capabilities, rather than what children lack (Siegler, 1997).

When construed as the logical knowledge needed to solve a task, competence may or may not be revealed in behaviour on a particular task. As well as this 'false negative' situation, a 'false positive' is also possible; a child may succeed in a particular experimental situation, whilst not actually possessing the logical knowledge being investigated. Consequently, a fundamental outcome of the competence/performance distinction is that in order to accurately characterise children's knowledge, it is necessary to consider not only the answers they generate but also how they arrived at those answers (Sophian, 1997, p. 282).

Competence consists of the individual steps required to complete a task, together with the knowledge which allows a child to select those steps and not other, inappropriate steps. This fact is incorporated into competence models such as Greeno's through the separation of conceptual knowledge from processes which generate a specific cognitive act (Sophian, 1997). In Sophian's presentation, competence models focus on characterising the conceptual principles that determine how cognitive processes are put together. These principles then facilitate the acquisition of task-specific procedures; developmentally, competence models suggest first conceptual principles, then later problem-solving procedures. Within a competence model of knowledge, failure to succeed at a task can have many causes, and does not necessarily imply a lack of necessary conceptual knowledge. This, as Sophian notes, insulates claims of conceptual competence from empirical verification or refutation (Sophian, 1997).

One positive outcome of this includes a great deal of experimental data from different tasks, since one remedy for this situation is to carry out detailed systematic variation, in order to try to definitively isolate the cause of task failure. Another outcome, according to Sophian (1997), is that competence models are strongly biased in favour of nativist conclusions, in effect because they cannot be disproven within the competence model. Sophian asserts that the extent to which difficulties are conceptual or task-specific ought to be an empirical question; although competence models tend to dissociate the variability in performance across tasks from the underlying conceptual knowledge, the possibility that children's susceptibility to task factors itself reflects limitations of their knowledge should not be ignored. The variations can be a crucial source of information about how children are generating responses, and hence what implications their performance has for conclusions about their conceptual knowledge (Sophian, 1997).

The important “Beyond Competence” paper (Sophian, 1997) proposes a significantly more interactive relationship between conceptual principles and problem-solving. Sophian (1997) hypothesises a bi-directional relationship between competence and performance, so that cognitive competencies both guide and are shaped by performance. She proposes that an aim should be to integrate varying findings, and that where studies conflict, it is likely that neither alone provides a full account. In particular, it is not possible to gain a full developmental picture by considering only those tasks in which children perform well.

Once it is granted that conceptual knowledge does develop significantly over time, its change mechanisms have to be inferred – since they can rarely be observed. A strategy change model proposes that children typically possess a collection of strategies for thinking about a problem situation, rather than just one, and that developmental change consists as much in changes in the choices of strategy as in the acquisition of new strategies (Sophian, 1997). Characterisation of the strategies used does not provide a characterisation of conceptual knowledge, since a strategy may be known and evaluated but not seen in use.

A competence model’s structure of first conceptual knowledge, then procedural knowledge, contrasts strongly with the theoretical accounts of Piaget and Vygotsky, and other developmental research, which all posit a more dynamic system in which change is a natural consequence of interaction with the world (whether primarily action – Piaget, or social interaction – Vygotsky) (Sophian, 1997). If interaction with the world is not held to influence developmental change, then it is unclear how to begin understanding development.

One idea is that key conceptual knowledge is present from birth, and needs only to become explicit and available (Sophian, 1997). An alternative, more recent, theory is that conceptual principles themselves change with development. Sophian (1997) notes that this is a fundamentally new idea to competence model theory. An idea from Chomsky (1959, cited by Sophian, 1997) is that internal biases and restraints focus children’s minds on some possibilities over and above other possibilities. In this account, older and younger children learn different things from the same experience, which supports the idea that constraints on learning change with development. Sophian suggests that the key may be to identify these dynamic constraints, those which are both outcomes of and determinants of further development. Siegler (1997) describes Sophian’s work here as an “important insight”, writing that “constraints cannot, in general, act as prime movers” but “must soon be supplemented by other constraints that themselves reflect the individual’s experience” (p. 327).

Sophian proposes 'goals' for the role of these constraints. In general, performance is a form of intentional action, carried out because the actor believes it will help fulfil some desire. Beliefs have been well researched and feature in competence models as conceptual principles. Desires should be interpreted as the goal or purpose of the action, and the criterion against which success or failure is measured. There is evidence that children's goals, when viewing a set of materials, affect what they remember from the materials (Sophian, 1997). Furthermore, there appears to be a dynamic relation between children's conceptual knowledge about numbers and their goal-based numerical activities; conceptual advances facilitate new goals and corresponding activities, which in turn provide the input for further conceptual advances. Sophian (1997) writes that this account needs further work, but the advantage it presents is a route out of the impasses to which innate competencies theories lead.

2.4.6.2 Overlapping waves theory

Supporting Sophian, Siegler (1997) first identifies positive contributions of competence theories. First of these is appreciating that children know more than previously believed. Second is awareness of great variation between tasks that ostensibly measure the same conceptual knowledge. Third is the realisation that age-related differences can arise from procedural as well as conceptual difficulties. The fourth and final type of contribution is constructs such as principles and constraints, which can explain why children learn some concepts more easily than others.

Siegler (1997) also identifies the weaknesses of competence approaches, which again overlap with Sophian's assessment. Five particular problems are: ambiguity about what children know of a concept; selective focus on successful performance (dismissal of significance of failures); inattention to within-task variability; failing to see the bidirectional relationship between performance and conceptual understanding; failing to specify the mechanisms of change.

Like Sophian, Siegler asserts that the bidirectional influences between performance and conceptual understanding, variability (within and between tasks), and the role of children's goals as constraints on learning, are key to gaining better understanding of development. It is pointed out that variability within children's performance on a single task is seen in all areas of cognitive development, not merely in logical tasks. Changes in thinking, when measured, appear typically to involve "ebbing and flowing" of multiple way of thinking (Siegler, 1997, p. 326). Within a framework that takes variability seriously, the extent of variability is vital, both to accurately describe cognitive change, and to understand how the change is actually occurring. Recent research supports Piaget's view that greater variability (or disequilibrium) is positively correlated with the rate of cognitive change – in other words, a high level of variability is seen when new things are being learned. The proposed explanation for this is that variability of approach allows children the opportunity to observe the consequences of using them (and hence learn).

Like Sophian, Siegler (1997) proposes that goals are a particularly important form of learning restraint. Not only do a child's own goals influence what is learned, but additionally the child's understanding of which goals a 'good' strategy for a particular problem must achieve (known as their 'goal sketch') allows unsuitable strategies to be rejected, in some cases without even trying them out. Siegler supports Sophian's view that the social world, for example parental feedback on attempts, is particularly important in shaping goal sketches.

Overlapping waves theory differs from traditional developmental theories such as stage theories and early competence theories by claiming that individuals generally know and use multiple, co-existing representations (Opfer and Siegler, 2007). The representational changes that occur within this theory are usually incremental: "children gradually increase their reliance on more advanced representations, as well as occasionally adding new representations to the mix" (Opfer and Siegler, 2007, p. 189). However, "broad and abrupt" change can also occur, due to "situations in which children are exposed to novel information that suggests that a representation that they use in other contexts yields much more accurate performance in a new, relatively similar, context" (Opfer and Siegler, 2007, p. 189).

Thompson and Siegler (2010) state that overlapping waves theory demonstrates good explanatory and predictive power with regard to numerical magnitude representation:

> *Our predictions and findings regarding heightened reliance on logarithmic representations for small numbers among children whose overall representation is linear suggest that the development of numerical representations involves trial-by trial variability, adaptive choice among representations, and knowledge-driven change like that described in overlapping waves theory. Within this theory, representations and strategies that are generally less effective continue to be used in specific situations in which they are effective.*
>
> (p. 1280)

This first section of the Literature Review has reviewed the literature relating to the research problem in terms of theoretical aspects. The research area is theoretically complex, and this section has looked at the ways in which the key ideas of representation, concepts, and the development of concepts of number have been theoretically approached, defined and researched. The different theoretical perspectives are vital contexts for the empirical research conducted into representation of number. It is only with the theoretical landscape thus reviewed that the review of empirical findings, following in the next section of the Literature Review, is able to take place.

Part 2: Empirical aspects

In this second section of the literature review, empirical findings relating to the research problem are reviewed. This involves considering findings from a range of fields of research, and the theoretical material covered in the previous section provides the necessary framework within which their varying approaches and findings can be understood.

2.5 What is represented by a representation of number?

As the theoretical section of the literature review showed, defining what is meant by concept of number is not trivial. The extent to which individuals' concepts of number become fully abstract and unify context-specific conceptions is still a matter of theoretical debate, and to complicate matters further, a given cognitive representation of number may represent only selected aspects.

The concept of number can be plausibly subdivided many ways, for example mathematically, into natural numbers, integers, rationals, etc., or by usage, as representations of ordinality, and representations of cardinality. Empirical results confirm that *number* may often not be fully abstract. Sophian and Wood (1996) found that both children and adults adopted different interpretations of "number" in answering different questions, for example as "written numeral", "spoken word" or "counting tag" separately, leading these authors to question "whether 'number' is a single ontological concept or several" (p. 355).

Strongly related to this are findings in the cognitive science literature which imply differences in how differing numerical inputs (for example symbolic or non-symbolic visual representations) are processed. Unless we can be certain that all inputs relate immediately and similarly to one abstract concept of number, the mode of numerical input in research studies remains highly important (Cohen Kadosh and Walsh, 2009). Evidence of input-sensitive results can be found in research by De Smedt and Gilmore (2011), Hubbard et al. (2009) and Mundy and Gilmore (2009), besides the examples discussed in depth by Cohen Kadosh and Walsh. An important point is that even where different notations/modalities "yield similar behavioural effects", they may not necessarily share the same representation (Cohen Kadosh and Walsh, 2009, p. 315).

2.5.1 Isolating aspects of number for research

For theoretical reasons, research may deliberately choose to investigate only isolated aspects of number. This is not a trivial task; cognitive representations can often be investigated only indirectly, and may encode aspects of number in indirect or unexpected ways (spatial, temporal, dynamic or chromatic for example). Fias and Fischer (2005) raise the interesting possibility that different aspects of number may be conveyed by different systems altogether, further complicating investigation. A hazard that arises is ambiguity regarding the relation between the isolated aspect and number itself. Fias and Fischer, for example, state that "the meaning of numbers is indeed spatially coded" (p. 52). However, the more precise statement of this assertion is actually that "**semantic representations of number magnitude** are indeed spatially defined" (p. 44, emphasis added).

A great deal of research has been carried out into the representation of numerical magnitude, and with good reason. Magnitude is a highly privileged aspect of number and, for some authors, it uniquely captures the meaning of number (Fias, Brysbaert, Geypens and d'Ydewalle, 1996). The representation of magnitude should consequently be vital for a wide range of mathematical and numerical activities, and empirical studies have verified this in children (see section 2.10: Numerical representations and mathematics).

Cohen Kadosh and Walsh (2009) give an explicit delineation of numerical representation in cognitive science; they state that "numerical representation relates to patterns of activation that are modulated by the **numerical magnitude conveyed** by the number" (p. 314). This directly addresses something often implicit in cognitive science literature, i.e. that representation of number consists primarily or even exclusively of representation of numerical magnitude. Researchers in both education and cognitive science agree that magnitude is a privileged and essential aspect of number, but it is easy to lose sight of distinctions. Where the hypothesised relationship is not clearly stated, it can be unclear what assumptions have been made about the relationship between magnitude and number itself. In the case of Cohen Kadosh and Walsh's review, it is implied that representation of number consists largely of representation of magnitude. It is not explicitly said that nothing else influences the representation of number, but nothing beyond magnitude is openly considered.

2.6 Number and space

The spontaneous association of numbers and space is seen across cultures (De Cruz, 2012). This association may manifest itself in a wide variety of ways, with mappings to everything from abacuses to parts of the body, and "space-number association appears to be a human universal" (De Cruz, 2012, p. 138). Furthermore, the association occurs in both intentional and automaticised representations.

Research in the past twenty years has found strong links between spatial coding areas of the brain and the representation of number (Fias and Fischer, 2005). This confirms a connection between visuo-spatial and numerical thought that has been long been present in self-report and anecdotal evidence, for example Galton's (1880) investigation of participants' mental images of number. Like the images reported by Galton (1880), Fias and Fischer (2005) report that the spatially coded representations of number in their study were "mostly automatically activated, were stable in time and had emerged in childhood" (p. 43). A great deal of research attention has subsequently been paid to the exact spatial associations of number representations. According to some, the spatial aspect is in fact dominant: "The non-verbal representations of quantity appear to be largely spatial … though other sensory modalities also seem to be included" (Siegler et al., 2011, p. 4).

With regard to the specific details of the representations, Fias and Fischer (2005) report that adults' visuo-spatial representations of numbers are "predominantly oriented from left to right" (p. 43). However, these authors also find that "numerical information can be dynamically allocated to different representationally defined reference frames" (p. 49), with the "left-right line-like spatial coding being merely a default" (p. 49). The key finding is that "spatial cognitive representation of numbers should not be considered as fixed and unchangeable … the characteristics of spatial number coding are largely determined by numerical and spatial parameters specific to the task at hand" (p. 44).

Since the left-right or right-left orientation of spatial-numerical association is culture specific (Dehaene, Bossini and Giraux, 1993), it has often been supposed that the association is linked to children's beginning reading. Opfer, Thompson and Furlong (2010) however found evidence of spatial-numerical association "long before children begin formal reading instruction" (p. 769). Interestingly, the results supported neither the idea that spatial-numeric associations arose as late as reading acquisition, nor the idea that they arose as early as numeric symbol acquisition, but instead suggested that the spatial association is formed at some time in between these two stages. With regard to the value of the spatial-numeric associations, it was found that found that "pre-schoolers showing spatial-numeric associations … displayed more mature, linear representations of symbolic value" compared to preschoolers of the same age lacking spatial-numeric associations (p. 769). Furthermore, the children showing robust spatial associations performed markedly better than others on those tasks where accessing representations of numerical magnitude (as opposed to merely reciting the number list) was required (p. 769).

2.6.1 Mental Number Line

The left-to-right spatial representation associated with numbers is generally known as the *mental number line*. Precise interpretations of the *mental number line* vary; they include the line as metaphor (Fias and Fischer, 2005), as mental imagery, or as synonymous with the AMS (Dehaene, 2001).

For Siegler et al. (2011), a mental number line is where "number symbols (e.g., ''7'') are connected to non-verbal representations of quantity in an ordered, horizontally-oriented array" (p. 4). According to Fias and Fischer (2005), the mental number line is a "useful metaphor" (p. 44), but there exist numerous questions about its nature and utility. The usual orientation - the "left-right line-like spatial coding" - is "merely a default" (p. 49), and indeed, the evidence of the flexible nature of spatial associations with numbers "challenges the appropriateness of the number line metaphor" (p. 52). In addition, it remains uncertain "whether the mental number line is a single, analogue continuum" or whether instead, there exist separate mental representations for the single- and multi-digit numbers (p. 46).

Dehaene et al. (2008) state very confidently that "the mapping of numbers onto space is a universal intuition" and that "this initial intuition of number is logarithmic" (p. 1217). In contrast, the "concept of a linear number line" appears to be "a cultural invention that fails to develop in the absence of formal education" (p. 1217).

Thomas, Mulligan and Goldin (2002), investigating children's imagistic mental representations of number with a focus on structural elements, do not infer anything so specific as a 'mental number line' from their results. Where a number line is externally represented by a participant, it is considered one of a larger set of "numerals drawn in various formations" (p. 121). The authors infer *structural aspects* of the participant's cognitive representation and conceptions of number from the external line, from its particular structural features (for example, markings at decade intervals).

2.7 Representational changes

A substantial body of research concludes that children's cognitive representations change from a logarithmic placement of numbers on a mental number line, to a linear placement (see for example Berteletti et al., 2010; Dehaene et al., 2008; Opfer and DeVries, 2008). Not all authors agree with these conclusions; some assert that there is still room for debate as to whether the mind inherently maps numbers onto space at all, as well as with what placing (Cantlon, Cordes, Libertus and Brannon, 2009). Another alternative proposition retains the mental number line, but in segmented form – composed of differently scaled sections. Ebersbach et al. (2008) propose a model in which the section break occurs at the end of an individual child's familiarity range, as measured by a counting exercise. Another suggestion is that the section break is located at 10, linking to children's difficulty or different processing of two-digit numbers and their magnitudes

(Moeller, Pixner, Kaufmann and Nuerk, 2009). Both segmented models are rejected by Thompson and Opfer (2010) on the basis that, for individual children, any inferred “change point” between linear sections of the mental number line varied wildly across different tasks; the same child placed numbers differently and showed differently scaled representations depending on the boundary conditions of the task.

2.7.1.1 Logarithmic to linear placement

Siegler et al. (2009) put together a convincing argument for the logarithmic to linear change, citing studies that appear to deduce the representational change occurring from multiple task contexts and upon multiple numerical ranges. The primary source of evidence is a large body of cross-sectional studies using number line estimation tasks (e.g. Booth and Siegler, 2006; Geary, Hoard, Nugent and Byrd-Craven, 2008; Laski and Siegler, 2007; Opfer and Siegler, 2007; Siegler et al., 2009; Thompson and Opfer, 2010).

Particularly strong evidence is provided by a study in which both adults and children were asked to estimate the position of salaries on a number line, where the salaries were expressed as fixed-numerator fractions (Opfer and DeVries, 2008). In this task, children outperformed adults in accuracy; both adults and children focused on their attention on the fraction denominators, but children’s tendency to use a logarithmic representation of number provided them with a task advantage, since the relation of denominator size to the fraction’s magnitude resembles a logarithmic function much more closely than a linear one (as used by the adults in the study). The study thus provided evidence of logarithmic to linear representational change independent of any general ‘improvement with age’, a common criticism of developmental studies’ significance (Opfer and DeVries, 2008).

Evidence from tasks other than number line estimation tasks has also been found. Booth and Siegler (2006) identified the log-linear shift across four kinds of pure numerical estimation problem, and suggestive evidence for the same shift occurring in number categorization tasks has also been found (Laski and Siegler, 2007).

Whilst logarithmic representations of quantities are “widespread among species and age groups” since useful “in a great many situations”, the shift towards a linear representation of numerical magnitude is considered desirable since “In the formal numerical system … magnitudes increase linearly rather than logarithmically” (Opfer and Siegler, 2007, p. 172). Opfer and Siegler conclude that children’s logarithmic representations of numerical magnitude are “understandable” but that in school and modern life can “interfere with accurate estimation” (p. 172).

The log-linear model is seen as preferable to a segmented linear model for both theoretical and empirical reasons. Amongst other things, the segmented model comes purely from number-line estimation task data, whereas a logarithmic placement is implicated by multiple situations: "many tasks indicate that the function translating objective numeric quantities into a subjective number is logarithmic" (Young and Opfer, 2011, p. 59). These tasks include choosing the more numerous of two sets of dots, and comparisons of Arabic numerals – in both cases, results follow Fechner's law ($y = k*\ln x$) (Young and Opfer, 2011, p. 58). The segmented model may have appeared to accurately describe experimental results, due to the particular effect of a mixed population of logarithmic and linear-placing participants: "in a hypothetical mixed population made entirely out of linear and logarithmic subjects, averaged together, the logarithmic model does not provide a better account until the population is over 70% logarithmic, whereas the segmented linear model retains a strong fit across all mixtures *even though no responses were generated by a segmented linear function*" (p. 60).

2.7.1.1.1 What is the change that happens?

The hypothesised log-linear change is most commonly referred to as a "shift", in which an individual moves from primarily using a logarithmic representation of magnitude to primarily using a linear one, on a given range of numbers. The shift has only been deduced with respect to specific ranges of numbers, and an individual commonly continues to use a logarithmic representation on a larger range (e.g. 0-10,000) long after 'shifting' to a linear representation on a smaller range (e.g. 0-100). Overlapping Waves theory frames the logarithmic to linear change as a change in *choice* of representation.

To better understand the change, it is worth surveying the language used to describe it, for example in Thompson and Opfer's (2010) lengthy and in-depth investigation of the change process. Their choice of language corresponds to the language used by other authors writing about the logarithmic-linear change in the framework of overlapping waves theory. They describe the change between logarithmic and linear representations using words from five main categories, shown here with their frequency:

1. **Apply**, **use** [7 times]. E.g. "children's application of linear representations to large numerical scales" (Thompson and Opfer, 2010, p. 16); "children's use of logarithmic representations appeared unchanged" (p. 25)
2. **Generate**, **produce** [3 times]. E.g. "children immediately generated a linear series of estimates for all other numbers in the 0-1,000 numeric scale" (p. 9); "produced logarithmic estimates" (p. 25).

3. **Generalise** [7 times]. E.g. “generalizing a linear representation of numbers to much larger orders of magnitude” (p. 11).
4. **Scale up**, **extend** [6 times] “scale up their linear representations” (p. 10); “extending linear representations of number to ever-larger numerical scales” (p. 25).
5. **Shift**, **switch**, **abandon** [9 times], e.g. “abandon use of a logarithmic representation” (p. 25), “a logarithmic-to-linear shift” (p. 12).

In addition, there were single instances of other terms to describe the change in representation: “adopt” (Thompson and Opfer, 2010, p. 4), “change” (p. 16), “bootstrap” (p. 17) and “transferred” (p. 24). A similar survey of the language used by Opfer and Siegler (2007) reveals the four most commonly used terms to be the following:

- “shift” [7 times]
- “transition” in use of representation [3 times]
- “apply” [twice]
- “extend” a representation [8 times]

These terms carry a wide variety of implications, which perhaps suggests that more research is needed to accurately describe the change.

2.7.1.1.2 Mechanisms

The change from logarithmic to linear representation can occur rapidly – between tasks in a single experimental session, for example. In studies such as Opfer and Siegler (2007), this change occurred in response to feedback from the researchers designed to highlight the discrepancies between logarithmic positioning and the desired linear positioning. Opfer and Siegler conclude that the logarithmic to linear change can be “strikingly abrupt” (Opfer and Siegler, 2007, p. 169).

Thompson and Opfer’s (2010) investigation into the mechanisms of the logarithmic-linear change indicates an important role for analogy. Analogy is investigated by these authors because it has the potential to “reconcile two sets of seemingly contradictory findings” (p. 4). These are, on the one hand, the “slow rate of representational changes observed in cross-sectional studies” such as Siegler and Opfer (2003), who recorded the “logarithmic-to-linear switch in numeric representations between second and fourth grade [ages 7-10] on 0-1,000 number line problems”, and on the other hand “one-trial representational changes” evidenced in microgenetic studies such as Opfer and Siegler’s 2007 study, noted above, which demonstrated children’s adoption of linear representation after being presented with “maximally discrepant feedback” (Thompson and Opfer, 2010, p. 4).

The analogy functions through aligning contexts, so that children are invited to make a generalization and extend their representations. Thompson and Opfer (2010) "aligned contexts in which children were familiar (e.g., 0-100) with larger, less familiar numeric contexts (e.g., 0-1,000 to 0-100,000)" and this "apparently prompted second graders [ages 7-8] to scale up their linear representation of numbers" (p. 26). It did this by "highlighting the underlying structure of the decimal system" thereby prompting children to "bootstrap the linear representation they already possessed in a familiar numeric context (0-100) to less familiar numeric contexts" (p. 26).

The theoretical underpinning of the analogy explanation is structure mapping theory, according to which "children form analogies by aligning representational elements between a base and target domain. This alignment process facilitates transfer of information from base to target through children's comparison of surface-level features. This comparison process leads to subsequent highlighting of common underlying relational structure shared by base and target" (Thompson and Opfer, 2010, p. 7). In a more general context, there is evidence that an increase in the linearity of mental representations of numerical magnitude can be stimulated by increased interaction with external linear representations of number, for example board games (Ramani and Siegler, 2008).

The idea that the non-linear placement is *incorrect* is presented as central to the proposed account of how and why the logarithmic to linear change occurs. The driving force behind the "age-related trend" is posited to be "the inaccuracies produced by the logarithmic representation, together with extensive experience that children have with some estimation tasks and numerical ranges" (Opfer and Siegler, 2007, p. 172). Or, in other words, "discrepancies between children's estimates and the linear function … provoke the realisations that the underlying representation is wrong and that a new way of thinking about the task is needed" (Opfer and Siegler, 2007, p. 173). In a more sceptical reading, the proposed change mechanism requires only that children learn the number line estimation task response hoped for by researchers and other adults.

Both Thompson and Opfer (2010) and Opfer and Siegler (2007) stimulate representational change through the use of feedback between trials. The feedback consists of showing participants where number line marks 'should' be – according to a linear representation. This feedback consists of the researcher signalling to participants that the task is actually about the ability to produce a linear representation of number, and in response to this additional task knowledge, children do respond by switching to and extending their linear representations. The conclusions to be drawn from this finding are not clear cut; the studies do show how representational change can be triggered, but the social factor is an important factor that needs greater attention. From the perspective of this thesis, the most accurate wording is the reference to "Children's learned expectations that numerical magnitudes increase linearly" (Thompson and Opfer, 2010, p. 6). This highlights accurately the nature of the new knowledge ("learned expectation") and leaves open the question of 'correct' and 'incorrect' representations.

2.8 Strategy in estimation and its relation to cognitive representation

Siegler et al. (2011) draw attention to a number of assumptions made in the empirical investigation of cognitive representations of number, particularly that representation is an automatic process and separate from conscious control. This is often revealed implicitly in the design of studies: "The implicit assumption is that people invariably use a particular representation of numerical magnitudes and that the research task is to determine the characteristics of that representation" (p. 7).

The assumption of automatic activation is sometimes explicit, with Dehaene (1997) for example describing logarithmic representations of number magnitudes as occurring "'like a reflex' that cannot be inhibited" (cited by Siegler et al., 2011, p. 7), and Fias and Fischer's (2005) assertion that "the spatial coding of numbers … occurs automatically" (p. 44). This language contrasts strongly with Goldin and Kaput's (1996) account of representation, which uses the language of a tool, with capabilities, which the individual accesses.

As Siegler and colleagues (2011) note, "reviews of the literature on whole number magnitude representation … typically do not even mention strategies or strategy choices" (p. 7). Two suggested reasons for this absence are firstly that strategies may be actually unimportant in processing numerical magnitudes (if, for example, they are processed purely automatically by the ANS); and secondly that more research attention is given to strategies where they are more obviously in use, for example for fraction processing. There is evidence that fraction processing is "slower and under greater voluntary control" and thus "characteristic of tasks on which people can accurately report strategy use" (p. 7). Research on the representation of fractional magnitudes has tended to confirm that strategy is an important variable (see for example Sophian and Madrid, 2003).

An important source of evidence for the automaticity of magnitude representation is that estimation accuracy appears to be unaffected by the length of time permitted. For example, Siegler et al. (2011) cite results showing that "number line estimation with whole numbers is no less accurate under time pressure than without time pressure" (p. 7). Another source of evidence is the so-called distance effect in comparison tasks, in which numbers closer together elicit a slower response time (Dehaene and Akhavein, 1995).

Despite this evidence for automaticity, a number of results suggest that aspects of strategy could be involved in magnitude representation. Geary et al. (2008) found “greater variation in children’s use of one representational system or the other” than predicted from the literature (p. 293). Specifically, “for some trials children made placements that implicated use of a linear representation and for other trials they made placements that implicated use of the natural number-magnitude representation” (p. 293).

Siegler and Opfer (2003) adopt the language of ‘choice’ in their paper showing evidence for multiple representations. They conclude that “over a wide age range, people possess multiple numerical representations, with choices among representations changing with age and experience” (p. 242). Whilst the awareness of the individual regarding their choices is not discussed, the idea does run firmly against the idea that individuals possess exactly one automatically activated representation of number magnitude at any given time. Individual strategy is similarly referred to by Siegler and Booth (2004)

Cohen Kadosh and Walsh discuss a slightly different perspective on individual capabilities and strategic choice in representation of number. They hypothesise that an abstract sense of number, if it develops, could be a “consequence of the **intentional processing** of numbers, which leads to **explicit creation of connections** between different notation-specific representations” (Cohen Kadosh and Walsh, 2009, p. 326, emphasis added). They argue that “humans do not, as a default, represent numbers abstractly, but can adopt strategies that, in response to task configuration and demands, *can* create real or apparent abstraction” (p. 326). The positioning here of the individual as actor makes this perspective much easier to reconcile with constructivist-informed accounts of representation in education literature.

Whilst strategy use and conscious control may appear to be very closely related, it is important to acknowledge that strategy may also occur within automatic level processes. Thompson and Siegler (2010), for example, conclude that the development of numerical representations involves “trial-by trial variability” and “adaptive choice among representations”, as hypothesised in overlapping waves theory. Crucially for this discussion, “The mechanism that produces these adaptive choices is viewed as unconscious, and its workings have been illustrated in computer simulations that generate strategy choices highly similar to children’s” (p. 1280), see also Opfer and Siegler (2007).

The dichotomy between representation as tool or as automatic reflex finds an echo in the two primary views of number sense, described by Berch (2005) as the “’lower order’ characterisation of number sense as a biologically based ‘perceptual’ sense of quantity” and “a ‘higher order’ depiction as an acquired ‘conceptual sense-making’ of mathematics” (p. 334).

2.8.1.1 Strategy in estimation tasks specifically

Number line estimation tasks form a very large part of research into cognitive representations of number (see Methodology chapter). In brief, this is because estimation tasks require the translation of numbers or quantities between different representations. As the previous section describes, the dominant interpretation of results from these tasks has been that they reflect an automaticised mental representation of magnitude; however, recent research has begun to identify aspects of strategy. Children’s use of reference points, evidenced in a variety of ways, is an important example.

One account of number line estimation that includes children’s use of reference points is the proportion judgement account. Barth and Paladino (2011) argued for a model of number line estimation in which children estimated by making proportion judgements using reference points such as midpoints and the line endpoints. This interpretation of number line estimation removes the need to account for a qualitative log-linear shift in how children estimate, and also has a convincing psychological basis in more general psychophysical models of proportion judgement. These general models are derived from Steven’s power law, and the inclusion of one or more reference points changes the mathematical model to a one- or two-cycle power curve. Barth and Paladino fitted children’s estimations to power curve models incorporating fixed reference points (e.g. midpoint 10, on a line from 0 to 20), and concluded that their proportion judgment account offered “at least five advantages over previous accounts of number-line estimation” (Barth and Paladino, 2011, p. 134). The advantages are firstly that the account “is motivated by the structure of the task”, secondly “has been modelled and validated in other domains, with many tasks, in children and adults”, thirdly “makes specific, testable predictions” which found support in their data and were not explained by previous accounts, fourthly that the account was able to explain both the linear estimations of older children and the “roughly logarithmic-appearing” estimations of younger children, and finally that it explains why the log-linear shift is not observed in many other estimation tasks (Barth and Paladino, 2011, p. 134).

Power-curve modelling has been questioned, with Opfer, Siegler and Young (2011) for example offering a convincing response to Barth and Paladino (2011) which argues that the success of power curve models in the proportion judgement account is owed to noise-fitting. A further weakness from the point of view of this study is that Barth and Paladino’s study included feedback on the location of the midpoint during children’s introduction to the number line estimation task, thus introducing one of the reference points later examined. However, further options also point towards children's use of reference points.

Ashcraft and Moore (2012) carried out an unusual number line estimation study designed to compare the log-linear and proportion judgement accounts, and found clear support for the hypothesis that children begin to use an inferred midpoint when carrying out their estimations. The study was unusual in that it tested only position-to-number estimation, as opposed to the much more commonly investigated number-to-position estimation. Ashcraft and Moore concluded that their results were "largely consistent" with the logarithmic-to-linear shift reported by Siegler and colleagues (Booth and Siegler, 2006; Siegler and Opfer, 2003), whilst there was only some support, and indeed some "problematic" aspects, with regard to the proportion judgment account. The specific findings about the accuracy of children's estimations were that estimates near the origin of the line were "always" highly accurate, and that "This point is then joined by accurate estimates at the endpoint of the line and, with increasing knowledge of arithmetic, by the midpoint of the line." (Ashcraft and Moore, 2012, p. 265) The authors deduced from these findings that "first graders [aged 6-7] perform their estimates in an 'origin up' fashion, beginning at the origin of the line and working forward to the hatch mark, with increasing errors and latencies as the location to be estimated gets farther from the origin." (p. 266). Second graders, on the other hand, aged 7-8, "seem to be working from the end of the line closest to the hatch mark … in a strategically economical and accurate fashion". The final stage occurs with arithmetical knowledge of the midpoint, at which point "nearby points can be estimated from that landmark as well". It is worth emphasising that no qualitative data to investigate these hypothesised strategies were collected.

An important source of evidence for children's use of reference points has been developed by Schneider et al. (2008) and their work analysing gaze patterns using eye-tracking. For children aged 7-9 years (but not older and younger children), recorded eye movements were correlated with manual estimation results and supported Petitto's (1990) finding that children used count-up from left endpoint and reference to an inferred midpoint. Eye-tracking was also used by Heine et al. (2010), who investigated the connection between eye-movement and the position of children's estimations. They found, in line with other research into implicit knowledge (for example gesture in the work of Goldin-Meadow and Alibali, see Garber and Goldin-Meadow (2002); Perry, Church and Goldin-Meadow (1992)) that eye-movements revealed " a qualitative change in children's implicit knowledge about numerical magnitudes in this age group that precedes the overt, that is, behavioural, demonstration of explicit numerical knowledge" (Heine et al., 2010, p. 175). A particularly interesting aspect of the study from the point of view of this thesis is that Heine et al.

excluded children's first fixations from the data analyses, since children fixed their eyes on the left endpoint, midpoint or right endpoint "significantly more often … than could be predicted from the set of stimuli" (p. 180). The authors interpreted this as evidence of children's initial orientation, and excluded the eye movements in question from further analysis, but in the terms of this thesis children's initial orientations and their specific details are highly interesting.

Further eye-tracking research has been carried out by Sullivan et al. (2011) to investigate the number-line estimation processes of adults. Although this does not shed direct light on the processes employed by children learning about number, the findings, which strongly support Barth and Paladino's proportion judgement model, are of note. Sullivan et al. (2011) found that adults' estimations demonstrated "patterns of error predicted by psychophysical models of proportion estimation" (p. 562) consistent with Barth and Paladino's work with children. Their results were not however predictable using the alternative proportional-reasoning strategy mentioned by Siegler and Opfer (2003) which predicts reduced variability of estimations around anchor points but does not predict the particular direction of errors. Sullivan et al. also found evidence suggesting that adults' estimation processes are "dynamic", a finding consistent with previous studies in which participants "adjusted estimation strategies to incorporate information about numbers to be estimated" (Sullivan et al. (2011, p. 561), see also Izard and Dehaene (2008) and Sullivan and Barner (2010)).

Important recent work on representational change and strategy has been carried out by White and Szucs (2012). The purpose of their contribution is to go beyond investigation of the potential log-linear shift, and develop further insight into children's estimation strategies. White and Szucs find that children in Year 1 (aged 5-6 years) "did not demonstrate any clear anchor point application" and it is suggested this is "because they were limited to counting strategies and were unable to link the numerical value to the spatial cues provided by the number line" (p. 9). Children in years 2 and 3, however, did provide evidence of "more flexible strategies, and use of anchor points, that utilize decomposition and part-whole relations" (p. 9). White and Szucs' results lead them to conclude that "specific numbers could exhibit unique behaviours as a function of the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy" (p. 9). Importantly, the authors reflect that this flexibility should give cause to question the validity of the usual modelling methods: "given the flexibility of strategy application, is it in fact meaningful to try and model the mental representation of numbers using a fixed linear/logarithmic model?" (p. 9). The potential for unique behaviours, as described above, "represents a limitation of the linear/logarithmic modelling approach" (p. 9).

With regard to the log-linear hypothesis investigated as part of the study, White and Szucs (2012) find that a linear model "dominates from Year 2" onwards, but neither a logarithmic nor linear model is compelling for younger children (p. 9). The authors suggest that children in years 2 and 3 were "probably" using the above-mentioned strategies, and that further research should make this a focus for investigation (p. 10). Such scrutiny of estimation strategies themselves, in combination with statistical modelling, "could greatly increase the understanding of developing mental representations" (p. 11).

The explanations hypothesised by White and Szucs have been hinted at by much earlier work, namely Petitto (1990) and Newman and Berger (1984), as acknowledged by Schneider et al. (2008). Specifically, both studies identify a change from inflexible counting-on strategies used by younger children, to more sophisticated use of counting and incorporation of midpoints by older children. Neither study is fully comprehensive. Newman and Berger carried out only 21 trials per child, and investigated strategy use for only three trials. More importantly still, strategy use was investigated only using self-report data, a serious limitation, especially considering the nature of the estimation process as evidenced by more recent studies.

Petitto's (1990) results revealed "qualitative changes in children's strategies over the four grade levels tested", and these changes "indicate a shift from simple unidirectional counting by 1s in the earliest grades to the use of midpoint values and alternative counting intervals by the end of third grade [age 9]" (p. 70). Petitto concluded that the changes were related to "incremental acquisition of component skills" and that later strategies incorporated elements of proportional reasoning not seen in younger children's estimations. In terms of the process of strategy change, the results suggest two phases in strategy change; firstly, "a drop in the use of inappropriate and ineffective modes of counting" followed somewhat later by "an increase in the use of new strategy components (i.e. counting by 5s and 10s) and using the value of the midpoint" (p. 70).

The principal limitation of Petitto's study is that findings on strategy were based only on in-the-moment observations of children's behaviour. For this reason, rapid changes in strategy, and aspects only represented in gaze or gesture for example, are unlikely to have been captured. The study also shares the limitation of using very few trials per child – only six. In terms of investigating changes in strategy, it also finally shares the limitation – along with many numerical representation studies - of a cross-sectional design (see Methodology chapter).

2.9 Imagistic representations of number

As the previous section describes, cognitive representations of numerical magnitude in particular have received a good deal of attention from researchers in cognitive science and development. A separate body of relatively recent work has considered the imagistic representation of number more holistically, using methodologies from Education research. Findings from this body of research share important themes with findings from the cognitive science research, but in addition add new dimensions of understanding.

The methodologies of this research share the use of tasks designed to stimulate children's visual imagery relating to number, and qualitative analysis of this imagery. This clearly complements research previously described. In common with experimental studies in the cognitive science tradition, education research into children's representation of number has put particular emphasis on the structure of children's representations and what this reveals. A principal aim has been the inference of children's internal representations from their external representations.

Vivid imagery relating to natural number is a common phenomenon, reported in a range of empirical studies (Galton, 1880; Thomas, 1992) as well as many anecdotal accounts. Thomas, Mulligan and Goldin (1994) carried out a large cross-sectional study of children in Grades K to 6 (ages 4-12 years), and concluded that children's internal representations of number were very highly imagistic. The extent to which children represented aspects of number structure was variable, as were the ways in which they did so. Interestingly, whilst only 3% of children in a general sample (n=166) demonstrated dynamic images of the number sequence, in a separate sample of children assessed as high-attaining (n=79), 29.9% demonstrated dynamic imagery.

An exploratory study by Thomas and Mulligan (1995) focused on structural aspects of number and number representations, particularly motivated by the low understanding of base-10 structure shown by many children. Children aged 10-12 took part in structured task-based interviews, which allowed the researchers to makes inferences about children's cognitive processes during the construction phase of representations. On the basis of this study, the authors conjecture that children's cognitive representations of number do "give clues about their structural development of the numeration system" (p. 21). They found a wider diversity of imagistic representations than expected, and a higher percentage of dynamic imagery in middle/lower attaining children than expected from the results of Thomas et al. (1994). Children who demonstrated dynamic imagery of number did however achieve higher results on a numeration test than those not demonstrating dynamic elements. The study is limited by necessarily capturing only a "partial description" of children's representational capabilities; other prompts, activities, or follow-up questions for example could potentially have led to different results. The authors conclude that further research is needed "to shed light on how children construct their personal numeration systems, and how they structure them over time" (Thomas and Mulligan, 1995, p. 21). On the basis of the 1995 results, it was concluded that more structurally developed internal representation of counting numbers correspond to more coherent external representations and a wider range of numerical understandings (p. 22). An hypothesis to be tested by further, longitudinal work, is whether children "who have access to several forms of imagery with which to represent their internal structures, will be more capable of developing a relational understanding of the numeration system" (p. 22).

A follow up study that aimed to partially address the question of the development of children's representations was carried out by Thomas, Mulligan and Goldin (2002). As in their 1994 study, two samples were used, a general sample and a separate sample of children assessed as high-attaining. A limitation shared with the majority of number representation studies in both cognitive science and education is the use of a cross sectional design to investigate development trajectories. The authors stress that "Data taken in just one or two interviews per child do not permit us to trace the process of construction of internal representational system longitudinally in individual children" (Thomas et al., 2002, p. 129). In addition, "Our methods of inferring aspects of children's internal representation from their externally produced representations are still exploratory, and not yet subject to tests of validity or inter-researcher reliability" (p. 130). However, the study nevertheless provides interesting additional evidence of children's dynamic imagery and initial identification of stages of imagistic representation. The stages include the following:

> *… inventive-semiotic acts, of initially assigning imagistic meanings to or identifying imagery with mathematical words and symbol configurations; structural development acts, associated with sequences of numbers, groupings by tens, recursive grouping, and other mathematical structures; and autonomous acts, in which insightful, mathematical meanings for numerals are freely and flexibly found in new contexts, distinct from those used initially in constructing the numerations system.*
>
> (Thomas et al., 2002, p. 130)

These stages quite clearly identify ways in which children's increasing arithmetical knowledge and number understanding are connected with changes in their representations of number.

2.10 Numerical representations and mathematics

As previously noted, cognitive representations of number are assumed to reveal aspects of conceptions of number, and to play a vital role in mathematical activities: "From understanding the meaning of number symbols (e.g., knowing that ''6'' or ''six'' denotes six objects), to comparing the magnitudes of numerals (e.g., knowing that ''six'' is more than ''four''), to estimating quantities (e.g., knowing whether there are closer to 6, 60, or 600 candies in a jar), children must map between alternative quantitative representations” (Young and Opfer, 2011, p. 59).

The involvement of cognitive representations of number in learning and doing mathematics is one of the key reasons for interest in researching them. In this section, I review recent empirical evidence on the ways in which cognitive representations of number are connected with performance in mathematics.

2.10.1.1 Base-10 structure

An important area of early mathematics which can be investigated via cognitive representations of number is the conception of the base-10 structure of the Arabic number system. This is certainly the conclusion of Thomas et al. (2002), as noted in the previous section. Moeller et al. (2009) suggest that increasingly linear responses to number line estimation tasks could be interpreted as improvement in integrating the single digits’ magnitudes of tens and units in compliance with place value structure. Linearity of responses could thus be a valuable indicator, since longitudinal work has found first graders’ (6-8 year olds) understanding of place-value to be a statistically reliable predictor for specific aspects of arithmetic performance in third grade (9-11 year olds) (Moeller et al., 2011).

Fias and Fischer (2005) also make connections between cognitive representations and base-10 structure in number. The key finding of interest is that the SNARC effect (observed correlation between the spatial orientation of numerical stimuli and response time for comparison tasks) does not clearly extend to two-digit numbers. This raises the prospect that numbers of more than one digit are not processed holistically, as “27”, but rather in some sense as “2” and “7” separately. Fias and Fischer judge that “evidence is accumulating for a separate representation of decade and unit magnitudes” (p. 47). This conclusion supports earlier results also suggesting separate processing of units and decades in two-digit numbers (Nuerk, Kaufmann, Zoppoth and Willmes, 2004).

These findings link clearly to education research results which show children’s difficulty with multiple digit numbers and the additive composition structure underlying them. Nunes and Bryant (2009) point out that children have difficulty with the additive composition of number in general, as well as specifically for the purposes of the base-10 system, and that during the first two years of schooling children appear to be learning about both simultaneously. Thomas (2004) notes that even in later elementary school grades, children show poor understanding of the base-10 structure of the number system. A number of studies have linked understanding to number words in natural language, finding that children speaking languages which more transparently reflect number structure (e.g. Chinese) are quicker than others (for example English and French speaking children) to demonstrate understanding of teen quantities as cardinal tens and ones (e.g. Ho and Fuson, 1998).

2.10.1.2 Mathematical performance generally

An association between small exact number representation (also referred to as subitising) and performance in mathematics has been recorded by both Landerl et al. (2004) and Mulligan et al. (2006). However, there exists "an even larger body of evidence" connecting the analogue magnitude system (AMS) with mathematical performance, in particular, with arithmetical competence in young children (Obersteiner et al., 2013, p. 126).

Many studies have found correlations between approximate number representation and current or subsequent mathematics achievement in children (Aunio and Niemivirta, 2010; De Smedt, Verschaffel and Ghesquiere, 2009; Halberda, Mazzocco and Feigenson, 2008; Siegler and Booth, 2004). As Inglis, Attridge, Batchelor and Gilmore (2011) explain, several researchers have therefore speculated that differences in this form of numerical representation "provide the basis for individual differences in symbolic mathematical competence" (p. 2). The results of Inglis et al., however, present a more complex relationship, with the authors concluding that "the association between non-verbal number acuity and mathematics achievement changes with age, and that non-verbal number representations do not hold the key to explaining the wide variety of mathematical performance levels in adults" (p. 2).

One problem with understanding the relationship between early numerical representation systems and mathematics is that work has been based largely on correlation or regression analysis. This is not of course the only area of mathematical development so afflicted; Obersteiner et al. (2013) acknowledge that the same holds true for other basic number processing tasks, including counting. A rare intervention study used a game in which participating children landed rockets onto a number line, in order to promote approximate number representation on a mental number line (Kucian et al., 2011). The intervention lasted for five weeks, with daily training sessions of fifteen minutes, and participants both with and without development dyscalculia demonstrated improved arithmetical ability. Patterns of brain activation also changed after training, which can be interpreted as "a qualitative change of mental processing" (Obersteiner et al., 2013, p. 126).

Whereas Inglis et al. (2011) and Halberda et al. (2008) used AMS acuity as measured from numerical comparison tasks (participants choose which of two arrays of dots is more numerous), Booth and Siegler (2008) carried out an experimental study using a different methodology. In this study, a computerised number line estimation task was used to assess “knowledge of numerical magnitudes” (p. 1020) and the knowledge was measured by calculating the linearity of children’s estimates, the R^2_{Lin} of each child’s answers to the number line estimation task (p. 1024). Children’s knowledge of addition was also assessed pre- and post-test. The study yielded three important results; firstly, “linearity of children’s number line estimates correlated positively with their existing knowledge of addition”; secondly, “degree of linearity of children’s pretest estimates was predictive of their learning of answers to novel addition problems”; and thirdly “providing accurate visual representations of the magnitudes of addends and sums increased children’s learning of the novel addition problems beyond the level produced by simply presenting problems and answers” (p. 1027). Surprisingly, children’s learning was not enhanced by being prompted to generate their own representations. In summary, the study indicates that “numerical magnitude representations are not only positively related to a variety of types of numerical knowledge but also predictive of success in acquiring new numerical information” (p. 1027)

The first dual intervention involved two groups of children undergoing a three week intervention, with one group experiencing AMS training and the other group SENS training (Rasanen, Salminen, Wilson, Aunio and Dehaene, 2009). The study found that SENS training led to improvements in small number comparison, and AMS training led to improvements in large number comparison, but a weakness of the study was the difference in learning environments used for the two groups, making comparisons difficult.

The first study to compare the effects of exact or approximate training within a “rigorously controlled learning environment” and directly compare their effects on arithmetic performance was carried out by Obersteiner et al. (2013, p. 127). The authors found “no crossover effect” between improvement in exact and approximate number processing, supporting the theory that approximate and exact number processing indeed “rely on distinct cognitive systems” and that training in both is required to enhance the complete range of numerical skills (p. 132).

In terms of improvement in arithmetic performance, “both approximate and exact training led to equal performance in arithmetic” (Obersteiner et al., 2013, p. 133). As Rasanen et al. (2009) found, the effect of both numerical processing training programmes was, though significant, small. This corroborates accounts of the development of the number concept, number understanding and number sense by again revealing that “though basic number processing is certainly an important prerequisite, arithmetical achievement is a complex construct involving other important facets” (Obersteiner et al., 2013, p. 133). The authors point out that the precise relation of basic numerical processing systems to arithmetical ability is still a matter for research, and cite Schneider et al.’s (2009) study as illustration of the complexity, since this study found “use of the internal mental number line was virtually unrelated to mathematical achievement in 5th and 6th grades [ages 10-12]" (Obersteiner et al., 2013, p. 133). This study in fact demonstrates yet another reason for caution about "interpreting performance on single tasks as a measure of the related mental representation" (p. 133); in contrast to the summary given by Obersteiner et al., Schneider et al. (2009) specify that they assess only two aspects of the approximate number system: the distance effect and SNARC effect. They state that: "Individual differences in the use of the internal number line, **as assessed by these 2 effects**, seem to be of little importance when it comes to the acquisition of the content of 5th- and 6th-grade mathematics lessons" (p. 359, emphasis added). They conclude, as do Obersteiner et al. and the majority of researchers in this area, that the relation between the AMS and mathematical performance is not yet well understood. Specifically, there appears to be "no simplistic relationship between the ANS and symbolic mathematics achievement" (Inglis et al., 2011, p. 13).

2.11 Literature review postscript

Relevant empirical studies of children's number line estimations have continued to be published during and since data collection for this thesis. Whilst these studies clearly could not influence the shaping of the research questions, methodology or data collection of this thesis, they have informed the focus of the quantitative analysis and the interpretation of findings. This section of the literature review addresses the most recent research on number line estimation, and explains the impact for this thesis.

2.11.1 Recent number line estimation studies

Two recent papers, Rouder and Geary (2014) and Slusser, Santiago and Barth (2013), report number line research particularly relevant to this thesis. Both studies examined young children's number-to-position number line estimations, and sought to shed light on developmental changes in number line estimation, and the cognitive processes underlying number line estimations. The two studies share a number of important elements with the research design of this thesis: they investigated children in the first grade of school (in the US, children aged 6-7, compared to 5-6 for Year 1 in the UK, but both are the first year of main schooling), imposed no time constraints on children's estimations, and used target numbers in the range 0-100. Unusually for this field, Rouder and Geary (2014) used a longitudinal research design. The key differences between these two studies and this thesis research are, firstly, that they collected no qualitative data on children's estimations; secondly, that both studies gave feedback on the position of the midpoint in children's initial trials; and thirdly, that neither study investigated estimations in the inverse, position-to-number, task. As in this thesis, Slusser et al. (2013) investigated estimations on more than one range.

Using distinct methods of analysis, both Rouder and Geary (2014) and Slusser et al. (2013) conclude that a proportional reasoning account of number line estimation is better supported by the observed data than the logarithmic-to-linear shift account. The precise versions of the proportional reasoning account that each set of authors advocates, and the mathematical models that they argue describe the proportional reasoning, differ slightly.

In addition to adding new data and analysis to the debate on number line estimation, both Rouder and Geary (2014) and Slusser et al. (2013) substantially reinforce criticisms of the logarithmic-to-linear account: its psychological basis, the methods of analysis used in its investigation, and the validity of conclusions drawn from these. Importantly, both papers prompt a re-examination of several studies previously dismissed or overlooked by the field: the most notable of these are Barth and Paladino (2011) and Ashcraft and Moore (2012). Although neither Rouder and Geary (2014) nor Slusser et al. (2013) refer to it, their findings also relate strongly to hypotheses on number line estimation strategies put forward by White and Szucs (2012).

2.11.1.1 "Developmental Change in Numerical Estimation"

The research reported by Slusser et al. (2013) is a cross-sectional study of first graders' [aged 5-6] estimation accuracy in number-to-position estimation in the ranges 0-20 and 0-100, using Percent Absolute Error as the measure of estimation accuracy. The authors compared the log-linear shift account to a proportion-judgment account by fitting the models implied by each account and comparing their success using Akaike information criterion (AIC), a method of comparison that takes account of model complexity as well as goodness of fit (unlike the commonly-used R^2 measure). The models comprising the log-linear shift account were logarithmic and linear functions, and the models comprising the proportion-judgment account were firstly an unbounded power function, and then the one- and two-cycle versions of the proportional power model, as developed by Hollands and Dyre (2000) and first used in a number-line estimation context by Barth and Paladino (2011) (Slusser et al., 2013, p. 198). The study tested children's estimations in different ranges in order to investigate "the claim that different estimation patterns for different ranges within children indicate the presence of multiple types of mental number representations", as suggested by Siegler and Opfer (2003) (Slusser et al., 2013, p. 197).

Slusser and colleagues conclude that their proportion-judgment account describes a more accurate understanding of children's number line estimations than the log-linear shift account. Their data constitute "strong evidence" against the idea that number line estimations directly reflect a mental representation of number, and against the idea of a discontinuous shift from logarithmic to linear mental representation of number (Slusser et al., 2013, p. 205). In contrast to the log-linear shift account, the proportion-judgment account was able to explain observed patterns in children's estimates well. The authors argue that this account also provides evidence for "at least one gradual component" in the developmental changes seen in children's number line estimations (p. 205).

2.11.1.1.1 Explaining developmental change in number line estimation

The models of proportion judgment considered by Barth and Paladino (2011) and Slusser et al. (2013) are originally derived from Stevens' power law, which "expresses the relationship between the estimated magnitude of a stimulus and its actual magnitude" using $y=\alpha x^{\beta}$. In this formulation, β represents "a quantification of bias associated with estimating a particular type of stimulus magnitude (such as brightness, area, or length)" and "α is a scaling parameter" (Slusser et al., 2013, p. 195).

A proportion-judgment account of number line estimation based on these models explains developmental change in children's estimations in two ways. The first aspect is change in the value of β, which reflects the degree of bias in a child's estimations. With age and experience, the parameter approaches the value $\beta=1$, which results in estimations equivalent to $y=x$ (i.e. perfectly linear, with gradient 1). The second aspect of developmental change is the sequential incorporation of anchor or reference points, principally the two line endpoints and an inferred midpoint. As Slusser et al. explain, "Our theoretical account predicts that increased accuracy is linked to the number of reference points utilized by a participant and offers a quantitative explanation of this link." (p. 196) The example given is that of a participant with poor understanding of either the task or the number range in question, who might therefore only take into account the left endpoint of a number line, hence "treating the task as an open-ended magnitude judgment rather than a proportion judgment." (p. 196) In this case, the estimates produced should be well fit by an unbounded power function, the first model of three in Slusser and colleague's proportion judgement account. This explicit consideration of how task responses may be associated with patterns of estimation error is something that previous research had suggested in only vague terms, and makes Slusser et al.'s (2013) work highly relevant to the aims of this thesis. As in the vast majority of studies, the limitation of Slusser et al.'s research is that no data on children's task responses was actually collected, meaning that the connection remains hypothetical.

2.11.1.1.2 Contribution to critical appreciation of earlier work

Slusser et al. argue strongly that deducing internal representations of number from number line estimation tasks is problematic, and argue against some of the specific conclusions made on this basis in other research. Their first argument against deducing internal or mental representations is that multiple task responses can result in estimates well fit by the same function. For example, estimates well fit by an unbounded power function could result from counting-on from the left endpoint with inappropriately sized units, from incorrect referencing of the right endpoint (e.g. not understanding the magnitude of 100), or from ignoring the right endpoint altogether. Slusser et al. emphasise that “For these and other reasons, the applicability of a particular type of function to number-line estimation patterns should not be taken as evidence for a corresponding mental representation of number.” (2013, p. 196) This conclusion forms an argument against the logarithmic-to-linear account of number line estimation, since in this account the logic of the developmental sequence *depends* on number line estimations revealing children’s mental representations. Without assuming this direct connection, the theory is unable to account for the observed pattern of changes in children’s estimations.

Slusser et al. (2013) helpfully situate the logarithmic-to-linear shift account in the history of research into the internal representation of number. The argument that internal representations of numerical magnitude are logarithmically arranged stems from the Weber-Fechner law stating that “the magnitude of sensation is logarithmically related to objective stimulus intensity” (p. 194). In other words, the internal representation of numerical magnitude follows a pattern identified in a much broader area of psychology, the conclusion of Dehaene (e.g. 1997). The principal opposing viewpoint is that internal representations of numerical magnitude are linearly spaced, and that logarithmic patterns appear due to estimation variability which increases in proportion to the magnitude of the target. Notable proponents of this view include Gelman and Gallistel (e.g. 1992).

Number line estimation tasks have been used extensively to try to settle this debate. Siegler and Opfer (2003), for example, failed to observe scalar variability in children’s number line estimations and on this basis argued against a linear internal representation of numerical magnitude. Slusser et al. (2013) query the validity of this argument due to the fact that the vast majority of number line estimation tasks (including those used by Siegler and Opfer (2003)) present number lines with both upper and lower endpoints. Slusser and colleagues argue that this upper bound necessarily affects the variability of responses, and hence “the absence of scalar variability in number-line estimates does not imply a lack of scalar variability in mental representations of numerical magnitude” (2013, p. 195).

This conclusion – that lack of scalar variability in estimates does not imply lack of scalar variability in mental representations – is convincing. However, the extent to which and way in which an upper endpoint affects estimates has to be stated carefully, and to acknowledge the potentially wide gap between an ideal task response and actual task responses. In their discussion of scalar variability in number line estimations, Slusser et al. at one point claim that “typical number-line tasks elicit estimates relative to marked endpoints, prompting participants to make judgments about relative numerical magnitude within a restricted range.” (2013, p. 195) This claim is in fact one of the aspects of number line estimation under investigation. Making judgments about relative magnitude within a restricted range is an ideal way to mathematically solve the task, but there exists a variety of potential task responses – in participants of all ages – that do not involve making a judgment of relative magnitude within a restricted range. Elsewhere in the article, the authors acknowledge this gap clearly, and do correspondingly include models that only take account of a left (lower) endpoint (see Slusser et al., 2013, p. 196).

Considering the question of "how (and whether) to draw conclusions about internal scales of magnitude" from estimation tasks at all, Slusser et al. summarise the difficulties involved (p. 194). The difficulties are firstly, the questionable way in which most studies only consider the two possibilities of logarithmic or linear; secondly, the fact that in modelling number line estimations with logarithmic and linear functions, researchers systematically fail to take into account salient task features; and thirdly that “typical analyses of these tasks attribute variations in number-line estimation solely to numerical processing and numerical representations, assuming that the spatial components of the task do not introduce their own variations.” (p. 195) As Slusser et al. observe, “This assumption is deeply problematic given a wealth of research on the estimation of spatial position in children and adults” (p. 195). Qualitative evidence can illustrate these difficulties with episodes in which intentions (for example a verbalised comment that “I’ll just put it halfway”, or an attempt to appropriately scale counting on a particular number line) do not match spatial actions.

Two previous studies (Barth and Paladino, 2011; Sullivan et al., 2011) had already looked outside the traditional account of number line estimation and interpreted data using psychophysical models of proportion judgment. These studies found success in terms of model fit and in providing a psychologically convincing theoretical underpinning for the models, but the impact of the work was slight. Slusser et al. (2013) identify two key reasons for this lack of impact. First is that the vast majority of research designs focused on the smaller numbers within a given range, since this is the point of maximum discrepancy between linear and log models, and this “yields little data to reveal the details of underestimation patterns for larger numbers” (Slusser et al., 2013, p. 196). Secondly, “contingent on the value of the exponent (β) and on the participant’s use of reference points, unbounded and cyclical power models may closely resemble logarithmic or linear” (p. 196), in other words, the models genuinely appear very similar, and this effect is exacerbated when there are few target numbers chosen in the upper part of a range.

2.11.1.2 *"Children's cognitive representation of the mathematical number line"*

Whereas Slusser et al. (2013) conclude that it is inappropriate to deduce internal representations of number from number line estimations, Rouder and Geary (2014) set out with this precise aim. Although the authors do not give attention to arguing why their deduction of internal representation is valid, their modelling of number line estimation in fact addresses most of the problems identified by Slusser et al. (2013) (see above). Rouder and Geary (2014) do not assume that the internal representation follows either a logarithmic or linear placement, and their modelling of number line estimation is highly focused on task features and the variation introduced by spatial representation.

The central point that Rouder and Geary (2014) emphasise is the issue of estimation variance. Like Slusser et al. (2013), Rouder and Geary (2014) test competing models, but they make a persuasive case that "the distribution is often more diagnostic for adjudicating between theories than the mean alone" (p. 2). In the case of number line estimations, this means not only deciding which model of mean estimation placement to test, but whether or not to assume equal variance for each target within the range covered. Rouder and Geary (2014) argue that the assumption of equal variance is problematic.

Rouder and Geary (2014) test a hierarchy of models, the first of which is a compression model of the form $\log y_i = a + \beta * \log x_i + \epsilon_i$, where ϵ_i is a zero-centred, normally distributed noise term with standard deviation sigma. The authors describe this model as "an amalgam of elements from Dehaene (1997, 2003) and Gallistel and Gelman (1992)" (p. 4). The mean placement of estimations in this model follow a power law, as in Dehaene's theory, but variance of estimations increases with target size, as in Gelman and Gallistel's work. The two following models in the hierarchy tested by Rouder and Geary are models for proportional reasoning responses, the first for participants incorporating the right endpoint (in addition to the always-assumed left endpoint) and the second for participants incorporating both endpoints and also a midpoint. The models consist of S-shaped curves, as in Slusser et al. (2013), the first a one-cycle curve and the second a two-cycle curve. The authors explain that "predictions for the proportional reasoning theory are based on a single principle – variation should be greater the further the number is from the nearest anchor" (p. 4). They note that this principle is "taken as axiomatic, perhaps almost self-evident" (p. 4) and point out that "The minimal requirement is that the variability in this physical distance is a function of the physical distance itself, which strikes us as reasonable." (p. 4) Most convincingly, the authors point out that the assumption is highly plausible independent of internal represenation of number: "The notion that the noise in placement is increasing with the distance to the nearest physical anchor is realistic even if it describes only the translation between precise mental distances and realized physical ones." (p. 4)

The precise predictions about the variance for different target numbers offer Rouder and Geary a "diagnostic" for evaluating models, and a distinct advantage, since "Previous researchers who used regression with its equal variances assumptions are unable to capitalize on the differential predictions about variance …" (2014, p. 6). Rouder and Geary (2014, p. 10) conclude that when analysed with conventional models that seek to explain mean estimations, "our data are consistent with many other developmental studies" including Ashcraft and Moore (2012) and the work of Siegler and colleagues. However, from the analysis using their own models, Rouder and Geary draw both more significant and more specific conclusions. Firstly, they conclude that "The compression model can be unified with the proportional reasoning models, and all of them understood in terms of placements guided by one, two, or three anchor points." (p. 10) This is a significant unification of models, and helps to dismantle a dichotomy between the log-linear shift and proportional reasoning accounts that is not only unhelpful, but perhaps unnecessary. The second conclusion regards developmental change in estimation accuracy. Here, Rouder and Geary conclude that it "results from incorporation of additional anchors, one at a time, that partition the line into segments. Placements are made within these segments, with numerals close to an anchor placed with greater accuracy than those farthest from an anchor." (Rouder and Geary, 2014, pp. 10-11) This conclusion firmly supports that of Ashcraft and Moore (2012) and Slusser et al. (2013), despite other differences between these groups of authors.

2.11.2 Summary and conclusions

Recent research continues to investigate number line estimation on the basis that it is an important element, or at the very least indicator, of early mathematical development. Performance patterns on number line estimation tasks are "correlated with performance on standardized math tests and other measures of mathematical ability" (Slusser et al., 2013, p. 193). The two measures are not only correlated; in fact, "children's ability to accurately place numerals on the line is predictive of their later mathematics achievement, controlling for other factors" (Rouder and Geary, 2014, p. 1). An even stronger claim comes from Fazio, Bailey, Thompson and Siegler (2014): "Precise representations of numerical magnitudes are foundational for learning mathematics. Both correlational and causal evidence link the precision of individual children's numerical magnitude representations to their whole number and fraction arithmetic skill, memory for numbers, and other aspects of mathematical knowledge." (pp. 53-54) In short, the interest of number line estimation tasks is agreed upon.

It is also an agreed, observed fact that children's number line estimations change with age and experience. The underlying characteristics and mechanisms are the aspects of developmental change under dispute. Many previous number line estimation studies provide evidence of a numerical spatial association, but even now "the exact nature of the connection remains controversial" (Siegler and Thompson, 2014, p. 40). In their review of spatial-numerical association research, Patro, Nuerk, Cress and Haman (2014) similarly conclude that "The mental processes and strategies underlying the number-space associations in this task [number line estimation] are thus still controversial." (p. 4)

The major debates in the field of number line estimation and numerical magnitude representation remain unanswered. These debates include:

- Whether numerical magnitude is internally represented with a compressed-scale function (e.g. logarithmically) with uniform variability, or linearly with scalar variability (increasing with increasing numerical magnitude)
- The relationship between internal representations of numerical magnitude and number line estimation tasks
- Whether number line estimation tasks are more appropriately interpreted as proportion judgment tasks, and hence modelled using cyclical power functions
- What causes the apparent increase in linearity of number line estimations

In terms of the efforts of research to answer the above questions, recent studies have acknowledged the need to focus more on the estimates of individuals (as opposed to the median estimates of cohorts). However, there remains a lack of longitudinal research, and an absence of qualitative evidence to support hypotheses about children's task behaviour; the most recent known study to collect qualitative evidence on children's number line estimations remains Petitto (1990), as discussed in the main literature review. The absence of qualitative evidence is particularly pressing given the increase in hypotheses connecting changes in children's number line estimations with progressive use of specific mathematical reference points (Ashcraft and Moore, 2012; Rouder and Geary, 2014; Slusser et al., 2013; White and Szucs, 2012). Plausible and specific hypotheses have been proposed, but recent research has not yet collected evidence on children's actual estimation processes.

With regard to developmental changes in number line estimation, although researchers have devoted a great deal of energy to debating the precise form of the mental representation underlying estimations, or alternatively the psychological model of proportion judgment most appropriate, it is important to remain focused on the documented changes in children's estimation. Fazio et al. (2014, p. 54) summarise the repeatedly-observed change as follows: “As children gain experience with increasing ranges of numbers, their number line estimates become more accurate and more closely approximate a linear function.” The tone of Fazio et al. differs remarkably from earlier work, in which Siegler and colleagues fairly vociferously argued for the logarithmic-to-linear account of children’s number line estimations. The most recent claims about number line estimation are more moderate: firstly that with age and experience, children’s estimates more closely *resemble* a linear function, and secondly that linear accuracy of number line estimations for symbolically represented numbers is closely related to mathematics attainment (Fazio et al., 2014, p. 54). One logical outcome of this simplified attitude is to compare children’s estimations over time simply in terms of their linearity, rather than by fitting and comparing multiple competing models for each child, condition and cohort.

This chapter has reviewed the literature in order to situate the current research problem and questions amongst existing research. The review of theoretical material demonstrated the theoretically complex approaches that have accompanied the various strands of relevant research. The latter sections, focusing on empirical studies, have shown that relevant empirical findings have been reached in a variety of research areas. In the next chapter, the theoretical framework which supports the thesis is explained. The theoretical framework provides a vital scaffold to the research problem, and enables a theoretically coherent interpretation of the literature reviewed.

Chapter 3 Theoretical Framework

3.1 Overview

In this chapter I explain the theoretical framework supporting this thesis. This framework supports every stage of the research; it articulates and clarifies the particular perspective that led to the identification of the research problem and the formation of the research questions; it guides the literature review; it informs and supports the chosen research methodology; and it directs the purpose of the analysis and interpretation of results. The theoretical framework has already been presented at the Young European Researchers in Mathematics Education summer school (YESS-6) held in Faro, Portugal in 2012.

The key theory of the framework is Duval's theory of cognitive representation, vision and visualisation (Duval, 1999, 2002). This theory sits within a cognitive tradition and aims to provide a framework for analysing the cognitive functioning of mathematical thinking and conditions of mathematical learning. The two fundamental aspects of Duval's theoretical perspective are the ideas that representation and visualisation are central to mathematical understanding, and that the particular character of mathematical knowledge necessitates a more detailed analysis of representation and visualisation than domain-general theories have provided. Studies of vision, visualisation and representation which do not pay sufficient attention to the uniqueness of mathematical knowledge are "deceitful" and unable to illuminate the processes of mathematical learning, and students' difficulties with that learning (Duval, 2002, p. 311).

In addition to Duval's theory of cognitive representation, the theoretical framework of this thesis incorporates aspects of Presmeg's theory of imagistic representation, as adapted for the context of the cognitive representation of natural numbers by Thomas, Mulligan and Goldin (2002).

The important remaining theoretical aspect of this thesis is Resnick's (1983) account of number development. The crucial role played by this theoretical work is not in the overall theoretical framework supporting the thesis, but in the analysis stage. Resnick's (1983) work is therefore not included in this chapter, but addressed separately in the theory for analysis section (4.4.2) of Chapter 4.

This chapter begins with a detailed discussion of Duval's framework of cognitive representations and the distinctions and classifications that can be made. It then moves onto discussion of the ideas of vision and visualisation, with respect to representations, and the consequences for mathematical learning. Attention is given to how this particular theoretical approach is appropriate to enable and support the current research. The chapter then goes on to introduce and explain the need for Thomas et al.'s adaptation of Presmeg's theory of imagistic representations.

3.2 Duval's theory of cognitive representations and visualisation

Duval's theory of cognitive representation and visualisation was laid out in a comprehensive form in a plenary address at the 21st Annual PME-NA meeting in Mexico (Duval, 1999) and subsequently republished by Fernando Hitt in a volume representing the progress of the PME-NA 'Working Group on Representation and Mathematics Visualization' over the period 1998-2002 (Duval, 2002; Hitt, 2002). The theoretical ideas from the 1999 paper were further developed in a theoretical analysis of cognitive difficulties in mathematics learning published in *Educational Studies in Mathematics* (Duval, 2006).

The 2002 volume published by Hitt contains strong theoretical analyses of representation and visualisation in mathematics education from researchers including Norma Presmeg, James Kaput, Patrick W. Thompson, Luis Radford and Sylvette Maury in addition to Duval and Hitt. Presmeg, in her preface to the book, describes it as "an uneven collection of papers, as individual as their individual authors" but singles out Duval's contribution for expressing well the "collective apprehension" reflected by the various papers taken as a collection (Presmeg, 2002, p. ix). Presmeg frames the introduction of the whole collection in terms of Duval's contribution, and describes Duval's chapter itself as "amazingly dense and authoritative", in fact "quite startling". In contrast to the précis and brief evaluation she offers for the other chapters, with respect to Duval's chapter, Presmeg (2002, p. xiv) states "I cannot do justice to Duval's work here: in fact I cannot even introduce it adequately."

3.2.1 Starting point: the unique character of mathematical knowledge

Duval's starting point is that mathematical knowledge differs in fundamental ways from knowledge in other fields, and until this is taken into account, deep understanding of mathematical thinking and learning will not be forthcoming.

Duval characterises the nature of mathematical knowledge as "paradoxical" (Duval, 1999, p. 4). It differs most obviously from other fields of knowledge in respect of the fact that mathematical objects, unlike the objects of study in other domains, are not available to perceptual senses and can only ever be accessed via the production of semiotic representations. At the same time, mathematical understanding requires absolutely "not confusing the mathematical objects with the used representations" (p. 4). Because of this paradoxical character of mathematical knowledge, Duval holds that "Representation and visualization are at the core of understanding in mathematics" (p. 3).

Duval observes that "explanations of the deep processes of understanding and learning in mathematics" have progressed less in recent years than innovations in curriculum and teaching. For Duval, this is directly linked to the unique character of mathematical knowledge and the consequent centrality of representation and visualisation. The concepts of mathematics are not what sets it apart – after all, "there is no domain of knowledge that does not develop a set of more or less complex concepts" (Duval, 2006, p. 106). The difference arises from the "very specific epistemological situation of mathematics" which means students have to face a "radically" different cognitive use of signs than that of other disciplines (p. 107). Deep understanding of mathematical learning therefore necessitates going beyond "local studies of concept acquiring at each level of the curriculum", "mere reference to very general theories of learning" and "global description of student's activity in classroom" (Duval, 1999, p. 3). The research here alluded to by Duval has not been, is not, and can never be sufficient to further illuminate the deep processes of mathematics learning.

3.2.2 A framework of cognitive representations

3.2.2.1 The internal/external distinction

Duval's framework encompasses all *cognitive representations*, that is to say, representations involved in the act of cognition. As discussed briefly in the Introduction, the framework rejects a primary distinction between internal and external representations. This sets Duval's theoretical perspective apart from many accounts of representation (e.g. Goldin, 2002) but is a key point in common with Presmeg's theoretical work on visualisation. Presmeg's decision not to develop the internal/external distinction in her theoretical work follows from a key assumption of Piaget and Inhelder (1971), that whenever a person creates an external spatial arrangement, they are guided by an internal image (Presmeg, 2006).

Emphasis on a distinction between internal and external representation both arises from and leads to confusion. From a cognitive perspective, the opposition or distinction rests upon "confusion between the phenomenological mode of production and the kind of system mobilized for producing" (Duval, 2006, p. 105). The internal/external distinction is commonly treated "as though it was evident and primary", but in Duval's analysis the division is in fact "a misleading division … which brings about two very damaging confusions" (Duval, 1999, p. 5).

The first confusion the division produces is to focus undue attention on an aspect of representations that is *not* key to understanding their significance. When characterising cognitive representations, "the distinction mental/ external refers to their mode of production and not to their nature or to their form" (Duval, 1999, p. 5). Duval emphasises that the signs themselves "are neither mental nor physical or external entities" (p. 5). In particular, the internal/external distinction can lead towards mistakenly seeking or assuming a correspondence "between the distinction mental/material and the distinction signified/significant". This correspondence is erroneous, following a Saussurian explanation, because the significance of a sign is not determined by its "material realization" – it is determined only by its relation to other signs (p. 5).

3.2.2.2 The intentional/automatic distinction

The internal/external distinction is also misleading because it obscures a distinction which *is* significant. This is the distinction between intentional/semiotic representations and physical/organic representations (Duval, 1999, p. 5). Semiotic representations are here defined as cognitive representations "intentionally produced by using any semiotic system". In contrast, physical/organic representations are those "causally and automatically produced either by an organic system (dream or memory visual images) or by a physical device (reflections [physical], photographs)" (Duval, 1999, p. 5). The diagram below shows Duval's classification of cognitive representations. It is an updated version of the same classification diagram found in the 1999 paper (Duval, 1999); the structure and substance of the classification does not change, but Duval's labelling and examples mean that the newer diagram provides greater clarity on a number of points (see Appendix 2 for 1999 diagram and brief discussion of the differences). The diagram classifies cognitive representations as follows:

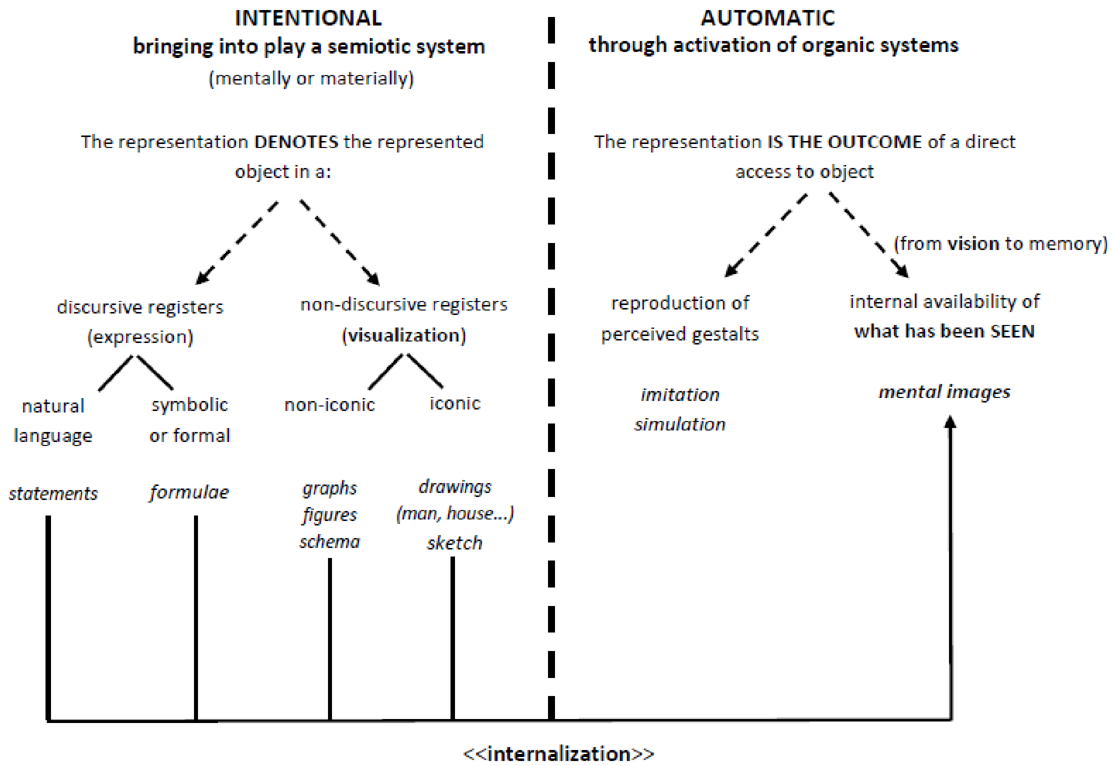


Figure 1. Diagram classifying cognitive representations (Duval, 2000, p. 66)

Three aspects of the classification of cognitive representations should be emphasised. The first aspect is that semiotic representations can be produced internally or externally. Duval's framework uses *cognitive* to describe representations involved in cognition; this, rather than their 'location', is the pertinent characteristic. *Cognitive* does not have the privileged association with *mental* or *internal* that it carries in many other situations.

The second particularly relevant aspect of the intentional/automatic distinction is that *mental images* are found within both categories. Explaining the existence of these mental images with reference to the diagram, Duval (1999, p. 6) writes: "We can notice the existence of two heterogeneous kinds of "mental images": the "quasi-percepts" which are an extension of perception (on the right) and the internalized semiotic visualizations (on the left)."

The third important aspect to note is that the intentional or semiotic representations (the left of the diagram) are further divided into discursive and non-discursive representations. The class of non-discursive semiotic representations consists of visualisations (showing "relations, or better, organization of relations between representational units" (Duval, 1999, p. 13) – see 3.2.3.2), which stand in analogical relation to the objects they represent. The class of discursive representations, on the other hand, are non-analogical in nature (Duval, 1999, p. 6).

3.2.2.3 Note: conscious representations

In the 1999 version of Duval's classifying diagram (see Appendix 2), the caption states that the diagram classifies "conscious representations" (Duval, 1999, p. 6). Duval's meaning is here unclear, since the term *conscious*, let alone *conscious representation*, is nowhere defined in the key English-language works presenting and discussing Duval's theoretical perspective (Duval, 1999, 2000, 2002, 2006). Since *conscious representation* is nowhere else mentioned in these works, it seems difficult to conclude that Duval considers the *conscious* specification a key point. Furthermore, the 1999 diagram caption states that the depicted classification "can be expanded more and includes all kinds of representations" (Duval, 1999, p. 6) – including, presumably, *non*-conscious representations, whatever these may be, making it valid to take the classification as applying to all cognitive representations. This interpretation is supported by the updated diagram (Duval, 2000, p. 66), in which Duval refers simply to *cognitive representations* as in the rest of his theoretical writing, without any reference to *conscious*.

The above argument notwithstanding, the word *conscious* is a highly significant term within theory of mind research, and it is interesting to ask what Duval intends by it in the 1999 classification diagram. It seems clear that *conscious* is not to be understood to mean *intentional*, since the classification diagram (Duval, 1999, p. 6) includes both *intentional* and *automatic* as contrasting subcategories within the overall scheme of "conscious representations". From Duval's (1999) comments on the multiple meanings of the term *representation* a tentative conclusion is that, by specifying *conscious* representations, he seeks only to exclude representations in the sense of 'information-coding in the brain', as studied in cognitive neuroscience. These representations are notably absent from consideration in all Duval's classifications of cognitive representations, and it is a plausible understanding of *conscious* that it should exclude this kind of representation.

As noted, no further mention of *conscious representation* is found in Duval's English-language theoretical works, but the term *conscious* is discussed at some length in the earlier *Sémiosis et pensée humaine* (Duval, 1995). This book neither presents nor uses the theoretical framework adopted by this thesis (as presented in Duval's works from 1999 onwards) but instead works with a more traditional theoretical framework based on the two classical dichotomies of internal/external and conscious/non-conscious (Duval, 1995).

Duval explains that the opposition conscious/non-conscious is the opposition between that which appears to an individual and that (s)he notices, on one hand, and that which completely escapes an individual and that (s)he cannot notice on the other hand ("L'opposition conscient/non-conscient est l'opposition entre ce qui apparaît à un sujet et qu'il remarque d'une part, et, ce qui lui échappe complètement et qu'il ne peut pas remarquer d'autre part" (Duval, 1995, p. 24)). An individual becoming conscious of something 'sees' the something, which then takes the status of an object for that individual. This clarification of *conscious* is compatible with the usage in Duval (1999); all cognitive representations incorporated into the classification diagram are capable of being noticed by the individual involved, and capable of objectifying for the individual that which they represent.

3.2.2.4 Registers of representations

Duval's concept of *registers* of representation is introduced in terms of the different semiotic systems developed and adopted throughout history. Duval writes that "Each new semiotic system provided specific means of representation and processing for mathematical thinking. For that reason, we have called them 'registers of representation'" (Duval, 1999, p. 6). It should be emphasised that a register of representation includes both the means of representation **and** the specific ways of processing (defined by Duval (1999) as the transformation of a representation within a register) associated with those means. Using this definition, the conclusion is that within the class of semiotic representations, "we have several registers for discursive representation and several systems for visualization" (p. 6). Examples of registers of representation are natural language, 2D shape representations and symbolic notation.

Thompson expresses uncertainty about how Duval's registers of representations do or do not differ from existing constructs. Thompson notes, for example, a strong similarity with Post, Behr, Lesh and Harel's "modes of representation" and Kaput's "representation systems" (Thompson, 1999, p. 2). Thompson also asks whether a register of representation is to be considered an ad hoc construct, "suggested to us by observing" (pp. 2-3) or whether it is "defined operationally by specifying cognitive operations that cohere into schemes" (p. 3). In Thompson's view, a register of representations must either be a "scheme of operations" or solely "determined by social convention" (Thompson, 1999, p. 3).

Duval expands on the notion of registers in the article *A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics* (Duval, 2006), and provides a more detailed analysis of the types that exist. It is argued that the common classification of representation systems, based upon contrasting language with image, is an insufficient characterisation. Instead, Duval proceeds on the basis of the cognitive *functions* that a system of cognitive representation is able to perform. Such functions include mathematical processing, communication, information processing, awareness and imagination. Some representation systems are able to perform only one of these (monofunctional) whilst others are able to perform many (multifunctional).

The following simplified diagram (Duval, 2006, p. 110) demonstrates how the above classification of semiotic representations into monofunctional and multifunctional, combined with the existing classification of semiotic representation systems into analogical and non-analogical (in the 2006 paper, labelled *discursive* and *non-discursive*), leads to the identification of four distinct types of register:

	Discursive or non-analogical representation (denotations, statement or inferences)	**Non-discursive** or analogical representation
Multifunctional registers	Natural language: oral and written (visual)	Iconic: drawing, sketch, pattern Non-iconic: geometrical figures
Monofunctional registers	Symbolic systems	Combination of shapes; diagrams; graphs

Figure 2 The four types of representation register (Duval, 2006, p. 110)

An additional category of "transitional auxiliary representations" encompasses those intentional representations which do not belong to a particular semiotic system, and have no particular rules for formation or transformation. The example given by Duval is that of matchsticks standing for small integers. This kind of representation is a *sign* to the user, but not part of a semiotic system (Duval, 2006).

3.2.3 Representations, vision and visualisation

Having introduced Duval's framework and classification of cognitive representations, the chapter now moves on to explore Duval's theoretical explanation of the processes by which representations are apprehended. The key concepts discussed are *vision* and *visualisation*, both theoretically defined in relation to representations and the idea of 'seeing'.

3.2.3.1 Vision

Vision, according to Duval, refers to visual perception and by extension to visual imagery. Like perception, vision performs two cognitive functions: the *epistemological* function and the *synoptic* function (Duval, 1999). The epistemological function of vision is to give "direct access" to any physical object. Vision is taken to give uniquely direct access to objects, and for this reason has historically been taken as the epistemological model for the idea of intuition – "Nothing is more convincing than what is seen" (p. 12).

The synoptic function of vision differs significantly from the epistemological function. The synoptic function apprehends multiple objects simultaneously, seeming to immediately provide "a complete apprehension of any object or situation" (Duval, 1999, p. 12). For this reason, Duval characterises the synoptic function of vision as "the opposite of discourse, of deduction, which requires a sequence of focusing acts on a string of statements" (p. 12). In actual fact, practical considerations mean that the synoptic function of visual perception is carried out very imperfectly (p. 12). The limitations are firstly that humans are only able to see one side at a time of objects in the three-dimensional world, reaching at best a "juxtaposition of successive sights", and secondly that even within one view, visual perception always focuses on one small part at a time and must "jump from one part to another". This is a necessary, not optional, facet of vision: "There is no visual perception without such an exploration" (Duval, 1999, p. 12).

Vision can perform both the above functions for physical objects. Duval raises the question for mathematical objects: "are there cognitive structures that can perform both the epistemological and the synoptic function for the mathematical knowledge?" (Duval, 1999, pp. 12-13). The particular nature of mathematical objects means that the answer is "no". However, in exploring what kind of 'seeing' of mathematical objects *is* possible, a clear characterisation of visualisation is achieved.

3.2.3.2 Visualisation

Whereas visual perception could represent an object in its 3D environment, or a 2D view of it, visualisation is based on the production of a *semiotic* representation which "shows relations, or better, organization of relations between representational units" (Duval, 1999, p. 13). Examples of such units are 2D shapes, coordinates, and words. As a cognitive activity, visualisation is "intrinsically semiotic, that is, neither mental nor physical" (p. 13). As the cognitive representation classification diagram (Figure 1) indicates, visualisations make up the class of analogical semiotic representations. Visualisation is contrasted repeatedly with discursive, non-analogical semiotic representation, particularly in geometry, where Duval argues geometrical figures always require the coordination of discursive and visualisation registers (Duval, 2006).

Visualisation is not primitive, and not mere visual perception (Duval, 1999). The paradoxical nature of mathematical knowledge means that "mental representation as direct grasping of mathematical objects" (p. 15) can never happen and visualisation is thus unable to perform the epistemological function (as vision does for physical objects). In summary, visualisation "performs only the synoptic function, is not intuition but representation" (pp. 14-15).

Duval emphasises that visualisation, by this definition, is necessary for mathematical understanding, since it is only through visualisation that mathematical structure can be apprehended. In a "string of discrete units" such as words, symbols, or propositions, only certain formations or organizations can be displayed, and since "understanding involves grasping their whole structure, there is no understanding without visualization" (Duval, 1999, p. 13). It is this which makes it so vital not to reduce visualisation to a kind of vision; vision sees the layout of the units as they are for example on the page, and visualisation 'sees' the structure represented by the units. Visualisation should not and cannot truly be reduced to vision, because "visualization makes visible all that is not accessible to vision" (p. 13).

Whereas vision requires physical movement to gain "a complete apprehension of the object", visualisation has the potential to "get at once a complete apprehension of any organization of relations" (Duval, 1999, p. 13). In practice, however, this may not happen, since visualisation requires a great deal of register-specific training.

3.2.3.3 Seeing

The 'seeing' of vision is not the same as the 'seeing' that occurs in visualisation. There are two potential points for confusion, however. Firstly, graphical productions of semiotic representations are in fact apprehended through visual perception, and so visualisation is consequently "always displayed within visual perception or within its mental extension" (Duval, 1999, p. 14).

In the second place, confusion can arise because of iconic representations, such as some drawings. These representations are semiotic in nature, but because of the "relating likeness" between representation content and represented object, it is often possible to recognise the represented object (e.g. tree) "at once, without further information" (Duval, 1999, p. 14). This possibility arises because iconic representations refer to "a previous perception of the represented object" and from that to the concrete object (p. 14). In mathematical visualisation however, this mechanism of seeing cannot happen – the nature of mathematical objects means that "to look at them [visualizations] is not enough to see, that is, to notice and understand what is really represented" (p. 15). It is not easy to deal with these confusions. Duval notes that when looking at visualisation there is a strong tension and discrepancy between the ways of seeing, "between the common way to see the figures, generally in an iconic way, and the mathematical way they are expected to be looked at" (Duval, 2006, p. 116). This is an example of how multifunctional registers can – deceptively – appear "common and directly accessible to every student", when in fact the mathematical way of using the multifunctional register "runs against the common practice" (Duval, 2006, p. 116).

The meanings of visualisations (which are semiotic representations) lie in the organization or structure that they represent. In technical terms, Duval characterises visualisation as the "bi-dimensional organization of relations between some kinds of units" (Duval, 1999, p. 15). Seeing what is represented by visualisation is not as simple as in the case of iconic representations. Duval emphasises that the significance of visualisations is in the organization of units:

> *The intricacy of mathematical visualization does not consist in its visual units – they are fewer and more homogeneous than for the images – but in the implicit selection of which visual contrast values within the configurations of units are relevant and which are not.*
>
> (Duval, 1999, p. 17)

Seeing the meaning of a visualisation involves therefore both the seeing of vision and the 'seeing' of mathematical structure that others might describe as 'understanding'.

3.2.3.4 Mental image

Duval describes the expressions "mental image, "mental representation" and "mental imagery" as "equivocal" terms (Duval, 1999, p. 13). They are commonly used to mean only an extension of visual perception. Duval quotes an insightful description of the scope of visual images in this sense, as an extension of visual perception, by Neisser (1967):

> *"[V]isual image" is a partly undefined term for something seen somewhat in the way real objects are seen when little or nothing in the immediate or very sensory input appears to*

justify it. Imagery ranges from the extremely vivid and externally localized images of the eidetiker to the relatively hazy and unlocalized images of visual memory.

(Neisser, 1967, p. 146, cited by Duval, 1999, p. 14)

The phrase “mental imagery” can also however refer to the mental form of visualisation, as indicated by Duval’s classification diagram (Figure 1). This is a separate meaning of the term, and here mental imagery is “the mental production of semiotic representations as in mental calculation” (Duval, 1999, p. 14). Most interestingly for the current research, both quasi-perceptual mental images and the visualisations can be operated upon: “Actions like the physical ones (rotation, displacement, separation) can still be performed on some quasi-percepts” (Duval, 1999, p. 6).

3.2.4 Contributions to understanding mathematical cognition

According to Presmeg, Duval offers a “cogent argument… for the distinction between vision or perception and visualisation in mathematical cognition and learning, and for the power of coordination of different registers of representation in these activities” (Presmeg, 2002, p. ix). Duval’s 1999 paper was especially important for English-speaking researchers because Duval’s extensive research had been previously published largely in French. Presmeg notes that Duval’s framework has been used extensively by other researchers, such as Acuña (Presmeg, 2006). Specific implications for the understanding of mathematical learning are offered by Duval’s theoretical work. The two most significant of these are outlined below.

3.2.4.1 Vitality of focusing on representations and not concepts

The first point contributed by Duval’s theoretical work is the need to focus on representations and not subordinate these to concepts. By emphasising the paradoxical nature of mathematical knowledge, and illuminating the complex cognitive processes involved in mathematically understanding representations that follow from this nature, Duval makes a powerful case for researching mathematics education from a representation viewpoint. In Duval’s (1999, p. 8) words, the work shows that

… semiotic representations constitute an irreducible aspect of mathematical knowledge and that wanting to subordinate them to concepts leads to false issues in learning. That amounts to forget the paradox of mathematical knowledge: mathematics objects, even the more elementary objects in arithmetic and geometry, are not directly accessible like the physical objects. Each semiotic register of representation has a specific way of working, of which students must become aware.

Each semiotic register not only affects the way of working with representations, but affects what aspects of the represented object are made explicit. Duval, following Frege, stresses the difference between the content of a representation and that to which the representation refers. The relation between the two is simply that the represented object is *denoted by* the representation; there need not necessarily exist any 'intuitive' relation at all (Duval, 2006, p. 114).

For the case of semiotic representations, Duval goes as far as to state that "the content of a representation depends more on the register of the representation than on the object represented" (Duval, 1999, pp. 40-46; 2006, p. 114). This differs from the case of automatic (non-semiotic) representations, those "produced by physical devices (mirror, camera, microscope, etc.) or by sensory and brain organizations", where there exists "something like a causality relation" (Duval, 2006, p. 114). This implies a further – third – consideration of the importance of registers of representation. The register not only affects the way of working, and the content, but also the relation between the content and the represented object (p. 115).

This emphasis of Duval's work influenced the research of this thesis by providing further theoretical justification for the representation viewpoint on mathematical learning. Specifically, it articulates the good reasons to research representations themselves, not just research learning 'through' representations.

3.2.4.2 Processing and conversion

The second key contribution of Duval's theory is to illuminate in greater depth the mathematical actions that are carried out with representations. Duval argues that mathematical processes consist of two kinds of transformation of representation: processing, and conversion. The term *processing* is used for transformations that stay within the same register of representation, for example algebraic computation that transforms between symbolic algebraic expressions (Duval, 1999, p. 8). In contrast, *conversion* is used to describe transformations between different registers of representation, for example the transformation of equations into Cartesian graphs (p. 9).

Linking the conversion of representations directly back to the paradoxical nature of mathematical knowledge, Duval states that "only students who can perform register changes do not confuse a mathematical object with its representation" (Duval, 1999, p. 10). In this light, the conversion of representations is the crux of mathematical understanding, and it is neither easily nor frequently achieved by students (Duval, 1999, 2006). Research that focuses on finding the "right" or "most accessible" register will advance nothing more than "surface understanding", since understanding proceeds from conversion. This is true for all representations; Duval holds that "Even auxiliary and individual representations, the most iconic or concrete ones, need to be articulated with the semiotic representations" (Duval, 2006, p. 128).

In the context of this research, Duval's theory offers theoretical motivation for the further investigation of number line estimation tasks. From the perspective of Duval's framework, such tasks are a *conversion* between representations of number in different registers, between a spatial representation on the one hand, and a symbolic and verbal representation on the other hand.

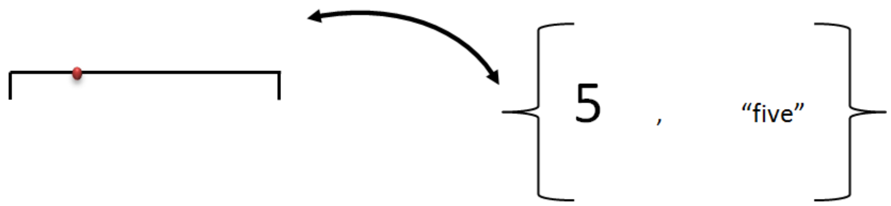


Figure 3. Conversion in a number line estimation task

Conversion of representations is often mistakenly considered as translation or encoding (Duval, 2006). Conversions can in fact be congruent or non-congruent. Congruent conversions, which seem significantly easier to carry out, are those in which "the representation of the starting register is transparent to the representation of the target register" and here conversion does feel like a simple translation from unit to unit (Duval, 1999, p. 10). A detailed analysis of the idea of congruence of conversion shows that it consists of three aspects (Duval, 2006, p. 122):

1. Whether a 1-1 mapping between meaningful component parts (symbols, words, or visual features) is possible or not
2. Whether there is a univocal choice for each meaningful component of the target representation
3. Whether the organizational structure of map-able components in the source representation is maintained in the target representation or not

The difficulty of conversion between representations is additionally affected by the direction of conversion. It frequently occurs that conversion can be "obvious" in one direction, "while in the inverted task, most students systematically fail" (Duval, 2006). The complexity of a conversion between representations is therefore a function of two phenomena: the congruence and the non-reversibility of the conversion. These phenomena are determined by the registers involved – the source register and the target register.

3.3 Cognitive integration

Duval's theory of representation and visualisation sits clearly within a cognitive tradition; it focuses upon and aims to illuminate the cognitive processes of mathematical thinking and learning (Duval, 2002). In relation to philosophical perspectives on how cognitive processing is related to ‘actions performed in the world’, the perspective of Duval’s theory fits well with the perspective identified by De Cruz as *cognitive integration*. This perspective rejects the accounts of both internalism and active externalism, and proposes an alternative (De Cruz, 2012).

The perspective of internalism is that cognitive processes take place “only in the skull” (De Cruz, 2012, p. 137). Whilst actions performed in the external world “can enhance and improve cognitive performance”, they are not themselves included in the act called cognition. De Cruz illustrates with the example of calculation on an abacus: from an internalist perspective “the only action that is cognitive is the retrieval from memory of abacus techniques and interpreting the result by converting the observed configuration of beads into an internal representation of mental magnitude” (De Cruz, 2012, p. 137).

The perspective of active externalism argues that in fact external tools are often part of the act of cognition itself. There is an obvious problem in that “if one simply grants cognitive status to every object that is somehow causally involved in cognitive processes, we end up with sentient pencils and notepads” (De Cruz, 2012, p. 137). A solution proposed by Clark and Chambers (1998) is the so-called parity principle, according to which if a certain brain process is characterised as cognitive, then a structurally similar process that takes place outside the brain ought to also be characterised as cognitive.

Cognitive integration similarly argues that external actions can be part of cognitive activity. However, it proposes that these external actions need not structurally resemble internal cognitive processes. Duval’s theoretical perspective on representation fits well into this cognitive integration view. Cognitive activity is certainly not confined to the ‘skull’: Duval’s framework of cognitive representations explicitly states that acts of cognitive representation may produce internal or external representations, and that transformations (processing and conversion) occur between and within both internal and external registers of representations. Duval’s framework does not, however, support an active externalist viewpoint, since the parity principle does not hold for all external activities qualifying as cognitive within Duval’s framework. In fact, Duval explicitly notes that the structure of cognitive activity with representations may *differ*, for example “mental arithmetic uses the same decimal system like written calculation but not the same strategies because of the cognitive cost” (Duval, 1999, p. 5).

De Cruz offers a critical comparison of cognitive integration, internalism and active externalism in the specific context of spatial representations involved in numerical processing, and concludes that cognitive integration is the only justified philosophical conclusion. Empirical findings, such as the development of a linear mental representation of number in Western children, and the effect of playing linear numerical board games (Ramani and Siegler, 2008), strongly support the theory that the internal and external aspects of numerical cognition are "complementary" and relate with dynamic interaction (De Cruz, 2012).

3.4 Duval's theory and the research problem

Duval's theory of cognitive representations shapes the research of this thesis at all levels, from the identification of the research problem to data analysis. Whilst the contribution of Duval's theory to the present research has been noted in various places, in this section I summarise the key ways in which Duval's theoretical perspective shapes the research.

In the first place, at a broad level, Duval's framework provides further theoretical justification for studying mathematics education from a representation viewpoint. This not only supports research into children's cognitive representations of number, but supports researching these representations as objects of interest in their own right, rather than as secondary to concepts. Duval's theory of representations stands apart from others in this sense, holding representations themselves to be the crux of mathematical learning.

A second feature of Duval's theoretical perspective that sets it apart from others in terms of its appropriateness for the current research is the way in which Duval's framework encompasses both intentional/semiotic and automatic representations. As is clear from the literature review, the status and nature of responses to number line estimation tasks is not clear. Whether they are quasi-perceptual responses (from evolutionarily developed systems), intentional visualisations, automatic learned responses or some combination of these categories is not yet certain. For this reason, it is absolutely necessary for this research to work within a framework of representation that encompasses both intentional and automatic representations, and acknowledges that relations exist between them (even though they may be unknown).

Beyond supporting of research into number line estimations, Duval's broad and comprehensive framework of cognitive representations motivates research questions into the different cognitive representations produced by children. Duval's theoretical work indicates that heterogeneous kinds of representation of mathematical ideas are to be expected, leading this thesis towards serious consideration of representations in multiple modes and registers. The theoretical framework indicates that 1-1 correspondence between internal and external representations, or any two different registers, should not be expected. It is for this reason that this thesis considers the external cognitive representations produced by children in their own right, and does not seek to infer internal representations underlying them. This aspect of the framework also motivates the research questions asking which modes of representation children use in particular tasks, and what aspects of number are made explicit in each case.

At the same time, the framework emphasises that the representations of number produced are all still cognitive representations, and that investigating the connection between them is a worthwhile task. The heterogeneous types of mental image identified by Duval's work may both be involved in children's representation of numbers, and it is a valid question to ask what aspects of number are cognitively represented in each kind, and how they compare.

Duval's work helps to emphasise that conversion between representations happens frequently in mathematical talk and tasks in the classroom, and at numerous points in the task-based interviews used in this thesis, not only in number line estimation tasks. In addition to helping to identify conversions of representations, Duval's work theoretically explains the mathematical significance of these processes. As explained, the process of conversion of representations is held to be the crux of mathematical understanding, and hence children's conversion of number representations is of high educational interest. Duval's theoretical perspective, together with the literature reviewed, makes examination of the *process* rather than just result of estimation and conversion a priority, and the research questions reflect this interest.

Finally, Duval's analysis of conversions contributes an additional motivation for the longitudinal aspect of the current research. It is to be expected that children's knowledge of number representations deepens and changes during the year of school studied, and their experience of representations of number increases. For this reason, children's conversions between representations, and hence their estimations, are expected to change throughout the school year.

3.5 The analysis of imagistic representations

Within its overall focus on cognitive representations, this thesis has particular interest in children's cognitive representations with visual, spatial or graphical characters. As shown in the literature review, such images are of high interest in investigating children's representation of the natural numbers. In this section, I discuss how Duval's framework of cognitive representation requires coordination with additional theory in order to analyse this kind of representation in detail.

No single term in Duval's theoretical framework of representations captures representations with visual, spatial or graphical characters as a particular class of representation, nor does his analysis of representations focus upon the details of this kind of representation. In Duval's classification, representations with visual, spatial or graphical characters are best encompassed by the category of *mental images* combined with the category of *visualisations* (i.e. analogical semiotic representations). To recap briefly, the category of *mental images* includes two heterogeneous types of representation: the “internalized semiotic visualizations” (on the left of Figure 1) and the “quasi-percepts” (on the right of Figure 1) (Duval, 1999, p. 6). By combining these two categories from Duval, both internal and external representations with visual, spatial or graphical aspects are included. This thesis will refer to such representations as imagistic representations, to avoid confusion with terminology already defined by Duval in his framework, and for consonance with the work of Goldin and Kaput (1996) and Thomas et al. (2002), whose theoretical work is incorporated into this section of the theoretical framework.

The finer detail required by this thesis in the area of representations with a visual, spatial or graphical nature is provided by the imagistic framework of Thomas et al. (2002). This combines Goldin’s model of internal representations (Goldin and Kaput, 1996) with a model of imagery based upon ideas from Presmeg (1986, 1998). The study adapts these ideas specifically for the analysis of children’s imagery of the counting sequence, so has very good applicability to the representations studied in this thesis.

3.5.1 Thomas, Mulligan and Goldin’s theoretical framework

3.5.1.1 Note: using a framework developed for internal representations

Thomas, Mulligan and Goldin’s study is titled “Children’s representation …” and the focus is described to be “internal imagistic representations” (Thomas et al., 2002, p. 117). The aim of Thomas, Mulligan and Goldin’s study, unlike this thesis, is to use children’s external representations to draw inferences about their internal representations.

In this thesis, cognitive representations that are produced using external means are analysed as representations in their own right rather than as ‘windows’ onto internal representations. It is assumed, following Presmeg, Piaget and Inhelder, that external productions are guided by internal cognition, but the external productions are analysed as they are. There are both empirical and theoretical reasons for this. Empirically, cognitive representations in different modes (e.g. speech and gesture) have been found to represent differing, and even apparently contradictory, information about participants’ mathematical problem-solving (Garber and Goldin-Meadow, 2002), making the processing of inferring one underlying internal representation problematic. Theoretically, Duval’s work emphasises the specific nature of the register of representation (e.g. Duval, 1999, 2006) as noted in previous sections. It is not to be assumed that children will, or can, represent ‘the same’ information in different registers.

Whilst the analytic framework of Thomas, Mulligan and Goldin is designed to investigate internal representations, this of course necessitates analysing external representations, since as with any study of internal representation, direct access is impossible. The authors apply their framework to all external representations produced during the task-based interviews held: interview transcripts, and external pictorial and notational representations (Thomas et al., 2002). This thesis therefore adopts the framework of analysis of Thomas, Mulligan and Goldin, and differs only in not proceeding to inferences about the internal representation underlying children's external productions. It extends the use of the framework by including children's gestures.

3.5.1.2 Goldin and Kaput's theory of representation

Thomas, Mulligan and Goldin's study is theoretically framed by Goldin's model of representation, in which internal imagistic representation includes the sub-systems of visual/spatial, auditory/rhythmic and tactile/kinaesthetic representation. The sub-system focused upon by Thomas et al. is the visual/spatial system.

Goldin and Kaput (1996) define imagistic representation as follows:

> *Imagistic or analogic representational systems refer to systems in which the fundamental characters, signs, and configurations are neither verbal nor formal in nature, but bear some interpreted sensory resemblance to what is represented.*
> (Goldin and Kaput, 1996, p. 414)

Imagistic representations therefore include a broad range of representations: "internal imagery and image-schematic representation—that which is 'imagined', visualised, represented kinesthetically and /or auditorily", and also "external enactive and pictorial representations, concrete embodiments and manipulatives" (Goldin and Kaput, 1996, pp. 414-415). Like Duval's *visualisations*, internal imagistic systems in Goldin and Kaput's framework "incorporate nonverbal configurations at the level of objects, attributes, relations, and transformations" (p. 418). Again like Duval's visualisations, it is students' imagistic capabilities which are necessary for "meaningful or insightful" mathematical understanding (p. 418).

The major contribution of Goldin and Kaput's theory to the framework of Thomas, Mulligan and Goldin is the idea that imagistic representations develop through three particular stages (Goldin and Kaput, 1996, p. 424). The first stage, following Piaget (1969a), is the inventive-semiotic stage in which new characters are created or learned, and crucially, in relation to an existing representation system. This stage is often problematic in learning mathematics, because the new characters are taken to "be" rather than symbolise the aspects of the previous system – leading to cognitive obstacles. The second stage of development is the structural-developmental stage, in which development is driven principally by structural features of the earlier representation system. The third and final stage is the autonomous stage, in which the new system of representation is considered mature, and separates from the old representation system.

3.5.1.3 Presmeg's theory of visualisation and images

While interpreting results within the overall framework of Goldin's theory, Thomas, Mulligan and Goldin's study is informed by literature specifically upon imagery, primarily the work of Presmeg. Presmeg's research considers imagery in general, as well as visual images and visualisation specifically.

Presmeg notes that imagery may occur in "one or more of five modalities ... visual, auditory, tactile, gustatory, and olfactory" (Presmeg, 1992, p. 596). The majority of Presmeg's work has focused on visual imagery, and the functions within mathematical thinking that it can perform, for example how "visual imagery may serve the purpose of abstraction" (Presmeg, 1992, p. 596).

Presmeg defines a *visual image* to be "a mental construct depicting visual or spatial information" (Presmeg, 2006). This corresponds with internal imagistic representations in the visual/spatial sub-system of Goldin's framework: in each case the image is internal (or mental), and either visual or spatial in character. Presmeg explains that the definition is deliberately broad, including "kinds of imagery which depict shape, pattern or form without conforming to the 'picture in the mind' notion" in addition to 'pictures in the mind' themselves (Presmeg, 1986, p. 42). The definition allows for visual imagery to include spatial arrangements of "verbal, numerical or mathematical symbols" (Presmeg, 1986, p. 42).

Within the mode of visual imagery, Presmeg's work has identified five forms or categories: concrete imagery (like a picture); pattern imagery (relationships without concrete detail); memory images of formulas; kinaesthetic imagery (involving physical movement); and dynamic imagery (the image itself is moved or transformed) (Presmeg, 1986, 2006, 2008). Of these five categories, the first three refer to the type of component sign of the representation, and the latter two refer to properties which may be present in representations of various types, within various modes. In the framework of analysis constructed by Thomas et al. (2002), representations are first classified according to their component sign, and then later classified according to their dynamic/static nature. Kinaesthetic imagery, whilst acknowledged and present, is not a basis for classification or analysis in Thomas et al. (2002); whether images include or do not include kinaesthetic aspects is not the focus of the work.

Concrete images include the archetypal "pictures in the mind" associated with the word 'imagery' (Presmeg, 1992, p. 596) and also "memory images" (p. 599). They are associated with imagery as prototype: in this function, individuals use their concrete images "to represent wider categories, concepts, or principles" in the course of mathematical problem-solving (p. 599). Prototypical images may also be used as metaphors (p. 599). Pattern imagery, by contrast, describes the imagery in which "concrete details are disregarded" (Presmeg, 1992, p. 602). In pattern imagery, what remains is a visual-spatial scheme depicting "pure relationships", a type of imagery "strikingly illustrated in the memory images of chess masters" (p. 602). In the category of memory images of formulae, fairly clearly, the component signs are mathematical symbols remembered as images.

In terms of the mathematical roles that imagery is able to perform, Presmeg identifies two ways in which imagery can depict mathematical abstraction. The first is "by concretizing the referent - that is, by making a concrete visual image the bearer of abstract information" (Presmeg, 1992, p. 603). Examples of this include the memory image of a formula, or the concrete image of a see-saw to metaphorically depict equality. The second way in which imagery depicts mathematical abstraction is by "by using pattern imagery which embodies the essence of structure without detail" (Presmeg, 1992, p. 603). Both functions of imagistic representation are of course clearly present in Duval's theoretical work. From Duval's perspective, all conscious cognitive representations, and so certainly the subset of imagistic representations, perform the function of objectification ("En ce sens, la conscience se caractérise par la visée de 'quelque chose' qui prend *ipso facto* le statut d'objet pour le sujet effectuant cette visée" (Duval, 1995, p. 24)). The second function, embodying "essence of structure without detail" (Presmeg, 1992, p. 603) in turn could be encompassed by Duval's generalised description of mathematical visualisation (Duval, 1999, p. 13).

Although visual imagery is defined to be a mental construct, in terms of imagery, generally Presmeg, like Duval, asserts that the internal/external dichotomy "does not seem fruitful" (Presmeg, 2008, p. 2). In *An Overarching Theory for Research in Visualization in Mathematics Education* therefore, although the examples Presmeg (2008) examines "concern almost exclusively the external mode, called *inscriptions*", the taxonomy of inscriptions developed in the paper "might well be applied to the corresponding forms of visual imagery as well" (p. 2).

Presmeg's definition of visualisation reflects the fact that the internal/external dichotomy is not key: "visualization is taken to include processes of constructing and transforming both visual mental imagery and all of the inscriptions of a spatial nature that may be implicated in doing mathematics" (Presmeg, 2006, p. 206). This definition corresponds well to Duval's, in terms of not restricting visualisation to either mental or external representation. The examples of visualisation identified by Duval (drawings, sketches, graphs, figures, schema, (Duval, 2000)) remain visualisations by Presmeg's definition. Presmeg's definition is however more explicit in stating that visualisations represent by *visual* or *spatial* means.

This chapter has set out the theoretical ideas providing the framework for this thesis, namely Duval's theory of cognitive representations, and the theoretical approaches to imagistic representation brought together by Thomas et al. (2002) in their work on children's imagistic representation of natural number. The chapter has explained how these theories support the current research, and critically examined the ideas within them. The following chapter explains the research methodology and methods adopted, and why they are the appropriate choices for this theoretical framework and research questions. This methodology chapter will also lay out the theory for analysis, including most importantly Resnick's (1983) account of number development, and the plan of analysis that implements this theory.

Chapter 4 Methodology

This chapter presents the methodological reasoning that led to the chosen research design. The methodological choices made were influenced by three main considerations. First of all, methodological choices were shaped by the implications of the underlying theoretical perspective, which includes the research focus on cognitive representations. Against this background, choices were then made to meet the particular demands of the research problem identified and research questions posed. A final and significant influence on the methodological choices was the wish to contribute new knowledge to the research area by adopting a methodology and methods which deliberately complemented the approaches of prior research.

4.1 Overview of research design

The methodological approach used for this research can be summarised as an exploratory longitudinal multiple case-study approach, within which video-recorded task-based interviews are used. The research project consists of twelve case studies, where each case is an individual child within a Year 1 classroom at a local primary school. The children worked with the researcher over the course of one school year.

4.1.1 An exploratory design

Yin argues that all research strategies can be used for three purposes: exploratory, descriptive or explanatory, and that it is a misconception to consider certain strategies appropriate only for certain purposes. Strategies should instead be distinguished by the three following conditions: "(a) the type of research question posed, (b) the extent of control an investigator has over actual events, and (c) the degree of focus on contemporary as opposed to historical events" (Yin, 1994, p. 4).

In the case of this research project, the research questions and their relation to the literature determine that an exploratory study is appropriate. The questions are about how previous findings relate to each other and about looking at the research topic from a new viewpoint: focusing on the developmental trajectory of individuals, in contrast to changes observed between cohorts; and looking in depth at the process of children's number line estimations, in contrast to focusing only upon the result of estimations. The questions are formed in terms of "how" and "in what ways", and lead naturally to an exploratory study. Exploratory research is one of the situations in which "all research strategies might be relevant" (Yin, 1994, p. 9), so the following section explains the methodological considerations that led to the choice of a longitudinal multiple case study.

4.1.2 A longitudinal design

The decision to pursue a longitudinal study stemmed primarily from the developmental aspect of the research questions, and the wish to complement previous, largely cross-sectional, research in the field.

The developmental focus of the research questions was directed by both the existing literature, and the theoretical framework of the thesis. Several parts of this framework suggest a longitudinal approach, because they hypothesise qualitative changes in how children cognitively represent numbers. The principal theory, Duval's theory of cognitive representations, identifies children's number line estimations (amongst other events) as *conversions* of representations, which can be expected to change as children's knowledge of particular representations of numbers increases during the school year. Fuson's account of children's developing number knowledge (see Analysis section) indicates the order in which children typically grasp particular aspects of natural number; together with Duval's theory, this leads to the hypothesis that as children begin to 'see' different aspects of number in particular representations, their ability to convert representations of number into a different representational system will change.

Goldin and Kaput's (1996) theory of representations as adapted by Thomas et al. (2002) suggests that children's new representations can be expected to pass through three stages: inventive semiotic construction, structural development, and finally autonomous use. Thomas et al.'s study, like the majority in this field, was cross-sectional in design, and the authors called for further longitudinal investigation.

Longitudinal or cohort studies are in many ways the obvious choice for researchers studying aspects of human development. By investigating the same sample over an extended period of time, they enable within-individual comparison over time and hence change to be analysed at the individual or fine-grain level. Longitudinal studies with statistically representative samples are "uniquely able to identify typical patterns of development and to reveal factors operating on those samples which elude other research designs" (Cohen, Manion and Morrison, 2011, p. 269). Even without a statistically representative sample, longitudinal studies "permit researchers to examine individual variations in characteristics or traits, and to produce individual growth curves" (p. 269).

This ability to capture individual variations is precisely what the research questions of this project seek to explore, and what previous cross-sectional studies have not been able to illuminate. Particularly relevant for this research project is that longitudinal studies enable not only change, but the *nature* of change to be captured; they permit "the dynamics of change to be caught, the flows into and out of particular states and the transitions between states” (Cohen et al., 2011, p. 272). In the context of the current research, this is highly relevant. Duval’s theory of representations, framing individual estimations as conversions of representations, focuses the attention of this thesis on a far more detailed view of children’s cognition than previous studies in number line estimation, which typically examine change by aggregating the estimations of multiple children, and identifying change in the sense of cohort differences only.

A particular example of how the longitudinal advantage will be used in this thesis is in the exploration of Siegler’s theory of overlapping waves (e.g. Opfer and Siegler, 2007), which holds that children have access to multiple cognitive representations of number in a given situation at any one time, and that children’s preference for a new representation may appear as ‘waves’ of usage that overlap with usage of the older representation. This theory was developed from Siegler and colleagues’ work on number representation, i.e. specific to the context of this thesis, but has not been investigated with respect of the representations of individual children – it remains an hypothesised account from cross-sectional data. The theory is one of several predictions informing the data analysis of the current research, and it is only through a longitudinal design that the theory’s ability to describe developmental change in individuals can be further investigated.

For researchers studying change more generally, Cohen et al. (2011) additionally note that “Individual level data are more accurate than macro-level, cross-sectional data” (p. 272) and that sampling error is lower than in other research designs due to the fact that the same sample is used throughout the study. In researching school-age children, longitudinal studies also provide valuable records deriving from “the known fallibility of any single test or assessment” (p. 269).

Despite these advantages, there are few examples of longitudinal studies in the literature on children’s changing representations of number (DeWindt-King and Goldin, 2003). Instead, the majority of studies so far carried out in this area have been cross-sectional in design. Thompson and Opfer (2010) describe a typical research approach in explaining that they “investigated long-term changes in children’s estimates of large numerical magnitudes… by examining estimates of second graders, third graders, sixth graders and adults” (p. 12). In this and the majority of studies in the field (e.g. Berteletti et al., 2010; Booth and Siegler, 2006; Ebersbach et al., 2008; Halberda and Feigenson, 2008; Laski and Siegler, 2007; Mundy and Gilmore, 2009), researchers have investigated long-term changes using purely cross-sectional methodology.

The situation is similar in development research more generally, and Cohen et al. (2011, p. 270) ask the obvious question: given that longitudinal studies are “particularly appropriate in research on human growth and development", why is it that so many studies in this area are cross-sectional? There are five principal reasons: cross-sectional designs are less expensive; they are usually able to include more subjects than a longitudinal design; findings are available more quickly; researchers find it easier to recruit participants for a 'one-off' occasion; and cross-sectional designs are less likely to suffer from the 'measurement effect' since measurements are carried out only once.

The specific weaknesses of cross-sectional designs correspond to the advantages of longitudinal work. Cohen et al. (2011) note that cross-sectional designs are "Unable to chart individual variations in development or changes, and their significance" (p. 273), and furthermore that "Sampling in the cross-sectional study is complicated because different subjects are involved at each age level and may not be comparable" (p. 270). These and other objections weigh strongly against cross-sectional studies "so much that one observer dismisses the method as a highly unsatisfactory way of obtaining developmental data except for the crudest purposes" (p. 270).

4.1.2.1 Disadvantages of a longitudinal design

A commonly cited problem with the use of longitudinal studies is the problem of 'sample mortality', that is the inevitable fact that during the course of a long-term study, some subjects are likely to leave the study. This causes a problem for the research in making it "unlikely that those who remain in the study are as representative of the population as the sample that was originally drawn" (Cohen et al., 2011, p. 270). In the case of this research, there are specific reasons why this objection should not outweigh the advantages of the longitudinal approach. In the first place, the research design consists of a small number (12-15) of case studies, rather than a sample which is intended to be statistically representative of any single cohort. The second point is that at 10 months long, this longitudinal project is relatively short in length, and is being conducted within one classroom at one school, in which the rate of pupil turnover is low and in which the researcher has good relationships with participants. For these reasons, although it is to be expected that some participating children will not take part in the full study, the proportion of participants leaving the study should be small.

A more worrying objection to a longitudinal design is what has been referred to as the 'measurement effect'. This is the effect whereby repeated interviewing itself influences the behaviour of participants, for example by "sensitizing them to matters that have hitherto passed unnoticed" (Cohen et al., 2011, p. 270). In the case of this project, participating children take part in similar or identical task-based interviews five times within one school year, and so the 'measurement effect' is a threat to be taken seriously. Specific steps were taken to minimise the disruption to results posed by the 'measurement effect':

1. The task-based interviews are held only once per half term, ensuring a gap of between 6-10 weeks between measurements. At an age at which children change rapidly, this gap represents a fairly significant period of time in which to forget about the previous interview.
2. Participating children are given no feedback on the 'correctness' of their interview answers. The interviewing researcher encourages their efforts with neutral comments and thanks, ensuring that children are not 'trained' at the tasks (although some familiarisation will still take place).
3. Where specific details of tasks (for example, target numbers) can be changed between interviews, they are changed, to ensure variety.

There are also specific aspects to the research and research context which limit the impact that a 'measurement effect' could have upon results. Firstly, participating children are carrying out tasks in classroom mathematics lessons throughout the study which deliberately 'sensitise' them to precisely the concepts being investigated: the relationship between numbers, and representations of the counting numbers – external and internal, in visual, spatial, symbolic, verbal, imagistic and enactive forms. Secondly, the research is concerned with the nature of change, rather than only whether change does or does not occur. It is entirely expected that participants become more sensitised to aspects investigated in the task-based interviews (for example, the relation between counting numbers up to ten), because this is an explicit goal of the teaching they receive during the school year. In comparison to this everyday teaching, the sensitisation caused by the 'measurement effect' itself seems likely to be minimal.

A practical objection to longitudinal studies is that their data "being rich at an individual level, are typically complex to analyse" (Cohen et al., 2011, p. 272). However, the advantages afforded by the longitudinal design for this particular research problem are so significant that the inconvenience or 'price' of complex data is willingly accepted. Another common objection to longitudinal research is that it is time-consuming and expensive. In the case of this research, the length of study is already constrained by the timescale of the PhD programme, but within this constraint it is perfectly possible to dedicate the necessary time and resources to the longitudinal study.

4.1.3 A multiple case study design

The decision to adopt a multiple case study design followed naturally from the reasons behind the choice of longitudinal approach. In summary, the motivation was to choose the approach which best afforded in-depth analysis of individual developmental trajectories. Such an approach would best allow new understanding of the nature of the representational changes previously hypothesised, and complement the majority of existing studies in the field, which had focused on cohorts of children rather than individuals. The fine grain detail afforded by case studies makes them ideal for this purpose; “to complement other, more coarsely grained – often large scale – kinds of research” (Cohen et al., 2011, p. 291).

In terms of the theoretical framework of research, Duval’s theory of cognitive representations focuses firmly upon fine-level detail and fits well with the in-depth perspective afforded by case studies. The identification of individual number line estimation trials as conversions of representations demands that the researcher examine the *processes* involved in estimation tasks rather than only the results. The hypothesis of White and Szucs (2012), that individual numbers may be estimated differently according to a whole set of variables, similarly provides strong motivation to examine the *processes* of estimations; there is very significant potential for patterns and differences to be obscured when only results are considered, and this is hypothesised to have occurred in many previous studies.

The primary justification for the use of case study designs is that case studies can capture some aspect – particularly of people – which is missed by other approaches. As Cohen et al. (2011) point out, “human systems have a wholeness or integrity to them rather than being a loose connection of traits, necessitating in-depth investigation” (p. 289). The intensity of observation and analysis required for in-depth investigation leads the researcher to focus on a small number of cases. A potential criticism of case study research is that the findings are too particular to cases, cannot achieve statistical significance, and have low generalizability. The response to this is that case study research sacrifices quantity of cases for the intensity of analysis of each case. Additionally, the case study researcher has fewer restrictions on the type and depth of data that can be collected, allowing them to “penetrate situations in ways that are not always susceptible to numerical analysis” (p. 289).

Underlying case study research is the belief that by researching and capturing the complexities of one instance, *insights* will be reached which in fact are likely to be generalizable. In studying representational development through case study, trajectories of change can be understood that would not have been captured through less fine-grained examination, and these trajectories of change are not limited to the specific individuals within whom they are studied (Cohen et al., 2011). An appealing characterisation is that a case study is “the study of an instance in action” (Adelman, Kemmis and Jenkins, 1980).

A case study design “is particularly valuable when the researcher has little control over events” (Cohen et al., 2011, p. 290). This is a consideration that (Yin, 1994) argues should be influential in deciding research design, and is certainly applicable to the case of children's cognitive representations of number. It is assumed that these cognitive representations will change, and that the researcher cannot control what affects them; they will be affected by explicit and implicit influences both inside and outside the classroom throughout the whole year.

Some "what"-type research questions suggest the use of a survey methodology (Yin, 1994) but there are clear objections to a survey strategy in this case, stemming from the complexity of the phenomena under investigation. There do not exist established 'normal science' methods for investigating the phenomena involved in the research problem, and the literature so far suggests that in order to gain further insight into the research problem, richness and depth of data are required, something a case study strategy is better fit to provide.

A further methodology that could be used is an experimental design. There is good evidence that children's cognitive representations of number can be strongly influenced – see for example studies in which alignment tasks were used to encourage the use of linear representations (e.g. Thompson and Opfer, 2010). However, this thesis aims to study children as they develop – it does not *want* to control changes in children's representation. The perspective of this thesis is that children's cognitive representation of number is a part of their growing understanding of number, and that it is affected by many factors: child development, formal teaching, and everyday experiences. From this perspective, the degree of available control is very low. Whilst previous researchers have carried out experimental work with cognitive representations, the research goals of this thesis would not be well-served by an experimental design.

4.2 Task-based interviews

Although case study research is commonly associated with qualitative research, the fine grain detail it collects can be quantitative or qualitative (Cohen et al., 2011). In the case of this research, I will collect and analyse both detailed quantitative and detailed qualitative data about each participant or case.

This section addresses the choice of task-based interviews as the main data collection tool. The description *task-based interview* is used following usage of previous researchers in representation (DeWindt-King and Goldin, 2003; Thomas et al., 2002). One advantage of this is to avoid falling into the strict definition of *clinical interview* or the *talk-aloud* procedure as defined by Ginsburg et al. (1983). As will be explained, the task-based interview developed for this thesis uses elements of each, chosen with extreme care, for reasons based on theoretical assumptions.

The data collection tool for this thesis needed to permit observation of children using and creating cognitive representations of number. Since the research questions require the investigation of number line estimation tasks, these tasks in particular needed to be incorporated. Since the thesis is concerned with the close analysis of individual responses, it was determined that a class or group class situation was not appropriate for the main data collection, a common conclusion in cognition research (Ginsburg, 1981).

Two particularly relevant studies in mathematics education demonstrate the potential of the task-based interview (DeWindt-King and Goldin, 2003; Thomas et al., 2002). In both cases, children were interviewed individually and the researchers carried out fine-grain analysis of children's cognitive representations. The task-based interviews enabled close analysis of children's overt behaviour, within which representations were identified and analysed. The task-based interview clearly permits observation, and in addition video-recording of the interview, to further enable detailed analysis.

A task-based interview incorporating number line estimation tasks was therefore chosen as the data collection tool. The following sections specify the exact features of the task-based interviews developed, and the reasons for this specification.

4.2.1.1 Clinical interviews and talk-aloud procedures

As briefly mentioned, the task-based interview is related to two established data collection tools: the clinical interview and the talk-aloud procedure. The original, verbal clinical interview method consist of "flexible questioning of individual children on a totally verbal level" (Ginsburg et al., 1983, p. 10). This method does, clearly, only provide verbal data, and according to Ginsburg the revised clinical interview method was developed by Piaget after he concluded that the verbal-only method was in certain situations inadequate. In the revised clinical interview method, concrete objects/tasks are incorporated, and the data collected are "both verbalizations and aspects of nonverbal behaviour" (Ginsburg et al., 1983, pp. 10-11). In terms of contributing to the understanding of cognitive development, the rationale of the clinical method is that it offers children "the opportunity to engage in various intellectual activities" whilst the researcher observes all aspects of their behaviour (Ginsburg et al., 1983, p. 11).

The talk-aloud method similarly sets an individual a task, but instead of being questioned, "the subject is instructed to say everything that comes into his or her head whilst solving a challenging problem" (Ginsburg et al., 1983, p. 8). The aims are to "elicit and describe the integrated activities constituting complex problem solving" with a minimum of researcher intervention.

Reflecting on the features and comparative advantages of these two protocol methods was important in order to develop the most appropriate data collection tool for this thesis. Questioning, in line with the revised clinical interview method, presents numerous advantages. There is explicit mention by Thomas et al. (2002) that by only recording spontaneous first attempts at a drawing task, the researchers accessed only a partial view of children's representational capabilities – "Other responses would very likely have occurred had the children had been prompted" (Thomas et al., 2002, p. 130).

On the other hand, there are good arguments against questioning in this research area, since it is not only researchers who are unable to directly access participants' internal representations. Goldin emphasises that the extent to which individuals can introspectively describe their *own* internal representation is also highly "questionable" (Goldin, 2008, p. 181). To this must be added the limitations of young children's verbal articulation. Furthermore, by asking children "how did you do that?" a researcher may easily bias results, by implying to children that an explicit, communicable strategy is expected and recommended. This is a particular danger in the context of this thesis, since both intentional and automatic cognitive representations are hypothesised to be involved.

The solution adopted was to limit direct questioning, and to invite but not require children to comment on their task solving. This attempted to combine some of the benefits of the talk-aloud method with the clinical interview method, whilst not obliging children to report or invent accounts of their actions to satisfy researcher interest. Children were reminded several times during the interview, "If you want to, tell me about what you're doing as you go along". Where children began an explanation and did not finish (e.g. "I know this one, because …"), or made specific comments about their attempts at a trial ("This is a really easy one!"), the researcher followed these up, as in the revised clinical interview method, but with questions deliberately conversational in tone. The examples given above could typically be followed up by "Because …?" or "Oh really? Why's that?", or simply facial communication between the researcher and participating child. Strong relationships between the researcher and participating children were established before the pilot study, and this seemed to make little prompting necessary.

The aims of the compromise were as follows:

- To minimise researcher influence where possible
- To leave open as many possibilities for the participants as possible, and imply as few expectations as possible (although by repeatedly asking children to complete certain tasks, the researcher will implicitly communicate an interest in children's estimations)

- To minimise the pressure on children to 'invent' accounts to satisfy researcher interest
- To maintain a relaxed interview atmosphere, in which children felt able to work at their own pace
- To encourage a rapport with research participants that encourages free talking
- To retain children's confidence (established in classroom interaction) that the researcher was interested in 'how we do maths' rather than in 'getting the right answers'

The downside of cautious questioning is that participants will not be pushed for a demonstration of full representational capability. This is, however, countered by specific steps: multimodal data collection, multiple tasks, repeated multiple times in an interview, and carried out on multiple occasions throughout the year. These steps are detailed in the following section.

4.2.2 Shaping the task-based interviews

The specific requirements of the research questions, and clear recommendations from the literature, determined the necessary features of the data collection tool. In summary, it needed to:

1. Include data from more than one task, situation or question
2. Examine cognitive representations involving more than one mode
3. Include data on children's responses to number line estimation tasks
4. Include data from sufficient tasks, situations, questions and times to gain insight into the variability of children's representations

These requirements, and the reasoning behind them, are explained more fully in the following sections.

4.2.3 The need for multiple tasks/situations

In viewing the individual as possessing representational capabilities, it is necessary to use more than one tool in order to more accurately describe and assess capabilities. DeWindt-King and Goldin (2003) recommend proving varied external representation opportunities, and there are both empirical and theoretical reasons for doing so.

Fuson and Hall (1983) emphasise that children demonstrate and successfully use wildly varying levels of knowledge when task designs prompt the use of particular knowledge, e.g. counting. This idea is evidenced in the field of number representations by findings such as those of Thompson and Opfer (2010); even within the strictly defined number line estimation, individuals appear to employ different numerical-magnitude representations depending on the particular numerical context, in this case, the scale indicated by the endpoints of the number line. The idea of different tasks eliciting different cognitive representations in the same individual is further supported by Siegler and Opfer (2003), Hubbard et al. (2009), and Geary et al. (2008).

From a more theoretical standpoint, Vergnaud (1987) concludes that “it is a good theoretical and methodological choice to study a set of situations” since “a concept refers to more than one kind situation and as the analysis of a situation requires usually more than one concept” (p. 231). At a more specific level, Duval also offers theoretical reasons to vary task situations. Research involving the conversion of representations, as in this thesis, "requires that students be given tasks that are varied systematically not only as a function of the original register but also as a function of internal variations within each register" (Duval, 2006, p. 121). In this case, this means that both the target number and register of the starting register are varied, and the endpoints of the number line target register are also varied. Additionally, the interviews developed for this thesis use two different formulations of number line estimation task: both the ‘number to position’ task and the ‘position to number’ task, requiring opposite conversions of representations. According to Duval, this methodology of variation then permits the observation of “a systematic variation of performances" (Duval, 2006, pp. 121-122). The thesis also incorporates two entirely separate task situations: an imagistic drawing task and the estimation of quantities of items (see detail on tasks in later section).

A further motive for using multiple tasks is the questionable validity and reliability of single hypothesised measures (Schneider and Stern, 2010). The question of reliability is particularly relevant considering the inherent variability and idiosyncratic meaning of children’s communications. This thesis takes the perspective that children’s variability should not be seen as a failing, but does strengthen the case for multiple data collection tools, and should be incorporated into the research conclusions.

Schoenfeld (2008) discusses the particular issue of construct validity in clinical interviews. The example of Piaget’s clinical interviews is discussed, where further research revealed that “although performance on certain tasks might be robust, the robustness was in part a function of the research design; other tasks aimed at the same mental constructs did not necessarily produce the same results …” (p. 486). Schoenfeld stresses that the analysis *within* the task situation was correct; it was “the mapping back to the conceptual framework (the attribution of certain logico-deductive structures on the basis of the analyses)” that had questionable validity (p. 486). This problem is one that further motivates this thesis to *not* map data back to the construct of concepts, but to focus upon cognitive representations, as Duval’s theoretical work strongly recommends.

By using multiple tasks, varied task details, multiple trials, multiple interviews, and acknowledging children’s variability, the research design of this thesis aims to provide insight into children’s cognitive representations that is both valid and reliable. The issue is a problematic one for research into the cognitive representation of number; as the literature review demonstrates, there is at present reliance on a too-narrow range of data, and a lack of consensus about the validity of competing accounts.

4.2.4 Examining in multiple modes

The primary motivation for considering multiple modes when examining cognitive representations is theoretical; Duval's framework explicitly notes that cognitive representations may appear in many different modes (e.g. Duval, 1999, pp. 5-7), and there is no theoretical reason to privilege one above another. Focusing on only a limited subset would limit the cognitive representations of number that this thesis is able to detect. In addition to this, Duval's theoretical perspective identifies the conversion *between* registers, and often consequently between modes, as the critical aspect of mathematical understanding. This gives further motivation to detect representations of number in as many modes as possible, in order that conversions are properly noted.

From many theoretical standpoints, examining multiple modes performs the methodological function of triangulation, but precision must be used in specifying how. In studies which infer an underlying internal representation from children's external representations (e.g. DeWindt-King and Goldin, 2003), data from multiple modes fulfils a triangulating function in terms of increasing the validity of claims made about a child's underlying internal representation of number. This is not the case in this thesis, where observed cognitive representations are considered in their own right, not used for the inference of internal representations (see Theoretical Framework chapter). Empirical evidence indicates that cognitive representations in different modes may represent different information about a mathematical problem altogether; in some circumstances, even apparently conflicting information (e.g. Alibali and Goldin-Meadow, 1993). Despite this, representations observed in multiple modes are still able to fulfil a function of triangulation for this thesis, in terms of validating claims about what aspects of number a particular child is able to cognitively represent. For example, a child may repeatedly appear to spatially represent the number sequence 10, 9, 8, 7, 6, 5 in gestures pointing at the page. The claim that the child is able to cognitively represent this number sequence is supported by the observation of the representation of the same number sequence in other modes, such as natural language or inscription.

4.2.4.1 Gesture

Chu and Kita (2011) observe that the majority of research into gesture focuses on co-speech gesture only. In the influential *Hand and Mind* for example, there exists an "obligatory presence of speech" in the definition of gesture (McNeill, 1992, p. 37); movement that is not accompanying speech is not gesture. McNeill acknowledges that “Many authors refer to all forms of nonverbal behavior as ‘gesture’” but views this as a failure to differentiate between behaviours that “differ fundamentally” (McNeill, 1992, p. 37).

For the purposes of this thesis, a broader understanding of gesture is required, specifically one that is able to encompass co-thought as well as co-speech gesture. Considering the observed behaviour of children working on tasks, where there is a continuum of loud talking, quiet talking, mumbling, whispering, and mouthing silently, it is difficult to accept that the audibility of the words turns the accompanying gesture into a behaviour that “differs fundamentally”. A more inclusive definition of gesture is the continuum conception of gesture by Kendon (1988), which includes the following forms of movement:

Gesticulation -> Language-like gestures -> Pantomimes -> Emblems -> Sign Languages

The movements farther to the right on the spectrum are less likely to be accompanied by speech, and more likely to resemble language in the way used (e.g. replacing a word in a spoken sentence). McNeill’s definition of gesture as “idiosyncratic spontaneous movements of the hands and arms accompanying speech” (1992, p. 37) corresponds to *gesticulation* in Kendon’s continuum.

4.2.4.2 Why analyse gesture?

The principal reason to include gesture within the data collection of this thesis is that Duval’s (1999) theoretical framework, as noted, emphasises always that the important aspects of cognitive representations are their nature and form, rather than mode of production. There is no theoretical justification for excluding a mode such as gesture, which is as likely as others to be used in cognitive representation.

Additionally, however, there are reasons for particular interest in gesture. Although it is very commonly analysed, speech is likely to be influenced by the wording of problems or interviewer questions, and to systematically omit any information that is difficult to verbalise. Since spontaneous gestures are not subject to these same limitations (though they may be subject to others), their analysis may provide a window onto knowledge that is not readily expressed in speech (Alibali, Bassok, Solomon, Syc and Goldin-Meadow, 1999, p. 327). This is a particularly relevant consideration when researching the knowledge of young children.

Empirical research suggests that gestures not only reveal important information about people’s representation of problem situations, but that speech and gesture together provide a more complete view of solution strategies than speech alone (Alibali et al., 1999). A phenomenon of particular interest is that gesture-speech mismatches in children may reveal transitional states of knowledge (Alibali and Goldin-Meadow, 1993; Perry, Breckinridge Church and Goldin-Meadow, 1988; Perry et al., 1992). Alibali and Goldin-Meadow (1993) for example found that mathematically accurate representations of a problem were found in gesture *before* speech for every child sampled, making gesture an exciting place to look for emerging conceptual knowledge.

Particularly relevant for this thesis and the conversion of representations is the demonstration by Garber, Alibali and Goldin-Meadow (1998) that the knowledge represented with gesture is not "tied" to the hands. In their study, Garber et al. (1998) showed that the knowledge children conveyed in gesture was also represented using other means, and used by children in solving other tasks.

The use of video assists in the capture of multimodal data, and enables more detailed analysis than permitted by in-the-moment observation. It is used to capture speech and sound, eye movement and gestures. It also captures body language and other action, though the literature does not suggest that these will be frequently involved in the representation of number.

4.2.5 Number line estimation tasks

There are clear reasons why the task-based interviews for this thesis need to incorporate number line estimation tasks: they are of high theoretical interest, consequently extremely common in number representation research, and a specific focus of the research questions of this thesis. This thesis uses number line estimation tasks for the data they are able to provide on children's cognitive representations of number, but additionally seeks to investigate the tasks themselves in order to engage with existing literature and hypotheses.

A 'number line estimation task' is a specific problem type in which participants are asked to indicate the position of a given number on an empty number line (a blank line on the page or screen, with the endpoints labelled, often with the values 0 and 100, or some other power of ten). The method described here by Opfer and Siegler (2007) is typical:

> *Each problem consisted of a 25 cm line, with the left end labeled "0" and the right end labeled "1000." The number to be estimated—2, 5, 18, 34, 56, 78, 100, 122, 147, 150, 163, 179, 246, 366, 486, 606, 722, 725, 738, 754, 818, and 938—appeared 2 cm above the center of the line. […]*
>
> *The experimenter began by saying, "Today we're going to play a game with number lines. What I'm going to ask you to do is to show me where on the number line some numbers are. When you decide where the number goes, I want you to make a line through the number line like this (making a vertical hatch mark)." Before each item, the experimenter*

said, "This number line goes from 0 at this end to 1000 at this end. If this is 0 and this is 1000, where would you put N?"
(Opfer and Siegler, 2007, p. 175)

Number line estimation tasks such as this are very widely used by researchers in cognitive psychology for studying the development of whole number magnitude representations (Siegler et al., 2011). A specific gap in the literature is detailed knowledge about what this task reveals about children's cognitive representations. It has been very frequently used, but insufficiently queried, so presents a key target for further investigation. The impact of children's *strategy* on responses is one obvious aspect for examination (White and Szucs, 2012).

Numerical estimation is of interest to researchers being not only pervasive in everyday life and education, but central to "a wide range of mathematical activities" (Opfer and Siegler, 2007, p. 170). Research over the past 25 years has found attainment at numerical estimation tasks to correlate well with specific mathematical skills (arithmetic, numerical comparison) as well as standardised mathematics test scores (Opfer and Siegler, 2007, p. 170). Clements and Sarama argue that "Improving children's number line estimation may have a broad beneficial effect on their representation, and therefore knowledge, of numbers (2009, p. 45).

Number line estimation tasks are considered revealing because they require genuine numerical estimation (not solvable through external, non-numerical clues). Specifically, solving a number line estimation task "requires translating a number into a spatial position on a number line or translating a spatial position on a number line into a number" (Young and Opfer, 2011, p. 59). It is this translation, or in Duval's terms *conversion*, which makes number line estimation tasks valuable to this thesis. The interest in representation translation or conversion differs according to theoretical framework. Working within Duval's framework, the number line estimation task is seen to involve a conversion between on one hand simultaneous representations of number in spoken natural language and written numerals, and on the other hand a spatial representation on paper. In contrast, the theoretical frameworks of many studies using this task type (e.g. Barth and Paladino, 2011; Berteletti et al., 2010; Thompson and Opfer, 2010) focus upon an underlying (and privileged) internal representation, to which to the external number line estimation is held to correspond. From this perspective, the conversion is a translation between "external symbolic representations of numbers (e.g., numerals) and internal, analog magnitude representations" (Young and Opfer, 2011, p. 59). In both cases, the translation or conversion is "highly revealing about how a cognitive system encodes number" (Young and Opfer, 2011, p. 59).

A concern with the validity of the use of number line estimation tasks is that the presentation of the task 'forces' participants into an external representation of a very specific and limited type, and that this is inadequately discussed in reports of its experimental use. There is strong evidence for associations between spatial and numerical representation in the brain (e.g. Fias and Fischer, 2005), but the exact nature of the association remains ambiguous. Furthermore, were it a known fact that the mental representation of number consisted exactly of a mental number line, for all participants, it would remain methodologically unjustified to assume a direct and transparent relationship between the mental number line and the task's (external) number line. In this thesis, the use of number line estimation is justified because there is no assumption that a number line representation corresponds to children's preferred, natural or otherwise privileged internal cognitive representation. Instead, the research is concerned with the process of children's conversion *to* the number line representation, and the aspects of natural number represented *when* children use the number line estimation.

A second concern regarding number line estimation tasks is that in a majority of studies, the positioning of target numbers with anything other than linear placement is classified as incorrect. A typical justification explains that "Just as 80 is twice as large as 40, so the estimated location of 80 should be twice as far from 0 as the estimated location of 40" (Siegler et al., 2011, p. 5). This claim, however, can only be true when the only *possible* mapping of natural numbers onto space is linear. It is a concern that a particular external representation must be produced for 'successful' task completion, despite the fact that no task instruction requests the production of this particular external representation. This thesis does not seek to judge the 'correctness' of children's number line estimations.

There are, nevertheless, reasons for a linear placement to be considered a more mature, and useful, representation, and so the linearity of children's representations does form part of the research questions of the thesis. For theoretical reasons, however, this thesis' interpretation of number line estimation responses will differ to those in the literature. For example, based on children's (approximately) logarithmically distributed number line estimations, Thompson and Opfer (2010) refer to "Children's initial expectations that numerical magnitudes increase logarithmically" (p. 6). From the theoretical perspective of this thesis, such a conclusion is not justified. In addition to the inference from number line task to mental representation, the assumed relation between representation and conceptual understanding is not discussed, and neither is the selective nature of representation – not every representation will represent every aspect of the represented. It is not obvious that children (before learning) know that the researcher in this task seeks an accurate representation of the 'spacing' of discrete numerical magnitudes. It is also conceivable that a child could possess the 'conceptual knowledge' that 80 is in some sense twice as 'far' from zero as 40, whilst not yet possessing the procedural knowledge that facilitates representing this understanding on paper, in the form specified by the task framework.

4.2.6 Data to assess variability

Three elements of the research design attempt to enable the observation of the variability of children's responses. Firstly, the longitudinal design allows children's performance on a highly specified set of tasks to be compared at five time points within one school year. Secondly, multiple tasks are used within an interview, inviting children to make similar estimations in different settings. Within each task, the variables of target number and number range are varied to give children multiple opportunities with each combination. Thirdly, if each task is considered as one task type (e.g. number to position estimation task) then the multiple trials (37 per estimation task) give children multiple opportunities to engage with that task type in each interview.

4.3 Research design: specific description

The chosen research design consisted of multiple longitudinal case studies, exploratory and qualitative in approach, which used video-recorded task-based interviews to collect data at intervals over one school year.

4.3.1 Sample and choice of age group

This thesis is concerned with children's changing cognitive representations of natural number as their mathematical understanding of natural number is still rapidly developing. The particular research interest in number line estimation tasks focused attention on the early years of schooling, where children's responses to these estimation tasks have been observed to change.

Experimental evidence so far suggests that the logarithmic-to-linear shift on number lines marked 0-100 occurs between ages 5 and 8 for US school students (Siegler and Booth, 2004), with similar shifts on larger number ranges occurring at later ages (Opfer and Siegler, 2007; Siegler and Opfer, 2003; Siegler et al., 2009). Data from studies with English school students is lacking, but it would not be surprising for these students to demonstrate linear representations of number on average somewhat earlier than American students, given the younger age at which they begin school (age 5-6 in the US compared to age 4 in the UK). Logarithmic-to-linear shifts have already been observed on smaller number ranges (e.g. 0-10) for Italian children aged 3½ to 6 (Berteletti et al., 2010).

An additional reason encouraging interest in the younger end of the 5-8 age range was the wish to study children during a year in which their exposure to standardised school representations increases most significantly. After the first years of school, children already will have experienced a great deal of systematic encouragement to use particular standardised representations of number such as illustrated 0-30 number lines. For these reasons, children in Year 1 (aged 5-6) were chosen as the population of most interest for a year-long study.

4.3.1.1 Sample

The sample invited to participate in the research consisted of 15 children selected from one Year 1 class at a local South of England primary school. Each participating child formed one case of the multiple case study. The sample was chosen with the expectation that not all participating children would remain in the study until the end.

Selection for inclusion in the study was carried out in consultation with the children's class teacher, after the researcher had spent one month in classroom mathematics lessons with the children. The criteria for selection were:

- children likely to want to participate, so in particular, no children with strong anxiety attached to doing mathematics
- an approximately equal number of boys and girls (8 boys and 7 girls invited to participate)

Within the above restrictions, a stratified sample was then taken based on the four teacher-assessed attainment groups within the class. In this way, the sample was constructed to be as far as possible representative of the class and the differing mathematical attainment levels. The sample is not fully representative due to the ethically necessary exclusion of children with mathematics anxiety or unwillingness, and the self-selection of actual participants from within the invited group. Since thirteen out of fifteen invited children did participate, the latter was not considered a significant concern. The research design does not involve a sample intended to be statistically representative of any single cohort. That said, the cases are intended and expected to be a fair representation of the typically-performing children within this particular school.

Of the fifteen children and their parents invited to participate in the research, parental consent for participation was received for thirteen children (6 boys, 7 girls), and all thirteen of these children themselves consented to take part in the study.

4.3.2 Ethical considerations

The significant ethical considerations of the study relate to researching with young children, and particularly video-recording them. To address this, the research was designed with the input of a classroom teacher to consider the impact on children, and both children and parents were asked permission. Parents were asked in writing, with the researcher available in person at the beginning and end of the school day to answer any queries and meet parents. The purpose and methods of the study were fully explained, in particular the process for ensuring the confidentiality of video data. No reservations or concerns were expressed by parents.

Children were asked for permission for each interview, and reminded that they were free to stop participation at any point. This was considered particularly important in order to ensure children did not confuse the interviews with the compulsory nature of the normal school day. Children were invited to view the recorded footage of their interviews after each interview; the majority were keen to see a few minutes of each recording. To protect children's anonymity, pseudonyms were used in all research reporting.

4.3.3 Data collection

Each child participating in the research was interviewed five times, once in each half term excluding the first half term of the school year. The intervals between interviews were therefore between six and eight weeks. This interval was chosen for several reasons: it was expected that children would make significant progress in mathematics between interviews, and have sufficient time to forget the precise details of the interview. Many practical considerations also made half-termly interviews a good choice; they fit well with the rhythm of the school year, avoided clashes with school holidays, and minimised disruption to class routines. Each interview lasted between 25 and 40 minutes, with children free to take as much time as they wanted for tasks.

Interviews were carried out by the researcher with individual children in an open-plan quiet study area of the children's school. Each interview involved the completion of four specifically designed tasks, and interviews were video-recorded for later analysis. Children were aware of the video recording, and invited to look at and test the camera themselves before interviews, and view the footage of their own participation afterwards. It was found during the pilot study that children were neither concerned nor overly interested in the video camera.

The tasks developed for the task-based interviews were:

- T1. "Imagine the numbers 1-100 …" task, adapted from Thomas et al. (2002).
- T2. Number-to-position estimation task, adapted from Thompson and Opfer (2010).
- T3. Estimation of quantities (Clements and Sarama, 2009, p. 53; Siegler et al., 2009).
- T4. Position-to-number estimation task, from Clements and Sarama (2009, p. 54).

The interview tasks were carried out in this same order for each child, in each interview. The drawing task, T1, was carried out first in order to capture children's cognitive representations before the potentially strong influence of the number line estimation tasks. The number placing task (T2) was carried out next, being the longest task in duration and so placed when children were likely to be feeling mentally 'fresh'. The quantity estimation task (T3) followed this, for children to refresh themselves with a break from the number line estimations. In the pre-pilot testing of Task 3, children did not seem to find the task taxing or tiring, and as a break from paper-based work this task seemed to function well. The final task was T4, in which children estimated the number represented by positions.

During each task, the researcher provided encouragement, and thanked children for any explanations and demonstrations given, but avoided giving feedback on the 'correctness' of any responses.

4.3.3.1 T1 – Imagine the numbers 1-100 …

This task invited children to give an external drawing and/or verbal description of their imagistic representations of the natural numbers from 1 to 100. The imagistic representation of number may occur spontaneously, and was expected to occur during the estimation tasks detailed above. However, the imagistic task used by Thomas et al. (2002) was incorporated into the interview in order to provide a deliberate opportunity for this kind of representation. The task has been successfully used by other researchers to elicit imagistic representations of number from children, and its inclusion allows a comparison of findings between Thomas et al. (2002) and this thesis.

The researcher introduced the task to each child in the following way: "This task is a bit different to normal maths, because there is no right or wrong answer. It's about the different ways people imagine or think about numbers. I want you to close your eyes … and imagine the counting numbers, from 1, 2, 3 … up to 10, and all the way up to 100 if you can. I want you to think about the way the numbers look in your mind, what picture you see when you think about those numbers, 1 up to 100. When you're ready, I want you to try to draw the picture that's in your head on this paper, and tell me about it, if you can."

Several questions and comments arose from participating children during classroom trials:

- "Can I use the colours?"
- "I don't know what to do."
- "Can I draw them in rows?"
- "Can I draw 200?"
- "Is this right?"

Responses to these questions and comments were therefore planned before the pilot study. The planned responses were as follows:

- Offer supportive encouragement whilst emphasising that there is "no right or wrong answer". The researcher can repeat that "this task is a bit different to usual activities in maths, isn't it?" Children in classroom trials were also interested to hear that a researcher in Victorian times (Francis Galton) used to ask his friends this question, so this can be used as encouragement.
- Children may use any of the drawing materials provided, and draw "whatever picture of the numbers you see in your mind".
- If children are hesitant or seem self-conscious, the researcher can also draw (with the image hidden from the child) so that the child feels more at ease.
- The researcher should repeat that "there is no right or wrong answer".

Children were invited to draw on plain A4 paper. The impact of this, and how it might constrain responses, was considered; alternative possibilities included using larger paper, offering children a 'blank wall', or asking them to interact with numbers in a physical space larger than themselves. Whilst A4 page boundaries might constrain a child's drawing, it is a paper size they are used to encountering and scaling their ideas to. Furthermore, its 'ordinariness' within school gives children confidence – it is not a precious 'special' art class resource, but one to be used however they wish to achieve their ends. It was felt that choosing something other than everyday class resources would have more impact on task responses, because children would notice it more, and feel pressure to invent something more 'special' than normal.

4.3.3.2 T2 – landing 'number rockets' (number-to-position estimation)

This task asked children to estimate the position of given target numbers on a blank number line, and stick the 'number rocket' in that position. As in number-to-position estimation tasks in the literature, the number lines presented were 25cm long, in the centre of a plain white sheet of A4 paper (e.g. Berteletti et al., 2010, p. 546; Siegler and Opfer, 2003, p. 238). The task differed from examples in the literature by presenting children with the target number on a sticker. Berteletti et al. (2010, p. 546) showed the numbers to be estimated "in the upper left corner of the sheet", and in the task by Siegler and Opfer (2003, p. 238) "the number to be estimated appeared 2 cm above the center of the line". The decision to present the target number on a sticker instead was taken for three important reasons. Firstly, it was felt that if the target number appeared *anywhere* on the number line page, it would be possible for its position relative to the line to influence children's estimated position for the number, however subliminally, and however small the effect. Secondly, handing children the target number on a sticker emphasised the physical and spatial act of taking a number in verbal and symbolic form, and assigning it to a position on the line, and emphasised that the child was in charge of this act. Thirdly, children were pleased by the context (rockets) and novelty of the materials, showing high levels of engagement when the task was trialled.

The task was introduced to children in the following way: “This task is about landing number rockets on number lines. Have a look at this number line – what’s different about it, compared to the number lines in your classroom?” If children did not comment that “it is missing all the numbers in the middle” or “it only has numbers at the ends” or similar, the researcher pointed to the endpoints and said: “Look, there are numbers here [point] and here [point], but we can’t see any of the numbers in between”. The researcher then explained: “I’m going to give you a number rocket, like this one [points], and I want you to put it on the number line where you think it should land. Don’t worry if you’re not certain, I just want you to have a think and put it in the most sensible place you can. It’s about *estimating* where the rocket should go – making a sensible maths guess.” The term *estimate* was included since it is part of the mathematics vocabulary taught during Year 1 and the children were beginning to apply it in mathematics lessons.

Children landed 37 number rockets in total, split over four different number lines. The four blank number lines used were marked with one of the following sets of endpoints: 0-10, 0-20, 5-15, or 0-100. The number lines consisted of 25-cm long lines in the centre of white, landscape, A4 sheets (see Appendix 3), as used by Berteletti et al. (2010). The target numbers to be positioned for each range were:

Range	**Target numbers**
0-10	1, 2, 3, 4, 5, 6, 7, 8, 9
0-20	2, 4, 6, 7, 9, 11, 15, 16, 18, 19
5-15	6, 7, 9, 10, 11, 13, 14
0-100	2, 3, 4, 6, 18, 25, 49, 50, 67, 71, 92
Total trials	**37**

Figure 4 Target numbers by range

Unlike Berteletti et al. (2010), on the range 0-10, every possible target number was tested, in line with the recommendations of White and Szucs (2012, p. 11). Saturation of target numbers would also have been desirable on the range 0-20, but as part of efforts to keep the overall interview length down only ten were selected. The selection included targets around each endpoint and the midpoint, as these were hypothesised to be points of particular interest.

The 0-100 target numbers were based on those used by Berteletti et al. (2010, p. 546), in turn following Siegler and Opfer (2003). Siegler and Opfer chose these target numbers in order to "maximize discriminability of logarithmic and linear functions and to minimize the influence of specific knowledge, such as that 50 is halfway between 0 and 100" (Siegler and Opfer, 2003, p. 238). Since this research did not want to minimise the influence of specific knowledge, target numbers close to 50 and 100 were added.

The range 5-15 was not tested by previous research. It was chosen for this research as an exploratory part of data collection, to investigate what would occur when endpoints were varied from zero and multiples of ten. In particular, it sought to further explore children's use of endpoint and midpoints to carry out their estimations. The range 5-15 was chosen for this non-standard range since it falls within the 0-20 range with which Year 1 children are familiar, and the number sequence *5 – 10 (midpoint) – 15* is one encountered during classroom practice counting in fives.

The order of the target numbers within each range was separately randomised for each child, again following Berteletti et al. (2010). The order of the ranges however was the same for each interview: 0-10, 0-20, 5-15, 0-100. A booklet was prepared for each child, with pre-prepared blank number lines in the correct order. Each child was then handed number rocket stickers by the researcher, one rocket at a time, to 'land'.

At the start of the task, the researcher completed a trial with the child to check task comprehension and build confidence. The researcher took a rocket labelled with an endpoint (e.g. 10) and asked "Where does this number rocket belong?" If the child did not answer by pointing to the correct endpoint, the task was re-explained. In pre-pilot testing of the task, all children confidently pointed to the correct endpoint. Most found the question funny, and made comments like "that's far too easy!" Rockets labelled with the endpoints were used so that the researcher avoided influencing children's estimations of rocket positions on the line itself (between endpoints).

4.3.3.3 T3 – estimation of quantities

This pure quantity estimation task was adapted from a task of Booth and Siegler (2006), in which children aged 5 to 9 years estimated the number of sweets in containers. In each trial, children chose between two possible answers: the correct answer and a different answer, either 0.5, 1.5 or 2 times the actual quantity. The quantities used by Booth and Siegler (2006) were 22, 34, 46 and 58. To gain additional data on estimations of small quantities, target number 14 was added for this thesis.

Sugar-coated chocolate sweets were used (diameter approx. 10mm), and an identical transparent container was used for each target number. The order of quantities and the choice of incorrect answer option were randomly chosen for each child. For each target number, the researcher asked “Have a look at this container. Do you think there are *n* or *m* sweets?” In case children recalled the answer options from the previous interview, alternate interviews used a slightly different set of target numbers: 18, 24, 32, 48, and 60.

4.3.3.4 T4 – naming ‘number rockets’ (position-to-number estimation)

In this task, children were shown an un-numbered rocket already positioned on a blank number line, and asked to estimate which number rocket they thought it was. The task was explained to each child in the following words: “Here are some number lines, like the ones you saw before. You can see I’ve already stuck a number rocket on, but I’ve covered up the number. Which number do you think this rocket should be?” The number lines and target numbers used in this task were the same as in Task 2, with the target numbers once again in a random order.

4.4 Theory for analysis

As the theoretical framework and methodology chapters have demonstrated, this thesis is primarily framed and supported by Duval’s theory of cognitive representations. This theory clarified the emerging research problem by encompassing within one framework children’s disparate cognitive representations of number. The framework explains the centrality of the representations to mathematical thought, explains why connections between representations can and cannot be expected, explains the significance of children’s conversions between representations, and explains why we can expect children’s conversions between representations of number to change.

What Duval’s theory does *not* provide for this thesis is the means to analyse the detailed features of representations and changes pinpointed by the research questions. This theoretical material, the theory for analysis, is presented now. The theory for analysis consists of three parts: firstly, imagistic representations; secondly, number structure represented by children; and thirdly, hypotheses in the literature with which the thesis engages.

4.4.1 Imagistic representation

Imagistic representations (representations with visual, spatial or graphical characters, as defined in the Theoretical Framework chapter) of number from any task, using any mode, are analysed using a framework adapted from Thomas and Mulligan (1995) and Thomas et al. (2002). Their framework uses Goldin’s model of internal representations with a model of imagery combining ideas from Presmeg (1986, 1998), Mason (1992) and again Goldin and Kaput (1996).

The basic framework first distinguishes three types of sign or character which constitute the fundamental elements of the external representation. The three categories are:

1. Pictorial: “pictures of objects, possibly together with oral descriptions of objects” (Thomas et al., 2002, p. 121). Pictorial elements in internal representations can also be inferred from verbal descriptions alone, for example if a child describes a collection of ten objects on the table.
2. Iconic: “drawings of tally marks, squares, circles, or dots” (Thomas et al., 2002, p. 121).
3. Notational: “the predominant use of numerals drawn in various formations such as number line, array, a 100cm ruler, or a vertical column” (Thomas et al., 2002, p. 121).

These three types of component may be combined in one drawing. The definition of the Pictorial category (above) contrasts with the Concrete/Pictorial category of Thomas and Mulligan (1995), which explicitly only includes “objects which do not have any quantitative relationship to the numbers” (p. 12). Correspondingly, the definition of the Iconic category also differs; in the 1995 paper, it is defined as “pictorial imagery which relates to a quantity” (p. 12) whereas the 2002 definition (used in the framework adopted for this thesis) specifies examples in which the elements of the image are to an extent abstracted; they play a role of signing the quantity as opposed to being depictions of concrete instances.

In general, it is reasonable to assume that the later paper (Thomas et al., 2002) reflects the authors’ deepened knowledge and insight into the research area. Additionally, the later definitions are preferred for use in this thesis since they provide greater clarity on an important distinction. The phrase “relates to a quantity” (Thomas and Mulligan, 1995, p. 12) is problematic; it does not specify the nature of the relation, and as such could include imagery better described as concrete pictorial, for example the depiction of a younger sibling “because she is two years old”.

The imagistic representation framework next examines the type of structure evident in the representation. The classifications developed by Thomas and Mulligan (1995, p. 12) are:

1. No structure: elements show no apparent relationship to equal groupings or sequence.
2. Linear structure: elements in linear formation (straight or curved), numbers in sequence.
3. Emerging structure: “one hundred represented by equal groups of objects, or linear sequence broken into equal segments” (Thomas and Mulligan, 1995, p. 12).

4. Emerging structure (m): representation shows some aspect of multiplication, such as multiple count and multiplication grid.
5. Partial array structure: elements are in rows and columns, but not a ten-by-ten array.
6. Array structure: elements in a ten-by-ten pattern.

Finally, imagistic representations are classified according to what Thomas and Mulligan (1995) refer to as "Nature of the Image" (p. 12). This categorisation refers to the movement in the imagistic representation, with the two classifications being:

1. Static: the representation is presented or described as fixed.
2. Dynamic: the representation involves changing or moving elements.

4.4.2 Representation of number structure

The above framework includes analysis of number structure, but the framework was specifically developed for the particular drawing task used by Thomas and Mulligan (1995). This section now presents theoretical material for the analysis of number structure, more generally, in children's cognitive representations across *all* tasks.

The principal theory used for this analysis is a microstage theory of number development in the early years of school, from Resnick (1983). It is congruent with the Piagetian account of number development, and consequently with the theoretical positions of Nunes and Bryant (2009) and other researchers favouring the Piagetian account. Resnick (1983) notes the key parallels with Piaget's account as "(a) emphasis on part-whole (class inclusion, for Piaget) relationships as a defining characteristic of number understanding, and (b) the proposal that ordinal (counting) and cardinal (class inclusion or part-whole) relationships must be combined in the course of constructing the concept of number" (pp. 146-147).

The convergence with the Piagetian account as "especially pleasing" since Resnick's analysis was carried out "quite independently of Piaget's work". The analysis set out neither to support or criticise Piaget's number understanding account; its aim was rather "to build a plausible account, from a current cognitive science point of view, or what number knowledge must underlie the various arithmetic performances observed in young children" (Resnick, 1983, p. 147). Resnick describes the methods used as "more bottom-up than those of Piaget" (p. 147), proceeding *from* children's behaviour in various tasks *to* the necessary number understanding.

The advantage that Resnick's theoretical account offers over others is the level of detail – the smaller 'grain size'. As a result of the 'bottom-up' methods, "we are able to detect – indeed, are forced to recognize – relatively small changes" (Resnick, 1983, p. 147). This leads to a microstage theory, “a theory that specifies many small changes in number representation and schematic interpretation of number in a period of development for which the Piagetian analysis recognised only the *macrostages* of preoperativity and concrete operativity” (p. 147).

4.4.2.1 Aspects of the number structure

The first aspect of number structure expected in children’s cognitive representations of number is **sequence structure**. Sequence structure is expected because it is incorporated into children’s number knowledge from the very earliest stages; when the numbers are known only as a string of words the sequence structure is already present (Fuson and Hall, 1983, p. 94). Resnick (1983) describes the number sequence as “a string, with the individual positions linked by a ‘successor’ or ‘next’ relationship and a directional marker on the string specifying that later positions on the string are larger” (p. 111).

An aspect of number structure identified as very difficult for children in the early stages of learning is **relative numerosity** (Fuson and Hall, 1983). Children in Year 1, like younger children, still demonstrate signs of difficulty with this concept. The words to describe relations are particularly difficult, with children confusing both the dimension of comparison (bigger, more, longer) and the direction of the relation (more than, less than) (Fuson and Hall, 1983). Determining the order relation on two cardinal words (e.g. the target number given and one endpoint) is “a very complex issue” (p. 75).

Children in the early years of school have been observed to use the sequence structure to solve relation questions on both number words and cardinal number contexts, for instance running through the sequence words for pairs of numbers to decide which is larger (Fuson and Hall, 1983). Children are able to use the order relation on sequence words to determine the order relation on two quantities. Fuson and Hall report research with children aged 4-6 years indicating that children appear to determine order relations on sequence words in the same way for all numbers up to twenty. Furthermore, “this same sequence process is used for the cardinal relations for words between ten and twenty” whereas for cardinal relations below ten, a different process – perhaps magnitude comparison – seems to be used (p. 98).

Even once the number sequence or ‘string’ is learned, children continue for some time to have difficulty beginning **sequences from numbers other than one**. In Resnick’s (1983) terms, this indicates “that the individual successor links are not fully established for some part of the string” (p. 112). A further specific stage of structural development is **decreasing number sequences**, involving the addition of ‘back’ markers to the number string, corresponding to the ‘successor’ or ‘next’ markers originally indicating the increasing sequence structure (Resnick, 1983). Fuson et al.

emphasise that these stages of development may take several years, and that it is to be expected that different ranges of the number sequence will be in different phases of development; typically, relations between numbers at the beginning of the sequence may be established whilst later parts of the number sequence are still in a very basic phase of development (Fuson, Richards and Briars, 1982).

The above sequence knowledge of number words "becomes a representational tool that is used for solving operations (addition, subtraction, multiplication, division) in cardinal contexts" (Fuson and Hall, 1983, p. 98). From it, skills such as the **counting-on** and **counting-back** procedures develop. Fuson and Hall suggest that the use of counting-on and counting-back helps children to see that addition and subtraction are inverse operations, allowing children in time to start choosing procedures for convenience or efficiency rather than because they directly model the problem context (p. 99). Resnick (1983), in turn, holds that counting-on and counting-back procedures "produce a quantitative interpretation" of the part-whole schema, which children appear to possess in primitive form before schooling (p. 146).

Resnick notes that so long as the number sequence remains the only number structure, "no precision" is possible in determining the relation of two numbers. The only way to compare the relative size of two quantities is "as a specification of the number of numerlogs that must be traversed between positions in the line" (Resnick, 1983, p. 114). The progression to interpreting numbers as compositions of other numbers, in terms of part and whole relations, is "the major conceptual achievement of the early school years" (p. 114). Children beginning to understand the compositional structure of numbers start to "**partition and recombine quantities** with some flexibility" (p. 122, emphasis added). Solutions to problems that **incorporate knowledge of number bonds to ten** are a particular application, one that signals the early stages of appreciation of the base ten number system (p. 121).

The special part-whole understanding that is knowledge of base ten number structure is a highly significant stage of development. With each number represented in terms of composition of tens and units, "in effect, that two-digit numbers are interpreted in terms of the Part-Whole schema, with the special restriction that one of the parts be a multiple of 10" (Resnick, 1983, p. 127). There exists some evidence of the base ten compositional structure of numbers early on in children's learning. Resnick cites evidence from Fuson et al. (1982) and Siegler and Robinson (1982), who found that "many 4- and 5-year-olds could count orally well into the decades above 20" and that their counting showed signs of being organised around decades: counting typically stopped at numbers ending in 9 or 0, and omissions and repetitions tended to be of entire decades (Resnick, 1983, pp. 127-128).

Whilst this primitive sense of part-whole structure exists before schooling starts, the early years of school are about "its systematic application to *quantity*" (Resnick, 1983, p. 146). Attaching the part-whole schema to counting-on and counting-back enables "a quantitative interpretation of Part-Whole" as noted, and the part-whole schema "in turn allows numbers to be interpreted *both* as positions on the mental number line *and, simultaneously,* as compositions of other numbers" (p. 146).

The basic procedures of counting on and counting back next develop to represent compositional and multiplicative structure by **counting-on and counting-back by tens** as well as by ones (Fuson et al., 1982). The stage developing from this is 'skip counting', in which children **count-on or count-back by any given number** appropriate to the problem context. In the context of number line estimations, employing the number sequence "five-ten-fifteen" for the number line 5-15 is an example of this appropriate skip counting.

These latter multiplicative stages are not trivial for children. Fuson and Hall (1983) explicitly link the difficulty of understanding the base ten structure of number with difficulty understanding the concept of numbers in *measure* contexts generally. The base ten system of numeration is a measure system, but comprehension of this is "very difficult" for children, and "At least through second grade [ages 7-8] and often later, words up to one hundred seem to elicit primarily counting, sequence, or cardinal meanings, rather than base ten measure meanings" (p. 85).

4.4.3 Specific predictions

In addition to the above theoretical material, specific predictions made by the literature influenced the research questions, data collection choices and data analysis of the thesis. These predictions do not form part of the theoretical framework, but are ideas with which the data analysis is designed to engage.

The first and main such hypothesis from the literature is the log-linear hypothesis. As discussed in the literature review, this hypothesis holds that both children and adults, at least in societies with highly developed writing and counting traditions, typically mentally represent number on a mental number line. In countries where text is written left to right, the mental number line is ordered left to right, and children's positioning of the numbers on this mental number line changes with age and development. Younger children represent numbers on the line with logarithmic placement, whilst older children and adults 'shift' towards linearly distributed representations (Siegler et al., 2009). The 'shift' occurs first on mental representations of small number ranges (0-10, 0-20) (Berteletti et al., 2010) and at later ages for larger ranges (0-100, 0-

1000) (Siegler et al., 2009). In terms of specific predictions for this thesis, the log-linear hypothesis predicts that:

1. Children's early number line estimations are likely to be best fit by a logarithmic curve in the case of T2 (number to position) and an exponential curve in the case of T4 (position-to-number)
2. Children's estimations in both T2 and T4 may 'shift' to a linear representation in later interviews
3. The shift is likely to occur first for smaller, familiar ranges (i.e. 0-10), and last for larger, less familiar ranges (i.e. 0-100)
4. The shift is likely to occur quite suddenly

The next sets of hypotheses presented suggest specific and even observable reasons why children's estimations over a given range come to resemble certain statistical distributions. Ashcraft and Moore (2012), who concluded that their data overall supported Siegler and colleagues' log-linear shift account of magnitude representation, concluded in addition that children's number line estimation was influenced by both the underlying representation *and* specific aspects of arithmetic knowledge (p. 265), due to evidence for the gradual inclusion of reference or anchor points. Their predictions for children's number line estimations are the following:

1. There will "always" be high accuracy for estimates of targets near the left endpoint, then accurate estimates for targets close to the right endpoint appear next, and finally accurate estimates for targets close to the midpoint (p. 265).
2. Over time, the graph of estimation error against target number will therefore increasingly resemble an "M" shape.
3. In T4, an early strategy is likely to be to be counting up from the left endpoint to the blank target (p. 266).
4. A more sophisticated strategy, likely to appear after this, is count from whichever line endpoint is closest to the blank target.
5. The further a target is from an endpoint, the higher the estimation error is likely to be.

The hypothesis of Barth and Paladino (2011) is that children's number line estimations depend on neither logarithmic nor linear positioning of numbers on a mental number line. Instead, the distribution of number line estimation results follows from the fact that each number line estimation is a proportion judgement, since for target number 30 on the range 0-100, for example, a child "cannot simply estimate the numerical magnitude of '30' in isolation; rather, they must estimate the size of a part (the numerical magnitude of '30') relative to the size of the whole (the magnitude of ‘100’)” (Barth and Paladino, 2011, p. 126). The variables involved in such a proportion judgement mean that systematic over- and under-estimations can be expected, depending on the relative distance of target numbers to endpoints and whether or not children mentally subdivide the whole range (for example with a midpoint). The predictions of this account of number line estimation are:

- T2 estimates in early interviews will be best fit by an S-shaped power curve, with overestimation of targets near the left endpoint, followed by accuracy around the midpoint, and then underestimation for targets near the right endpoint.
- There will be increasing use of a midpoint anchor, and estimates will then be best fit by a two-cycle S-shaped curve, with the cycle of over-estimation followed by under-estimation occurring twice within the given range, with accuracy at the midpoint (as before) but also around the 25% and 75% points of the range.

More recent studies have expanded on the work of Barth and Paladino (2011), as described in the postscript to the literature review. The research of Slusser et al. (2013) and Rouder and Geary (2014) broadly supports Barth and Paladino’s hypotheses, and also puts forward additional specific hypotheses. These are:

- Children may initially only use the anchor at zero, meaning that their estimations are “conceptually unbounded” (Rouder and Geary, 2014, p. 2).
- All models of number line estimation “can be understood in terms of placements guided by one, two, or three anchor points” (p. 10). Additional anchor points (from one to two, and from two to three) can be expected to be included progressively, one at a time.
- Targets close to anchor points will be “placed with greater accuracy than those farthest from an anchor.” (p. 11).
- Additional anchors (at 25% and 75%) are not expected: “We suspect there may be a general limit of one virtual anchor due to working memory constraints.” (p. 11).

A closely related hypothesis that the data analysis of this thesis is concerned with is the suggestion of White and Szucs (2012) that children may display unique behaviour for each target number in a number line estimation task. The representation of the number and the estimation of the number’s position are hypothesised to be a function of “the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy” (p. 9). The data analysis of this thesis will look for evidence that children’s estimations vary in these specific ways. This hypothesis differs from others in explicit consideration of familiarity and arithmetic strategy.

Overlapping waves theory is the final theoretical hypothesis with which this thesis is concerned. Again, as discussed in the literature review, the theory holds that children are able to access or use more than one representation of numbers at any given stage of development (Siegler, 1996). Children's use of particular representations occurs with the pattern of 'overlapping waves', that is to say, the introduction of a new representation of number occurs whilst the old is still active, and children may use either. In the context of the current research, the aim is to examine whether interpretation through the overlapping waves theory provides a plausible and coherent explanation of the data collected.

4.5 Plan of analysis

This section details the plan of analysis used. It first explains the initial stages required, for example, transcription and identification of cognitive representations, and then demonstrates how the different stages are brought together to answer the research questions.

4.5.1 Transcription and identification of representations

The first stage of data analysis for this study is the transcription of the video-recorded interviews. The audio parts of the video-recordings are transcribed first, noting speech, exclamations, pauses or hesitations, and any distinguishing features of speech (e.g. particular loud/soft, high/low pitch). Next, eye gaze is transcribed: the child's point of focus (e.g. left endpoint of number line, or the interviewer) is recorded, along with the timings of each transfer of gaze.

In common with previous studies, gesture is transcribed and coded separately from speech (Alibali et al., 1999; Garber et al., 1998). Gesture is analysed like other imagistic representations: for the nature of its component signs (pictorial or iconic – notational not likely in gestural representations), and any number structure represented. However, before this stage it is necessary to use gesture-specific codes to identify the fundamental components of the gesture and transcribe it.

A common route in the transcription of gesture is to use speech to identify the units of analysis (Garber and Goldin-Meadow, 2002), but this was not chosen for the present research given the importance of co-thought as well as co-speech gesture. Instead, units of gesture are identified using McNeill's definitions of gesture phrase (G-phrase) and gesture unit (G-unit) (McNeill, 1992). A *gesture phrase* consists of one or more phrases of movement beginning with *preparation*, various holds and strokes, and ending with *retraction*. A *gesture unit* is composed of one or more gesture phrases, and is defined as the period of time between successive rests of the limbs. A G-unit begins the moment the limb begins to move and ends when it has reached a rest position again (McNeill, 1992).

The basic elements of gestures in this thesis are identified using the iterative process set out by Perry et al. (1988). In the first stage, gestures are analysed at a fine level of detail and without interpretation, according to:

1. their form (e.g. pointing finger, grabbing motion, hand sweep left to right)
2. their placement (e.g. in 'neutral space' in front of the gesturer's face, or on paper)

A list of individual codes within these dimensions is found by inspection of the video data, with codes added until all basic forms and placements witnessed in the data had been exhausted.

After the transcription of speech, gaze and gesture, the transcripts, along with children's inscriptions, are examined for representations of number. Particular attention is paid to moments at which cognitive representations of number are most expected: during task explanations by the interviewer, during trials of a task, and after a question from the interviewer. Duval's theoretical framework identifies these as moments when children are likely to be performing one or more conversions between, say, representation of number in natural language and representation of number using another register.

4.5.2 Analysis of imagistic representations

Thomas et al. (2002) analyse the external representations produced by children with respect to three dimensions. The first stage is the categorisation of the components or sub-units of imagistic representations, as pictorial, iconic or symbolic, all of which Presmeg includes within visual imagery. Secondly the level of structural development (following Goldin) is identified, and thirdly the representation is examined for evidence of involving primarily dynamic or static imagery. Children's verbal explanations are used to help code the imagistic representations. When a representation has aspects of more than one representation type – e.g. a representation whose component signs are (discursive) algebraic symbols, arranged in an idiosyncratic spatial layout with the child's own scheme of colours – the representation is included within the analysis of imagistic representations, whilst not excluding it from further analysis as a discursive representation.

4.5.3 Strategy

Children's strategy cannot be directly observed; what can be observed is behaviour, and from this inferences about task strategy are made. In gesture analysis studies, a common method following Perry et al. (1988) for inferring the 'meaning' of gestures, after transcription, is to consider speech and gesture together in order to generate a task-specific *lexicon* of gestures, for example referring to the procedures within a task used by a participant. In this thesis, the different modes considered are brought together for the purposes of inferring the *strategy* a child appeared to use in a given estimation trial. An example is the gesture consisting of 'hopping motion with pointed finger [gesture form], left-to-right from left endpoint along blank number line on paper [gesture placement]'. In the pilot study, this gesture was frequently accompanied by whispered or mouthed numbers: "one-two-three-four-…". In such cases, this behaviour was interpreted as the child using a counting-on strategy to estimate the position on the number line.

4.5.4 Quantitative analysis of T2, T3 and T4

Several statistical analyses are carried out on children's responses to T2, T3 and T4 in order to address the research questions concerned with quantitative hypotheses in the literature, and to cross-reference with the qualitative analysis of children's task responses.

In order to work with children's estimations from Task 2, children's *target number estimates* are calculated for each rocket placed on the line (Booth and Siegler, 2006; Laski and Siegler, 2007; Siegler and Booth, 2004; Siegler and Opfer, 2003; White and Szucs, 2012). For each trial of each child, the *target number estimate* is the number that is 'hit' in each trial assuming a linear scale to the blank line:

$$\frac{\text{Distance from left endpoint of line to rocket (mm) x numerical range of number line}}{\text{Total length of line (mm)}}$$

For example, a rocket positioned 10mm from the left endpoint of a 250mm line marked with the range "0-100" equates to the target number estimate (10 x 100) ÷ 250, which is 4.

For all estimation tasks, the percentage absolute error (PAE) is then calculated, again following methods of analysis in the literature (Berteletti et al., 2010; Siegler and Booth, 2004). The following equation is used to calculate PAE:

$$\frac{|\text{Estimate} - \text{Target Number}| \times 100}{\text{Scale of number line}}$$

Descriptive statistics are then calculated to provide overviews of children's responses on each task and each range, at each interview round. The PAE is commonly described as a measure of accuracy, but it should be noted that it measures the accuracy of linear estimation only.

For the purposes of investigating the log-linear hypothesis, analysis again follows the established methods in the literature. In order to describe the best fitting model for estimates at group level, logarithmic and linear models are fitted to the group's median estimates (Berteletti et al., 2010; Siegler and Booth, 2004; Siegler and Opfer, 2003; White and Szucs, 2012). The best fitting curves for each model type are then compared by calculating the residuals of the median estimates to each model, and comparing the R^2 (variance explained) figure for each model. The significances of differences in model fit are tested using paired-sample t-tests comparing the residuals of each target number to each model (Berteletti et al., 2010). To investigate changes over time in the linearity of estimations at group level, one-way ANOVA tests on the PAE of estimates for particular ranges are carried out (Booth and Siegler, 2006).

The log-linear hypothesis, as previously discussed, has primarily used number-to-position estimation tasks for research. Task 4 however is a position-to-number estimation task, also used but less frequently. For number-to-position tasks, the models of theoretical interest are linear and exponential since "Consistent reliance on a logarithmic-ruler representation implies that mean estimates should increase logarithmically with numerical magnitude on the NP task and exponentially with numerical magnitude on the PN task" (Siegler and Opfer, 2003, p. 238). Consequently, the analysis will compare linear and exponential models rather than linear and logarithmic for Task 4.

To examine whether the group level analysis reflects model fit for individuals' estimates, a more interesting analysis for the purposes of this thesis, individual children's estimations on each range are also statistically compared to possible underlying models by curve fitting (e.g. model fitted to estimations for "Child A, Task 3, Round 1, Range n_1-n_2"). Following the method of Siegler and Opfer (2003) and Berteletti et al. (2010), the best fitting model for each child on each range is decided by comparing R^2 figures. In cases where neither model reaches significance, the child is classified as 'no model' for that range, following Berteletti et al. (2010). The best fitting model for a particular child on a particular task and range can then be compared over time, between interview rounds, and compared to the cases of other participants. To test for changes in the proportion of children for whom the linear or logarithmic/exponential model best fits for a given range, chi-square tests are used (Siegler and Opfer, 2003).

At a group level, the variation of estimation for particular target numbers and particular ranges is compared by calculating the mean and standard deviation of PAE. To compare how particular numbers are estimated within different ranges (for example target numbers 2, 4, and 6 are tested on ranges 0-10, 0-20 and 0-100) the PAE can be compared, but also the residuals to the individual's best fitting linear and logarithmic/exponential (as appropriate) model. Comparison of residuals of individual target estimates of each participant is the method used by White and Szucs (2012), using a target number x time (interview round) x model (linear/logarithmic) ANOVA test.

The quantitative analysis of Task 3 carried out is principally calculating the percentage of correct estimations by individual children, and comparing the proportions of under- and over-estimation (Booth and Siegler, 2006).

The above-described stages of analysis are brought together in specific ways to answer the research questions. The following table shows how this is done:

. **Research question**	**Relevant analysis**
1. **In what ways do children cognitively represent number during the different tasks of the interviews used in this research?**	
a. What are the modes and component signs used in the representations?	In each task, examine: speech, inscriptions on paper, gesture and body movement, and gaze. For each child, note the modes with which cognitive representations of number occur. Using Presmeg's classification as developed by Thomas et al. (2002), identify the primary component sign or signs in each representation: pictorial, iconic, or symbolic.
b. What aspects of number structure are represented?	In each representation, ask which aspects of the structure of number are evident. E.g., sequential structure, relative proximity, midpoints, quarter points.
c. What are the notable between-task and within-task connections between representations?	Within each child's interviews, compare the representations demonstrated in each task on the basis of all variables examined in Q1 [mode, component sign, number structure]. • What is the same and what differs? • Do representations appear simultaneously using different modes? If representations occur sequentially, what is significant about the order? Compare the representations demonstrated within each task within each interview, on the basis of all variables examined in Q1. • What is the same and what differs? In particular, do characteristics occur consistently, or vary with a detectable pattern? • Do variations in task detail (i.e. target number, target order, and range) correspond to variations in representations detected? • Do representations appear simultaneously using different modes? • If sequentially, what is significant about the order? Do different representations appear to perform different task functions?

2. What strategies can be identified in children's interactions with number line estimation tasks?	
a. What patterns can be detected in the way children use or do not use these strategies?	Look for the occurrence of the different strategies and patterns that occur across the variables of round (time in year), task, target number, range, child in tasks 2 and 4. • For a given range, do individuals' strategy uses change? • Do particular combinations of target number and range witness particular strategies more often? E.g., does target 9 on range 0-10 see more children more often using a 'count back' strategy? • How does the order of target numbers affect use of strategies? E.g. using previous estimations to guide or check current task part.
b. How do the strategies used relate to children's estimation results?	At range-within-task level: compare estimated positions, especially linear accuracy, with the strategies used for that range in that task in that interview.
3. How do young children's cognitive representations of number change during their first year of formal schooling?	
a. In what ways does evidence support or not support the log-linear hypothesis?	Do (any) individuals, and on which ranges of which tasks, show the log-linear progression as hypothesised by Siegler et al. (2009)? i.e. early in the year, best fit by log model, becoming increasingly linear.
b. What is the intra-child variability of children's numerical magnitude representations in estimation tasks at different times?	Variation in residuals calculated for best fitting model is the measure of variability to compare with variation in the literature. Additionally, cross reference to variation in observed cognitive representations and strategies. Q 3b should begin this work. Link statistical variation figures to the strategies: does high variability occur when a particular strategy is used/not used, or within same strategy?
c. Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?	• Map the changes in cognitive representations, strategies and quantitative estimations for each child. • Look for patterns of change based on ranges [e.g. for child X, trajectory seen on multiple ranges, either simultaneously or at different times]. • Check for inter-child similarities in their changes over the year.

Figure 5 Research questions mapped to analysis

This chapter has set out the research design, along with the theoretical and methodological justifications for it. In the process of drawing up the research design, potential tasks were trialled with children and a small scale pilot study was carried out. The next chapter presents a description and findings of the pilot study, highlighting the alterations that were made due to its findings, and the results of interest that it presents for investigation in the main study.

Chapter 5 Pilot study

In order to trial aspects of the task-based interviews, a small-scale pilot study was carried out prior to finalising the research design. This chapter discusses the pilot study design and findings, with a focus upon the ways in which the pilot study informed the main research design. The pilot study and its findings were presented in a research report at PME-37 in July 2013 (Williamson, 2013).

5.1 Setting and sample

Eight children from a Year 1 class at a local South of England infant school participated in the pilot study. The researcher had been observing and participating in classroom mathematics lessons for the previous three weeks, and individual tasks of the task-based interviews had been trialled with groups of children in the class. Children were chosen for participation in the pilot study based on the following criteria, as assessed by observation and teacher recommendation:

- willingness to participate
- had not already taken part in classroom trials of individual tasks
- equal numbers of boys and girls
- teacher-assessed to be middle or high attainment in mathematics, since it was not known precisely how challenging children would find the interviews

5.2 Study design

The pilot study consisted of small-scale case studies of eight children. The task-based interviews were largely structured as described in the Methodology chapter. Children completed the four tasks used in the main study, and the individual interviews were video-recorded for analysis.

Four tasks were completed, designed to stimulate and require translation of cognitive representations of number. The first task (T1) required children to close their eyes and imagine the numbers 1 to 100, then to draw and describe the picture in their mind. Following this, children completed an estimation task (T2) in which they were asked to position number rocket stickers onto blank number lines. A third task (T3) asked children to estimate the quantity of sweets in clear plastic boxes. Finally, children were asked to estimate the number represented by already-positioned rockets on blank number lines (T4).

Data analysis followed the plan set out in the Methodology chapter. The analyses possible on this small-scale, non-longitudinal pilot study were limited, but it was nevertheless interesting to begin exploratory work and get a sense of results that might be fruitfully followed up in the main study.

In addition to the planned data collection, observation notes were taken relating to children's reactions to tasks and materials in order to consider refinements needed before the main study.

This chapter reports and explains quantitative findings from all eight children's participation, and then presents one interview in detail in order to illustrate the qualitative analysis carried out, the conclusions drawn, and the implications for the main study.

5.3 Task 2: quantitative analysis

Task 2, the number-to-position estimation task, was presented as described in the Methodology chapter, with a few differences described here. The target numbers used during the pilot study were the following:

Range	Target numbers	Trials
0-10	2, 3, 4, 6, 7, 8, 9	7
0-100	2, 3, 4, 6, 18, 25, 48, 67, 71, 86	10
0-20	2, 4, 6, 7, 13, 15, 16, 18	8
5-15	6, 7, 8, 9, 11, 13, 14	7
Total		32 trials

Figure 6 Target numbers and trials by range

The 0-20 target numbers were taken from Berteletti et al. (2010), as were the target numbers for 0-10. The 0-100 target numbers were also taken from Berteletti et al. (2010), in turn following Siegler and Opfer (2003). The 5-15 number line was an original task, not seen in the literature, so these target numbers were chosen specifically for this study, using the same reasoning as other researchers' target number choices for the 0-10 range.

Unlike in the main study, the order of ranges tested was randomised for each child, following Berteletti et al. (2010), and the rocket stickers used were larger (see Appendix 5). The discussion explains why these aspects were changed in the final research design.

5.3.1 Modelling target number estimates

Firstly, children's median estimates were compared to possible linear and logarithmic models, with the following results:

- For the range 0-10, the fit of the linear model was significantly better than that of the logarithmic model (R^2_{lin}= 93%, $p<.001$ vs. R^2_{log}= 83%, $p<.01$, $t[6]=-3.10$, $p<.05$).
- For the range 0-20, the linear model provided a very good fit (R^2_{lin}= 99%, $p<.001$) but not significantly better than the best-fitting logarithmic curve (R^2_{log}= 98%, $p<.001$, $t[6]=-1.75$, $p>.05$).

- On the range 5-15, the linear model again provided a better fit, but not significantly better than the logarithmic model (R^2_{lin}= 95%, p<.001 vs. R^2_{log}= 90%, p<.01, t[7]=-0.84, p>.05).
- For the range 0-100, a logarithmic model provided a better fit, but the difference was only significant at 10% level (R^2_{log}= 97%, p<.001 vs. R^2_{lin}= 85%, p<.01, t[9]=2.05, p=.07).

These results are as expected for children aged 5-6; on the lower ranges, children's estimates are largely linear, whereas on the range 0-100 estimates are still better fit by a logarithmic model. These findings are in line with those of (Berteletti et al., 2010), whose research with 4-6 year olds on lower number ranges most closely matches the conditions of the pilot study.

The above modelling was carried out in order to get an overview of children's responses and allow comparability with previous key studies. The following table now shows the results of fitting linear and logarithmic models to the estimates of *individuals*. For each child on each range, the better-fitting model (with higher R^2 value) is highlighted green. In cases where no model achieved a significant fit, this is noted 'None'.

Range	Child	Code	R^2_{lin}	R^2_{log}
0-10	B	15	0.7694	0.6319
0-10	ED	11	0.9826	0.9555
0-10	EW	12	0.8792	0.7482
0-10	Imogen	14	0.9789	0.9176
0-10	JS	16	0.9016	0.8433
0-10	JB	17	0.8478	0.7392
0-10	O	13	0.9192	0.8582
0-10	T	18	0.6050	None sig.
5-15	B	25	0.9368	0.9507
5-15	ED	21	0.8858	0.9073
5-15	EW	22	0.8020	0.8528
5-15	Imogen	24	0.9852	0.9687
5-15	JS	26	0.6777	0.6592
5-15	JB	27	None sig	None sig
5-15	O	23	0.7743	0.8218
5-15	T	28	0.9900	0.9711
0-20	B	35	0.6533	0.6160
0-20	ED	31	0.9712	0.9476
0-20	EW	32	None sig	0.6521
0-20	Imogen	34	0.7157	0.7587
0-20	JS	36	0.9115	0.8715
0-20	JB	37	0.7908	0.6925
0-20	O	33	0.8046	0.8399

0-20	T	38	0.9134	0.7231
0-100	B	45	0.9289	0.9012
0-100	ED	41	0.8232	0.8896
0-100	EW	42	None sig.	0.6785
0-100	Imogen	44	None sig.	0.6872
0-100	JS	46	0.8233	0.7519
0-100	JB	47	None sig.	0.5602
0-100	O	43	0.6542	0.8833
0-100	T	48	0.8669	0.6558

Figure 7 Model fits by child and range

For all children, estimates on the 0-10 range were indeed better fit by linear models, supporting the group level result. For ranges 5-15 and 0-20, the picture is mixed, with children's estimates best fit by both kinds of model. For 0-100, the majority (5/8) of children's estimates were best fit by a logarithmic model, but the three exceptions help explain why the group's difference in R^2 figures was only just significant. These findings are again in line with those of Berteletti et al. (2010).

5.4 Task 3

In this task, children estimated five different quantities of sweets in a clear plastic jar. Each time, they were presented with two answer options: the correct quantity, and an incorrect option either half, one and a half, or twice the correct quantity. The main role of this task is to track children's ability to estimate quantities of different sizes, and to examine whether children's patterns of over- and under-estimation change over the school year. The analysis to be carried out after the non-longitudinal pilot study was, therefore, not extensive.

The task on estimation of quantities was adapted from the research in Booth and Siegler (2006). Unsurprisingly, Booth and Siegler found that kindergartners (aged 5-6, the same as the pilot study participants) were less accurate than older children, selecting the more accurate response on average 53% of trials (SE = .03, and chance score 50%, as in this task). The participants in this pilot study were more accurate, selecting the more accurate response on average 60% of the time (s.d. .174). Individuals' accuracy in one set of estimations (five trials) ranged from 20% to 80%.

Children were twice as likely to overestimate the quantity shown than to underestimate. The estimates were classified as follows:

Estimate	Percent
Underestimate	13.8
Perfect	60.0
Overestimate	26.3
Total	100.0

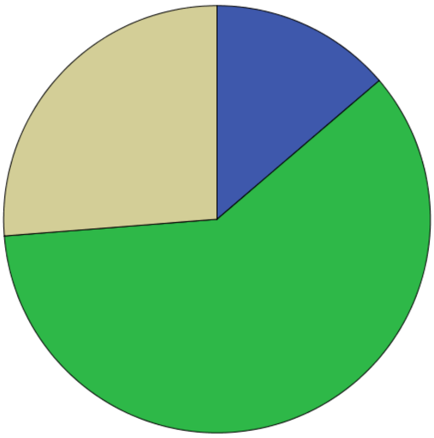


Figure 8 T3 estimation results

5.5 Task 4: quantitative analysis

5.5.1 Modelling target number estimates at group level

Once again, the median estimate of all the children was selected, and linear and exponential curves fitted (Siegler and Opfer, 2003; White and Szucs, 2012). Exponential curves were chosen because, as discussed in the Methodology chapter, if the logarithmic model is correct for children's positioning of numbers onto the number line, then the position-to-number task, as the inverse of the number-to-position task, ought to be best modelled by an exponential curve. The R^2 figures of the exponential model were in fact very low, so logarithmic curves were also fitted in order to compare logarithmic, linear and exponential models. The table showing the R^2 figures for the best fitting model of each type on each range is shown below. All models shown reached significance at $p<.05$ or better. The highest R^2 score is highlighted green for each range.

Range	R^2_{lin} (linear model)	R^2_{log} (logarithmic model)	R^2_{exp} (exponential model)
0-10	91%	96%	83%
5-15	74%	80%	48%
0-20	98%	95%	86%
0-100	95%	75%	95%

Figure 9 Table of model fits to median estimates

These results differ from those anticipated, since the exponential model is the best-fitting for none of the ranges, only equally the linear model in the case of the range 0-100. This clearly differs from Siegler and Opfer's (2003) findings with 7-8 year olds on position-to-number tasks, whose estimations were well fit by exponential models.

To examine whether the group median results reflect the models fitting individuals' results, linear, logarithmic and exponential curves were also fit to individuals' estimates. The following table shows the R^2 figures of the best fitting model in each case; once again, where no model reached significance, this is noted.

Range	Child	Code	R^2_{lin}	R^2_{log}	R^2_{exp}
0-10	EW	12	.615	.758	None sig.
0-10	B	15	.938	.983	.849
0-10	JS	16	.781	.838	.591
0-10	JB	17	.672	.650	.708
5-15	EW	22	.924	.918	.875
5-15	B	25	.974	.985	.924
5-15	JS	26	.895	.888	.862
5-15	JB	27	None sig.	None sig.	None sig.
0-20	ED	31	.976	.895	.941
0-20	O	33	.802	.849	.731
0-20	Imogen	34	.971	.924	.864
0-20	T	38	.989	.949	.884
0-100	ED	41	.951	.707	.946
0-100	O	43	.836	.574	.899
0-100	Imogen	44	.911	.743	.902
0-100	T	48	.973	.759	.912

Figure 10 Table of model fits, individuals

Supporting the picture given by the group level results, there are few cases in which an exponential curve was the best fitting model for children's estimates. In fact, the linear model provides the better fit in a majority of cases.

To better see differences between Task 2 and Task 4 estimations, a scatter plot was produced to show children's median estimates in both tasks for each range. For the range 5-15, this reveals the highly linear estimates in both tasks; differences are minor.

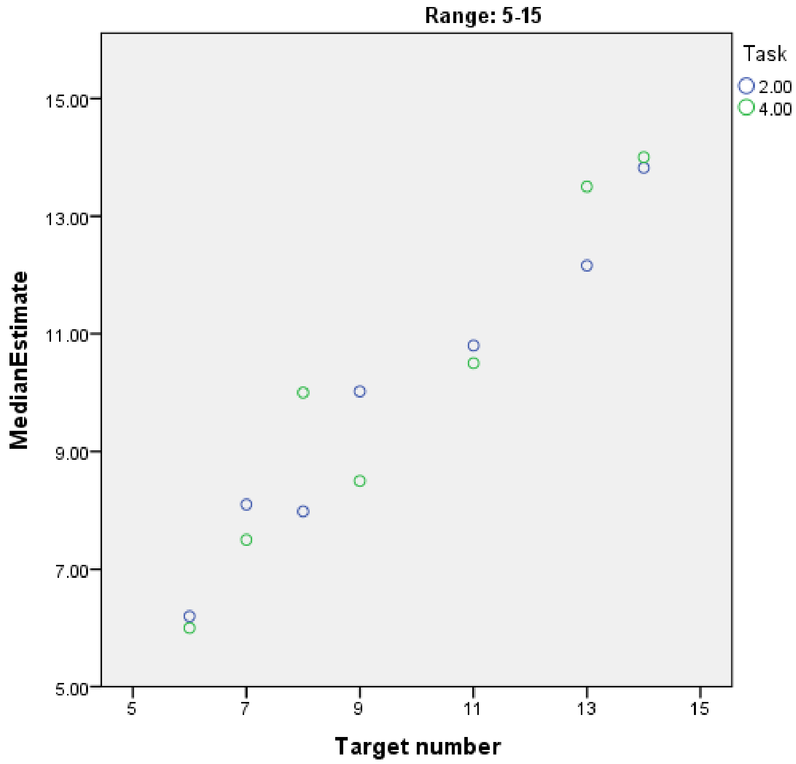


Figure 11 Median T2 and T4 estimates in range 5-15

For the range 0-100 (Figure 12, below), and to a certain extent 0-20, different patterns of estimation are clearly visible. Specifically, the 0-100 plot follows two s-curves, each a reflection of the other, and centred on the midpoint 50.

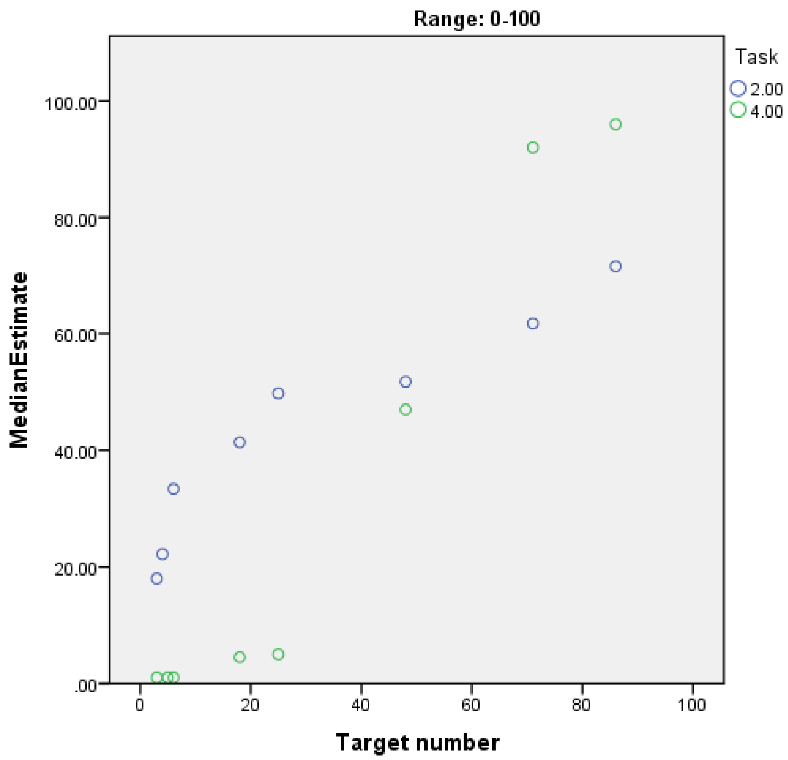


Figure 12 Median T2 and T4 estimates in range 0-100

This picture strongly suggests that the investigation of children's uses of midpoints is worthwhile, and could also suggest re-visiting the proportion judgement models of (Barth and Paladino, 2011). Though good arguments have been made against this type of model (Opfer et al., 2011; White and Szucs, 2012), it should not be ruled out that children's estimations do follow the proportion judgement model. However, it is also possible that the distribution of estimations shown above may be well explained by the observed estimation strategies children apply, for example the particular methods of counting on and counting back. The following section presents one interview analysed in detail, which supports this hypothesis.

5.6 Case study: Imogen

Imogen, whose case is discussed here, was assessed by the teacher as high-attaining in mathematics. Interviews were carried out at the end of the school year, so at the time of interviewing, Imogen was 6 years old.

5.6.1 Task 1

Imogen's initial response to T1 was to ask for clarification. During the exchange that followed, she gesticulated in reference to the number sequence:

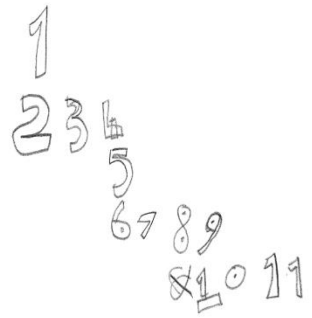
Interviewer: I'd like you to try to draw ... the picture you see in your imagination of all those counting numbers.

Imogen: So like <u>one two … and three four</u> [*right hand hovers over paper and traces stair-shaped path from left of page: right-down-right, twirling pencil*]

This same stair-shape appears in the drawing then produced by Imogen (Figure 13). The drawing is composed of notational signs with an idiosyncratic and pictorial aspect, evidenced in both the drawing and Imogen's unprompted explanations: "When I saw it all it was bubble writing". Though the interviewer did not comment or enquire, Imogen explained that the particular form was also the reason for drawing a limited range: "I'll just do it up to ten … Cos I don't want to waste all my time counting up to a hundred in bubble writing". The interviewer then pursued this:

Interviewer: If you did have time, where would one hundred go on that page?

Imogen: [*silently mouths the numbers one to ten, as right index finger jumps one by one along the number sequence already drawn*] <u>Twelve</u> [*finger jumps onto empty space to right of "11"*] … <u>thirteen</u> <u>fourteen</u> <u>fifteen</u> … [*finger jumps three steps to lower right, see* Figure 13 *right*] … I have no idea!



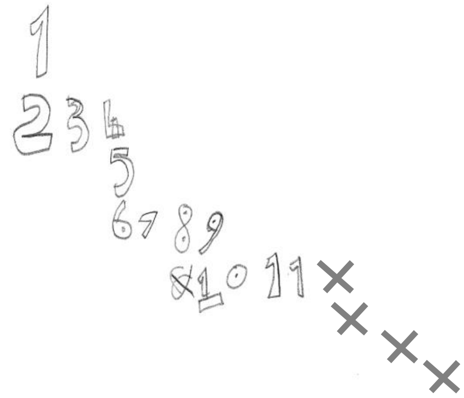


Figure 13 Imogen (T1) **Imogen (T1) with gesture**

The sequence structure of number is clearly represented in all modes, and the numbers as far as shown are evenly spaced. Numbers are also grouped, relatively consistently though unconventionally. Imogen’s comments and re-drawing indicate clearly that the “9” was originally intended to be positioned beneath the 8, consistent with her earlier grouping (changing direction on each multiple of two).

5.6.2 Tasks 2 and 4

Representation of structural elements of number occurred in all modes examined during these tasks: speech, gaze, and gesture. Representations occurred during trials both before and after giving an initial solution, and also in spontaneous comments and justifications. Clearly distinguishable strategies were identified in Imogen’s interactions with T2 and T4. The strategies, and the target numbers of the trials in which they were observed, were as follows:

Strategy	Task 2				Task 4	
	0-10	*5-15*	*0-20*	*0-100*	*0-20*	*0-100*
Counting on from L endpoint in 1's	4, 6		2		2	3
Counting on from L endpoint in 5's		9				
Counting on from midpoint in 1's		11, 13	6			
Counting on from other point		13				25
Counting back from midpoint in 1's			6			
Counting back from other point		11				
Referral to a previous trial		7, 14	7	48	15	18
Ref. to L endpoint	8, 2, 7, 3	6	4, 13, 18	2,48, 4, 67, 25, 6	16, 4, 7	71, 6, 86
Ref. to R endpoint	9, 8, 7, 3	13	16, 18	2, 48, 4	7	71, 86, 48
Ref. to midpoint				71, 67, 18, 3, 25		
Ref. to other point						
Ambiguous		8	15	86	6, 18, 13	67, 5, 3

Figure 14 Strategies identified in use in Task 2 and Task 4

In the following example from T2 (range 5-15, target number 13), the strategies identified were "Ref. to R endpoint", "Counting on from midpoint in 1's", and "Counting on from other point":

Interviewer: It's <u>thirteen</u>. Where do you think thirteen belongs? [*Imogen's gaze goes quickly to right endpoint (15) then to interviewer proffering rocket sticker*]

Imogen: [*takes rocket with right hand, transfers to left hand, pauses*]
'Cause <u>ten is here</u> [*right hand points onto midpoint and holds*]
[*right hand 'hops' to right; both hands stick rocket to right of the 'hop'*]
<u>Fourteen</u> <u>fifteen</u> [*right hand thumps line between rocket and right endpoint, then thumps right endpoint itself*]

Representation of aspects of number structure is apparent in these strategies. In the example above, the number sequence is represented in the two counting on strategies, and with a confidence that allows Imogen to start counting midway through the sequence. The units represented by gesture during the counting on represent a further aspect of structure: the spatial extent of each unit is approximately equally sized, and scaled so that Imogen's sequence from ten to fifteen covers the spatial extent from indicated midpoint to endpoint. Structure of number is also apparent in Imogen's use of the right endpoint (fifteen) as an appropriate 'landmark' for the target number thirteen. The midpoint structure of ten within the range 5-15 is clearly represented in speech and gesture.

Throughout Task 2 and Task 4, the sequence structure of number, with left to right orientation, was most frequently represented. Representations of number that Imogen spontaneously demonstrated on the ranges 5-15 and 0-20 encompassed further structure in the form of evenly spaced multiples of five. An example of this was Imogen's exclamation on seeing the first page of 5-15 trials: "Shouldn't it be five TEN …?" [*Right hand points onto midpoint of line and holds*].

As expected from previous research, Imogen's estimates were more linearly accurate on low ranges (0-10, mean PAE 4.17%; 5-15, mean PAE 2.86%) and higher on the range 0-100 (mean PAE 24.23% and 14.38% for T2 and T4 respectively). Interestingly, linear accuracy in Task 4 was higher than in Task 2 on both ranges tested; a paired samples t-test was conducted to compare the absolute percentage error found a significant difference between the error in Task 4 (mean=11.94, SD=10.00) and Task 2 (mean=21.59, SD=8.78); $t(16)=2.89$, $p<.05$.

In terms of model fit, on the range 0-10, Imogen's T2 estimates were best described by a linear model (R^2_{lin}=.980, compared to R^2_{log}=.918). On the range 5-15, the linear model again provided the best fit (R^2_{lin}=.985 compared to R^2_{log}=.969). On the ranges 0-20 and 0-100, Imogen's estimates were more consistent with a logarithmic model (0-20: R^2_{lin}=.716 and R^2_{log}=.759; 0-100: no linear model reached significance, and R^2_{log}=.687). Imogen's T4 estimates were, in contrast, better fit by linear models for both ranges 0-20 (R^2_{lin}=.971 compared to R^2_{exp} =.864 and R^2_{log}=.924) and 0-100 (R^2_{lin}=.911 compared to R^2_{exp} =.902 and R^2_{log}=.743).

The linear accuracy of Imogen's Task 2 estimates is in line with previous research, which expects that by the end of Year 1, estimates on the range 0-10 will demonstrate a good level of linearity, whilst those on larger number ranges do not. Overall, Imogen's linear accuracy was highest on the Task 2 trials on the range 5-15. Almost every strategy identified was in evidence during these trials; and, furthermore, during this part of the interview cognitive representations with more structural detail and greater accuracy were inferred from Imogen's spontaneous behaviour, for example with regard to visualising the subdivision of the 5-15 line into equally sized fives. The following scatter plot shows the linear accuracy of Imogen's estimates on the range 5-15 in Task 2:

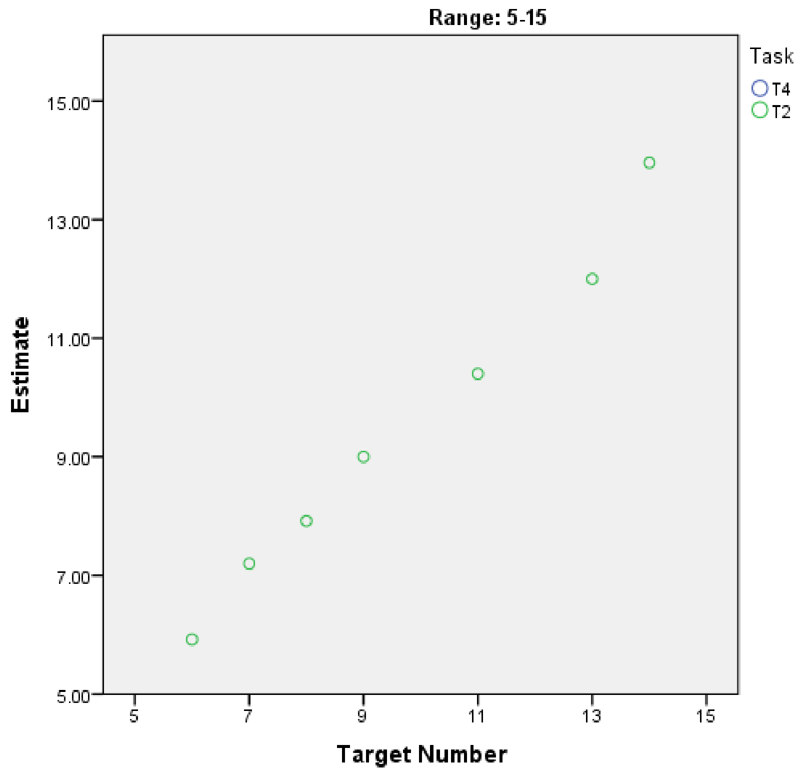


Figure 15 Imogen's range 0-20 estimates

As with the group median estimates, Imogen's Task 2 and Task 4 estimates differed most interestingly on the range 0-100. The following scatter plot demonstrates a pattern with similarities to that of the group level plot:

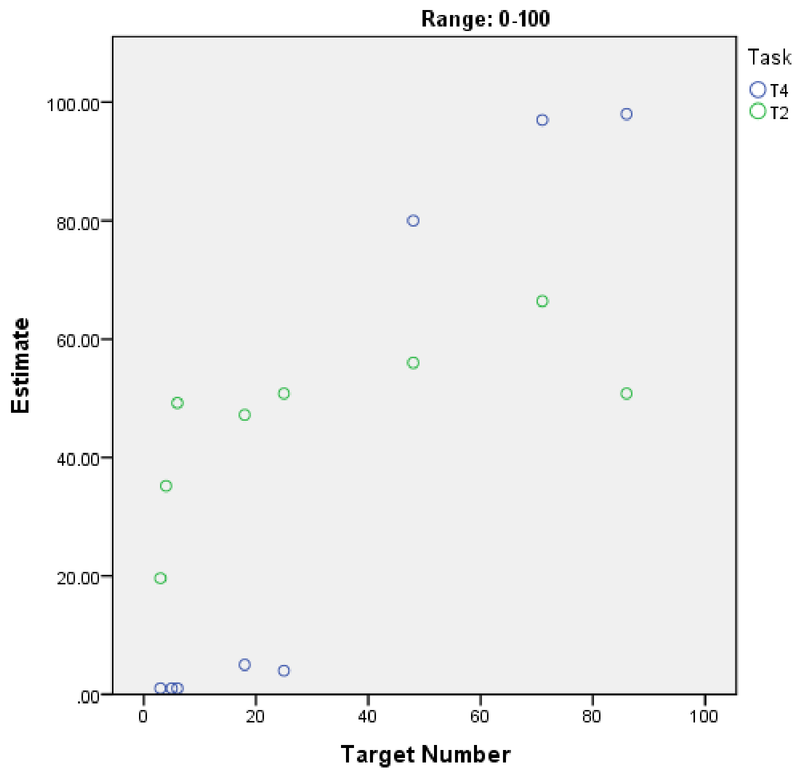


Figure 16 Imogen's 0-100 estimates

Comparing this to the observed representations of number during the task (see Figure 14), the most obvious difference is the explicit representation of the midpoint 50 in Task 2, absent in Task 4. A tentative conclusion, to be followed up, is that this is a factor in the 'clustering' of estimates around 50 shown in the Task 2 estimates (green marks) in the plot above.

Another final interesting contrast to be followed up concerns the estimation of the same target numbers within different ranges. Targets 4 and 6 were estimated in multiple ranges each, and the following chart indicates the estimates within different ranges (for Task 2 only).

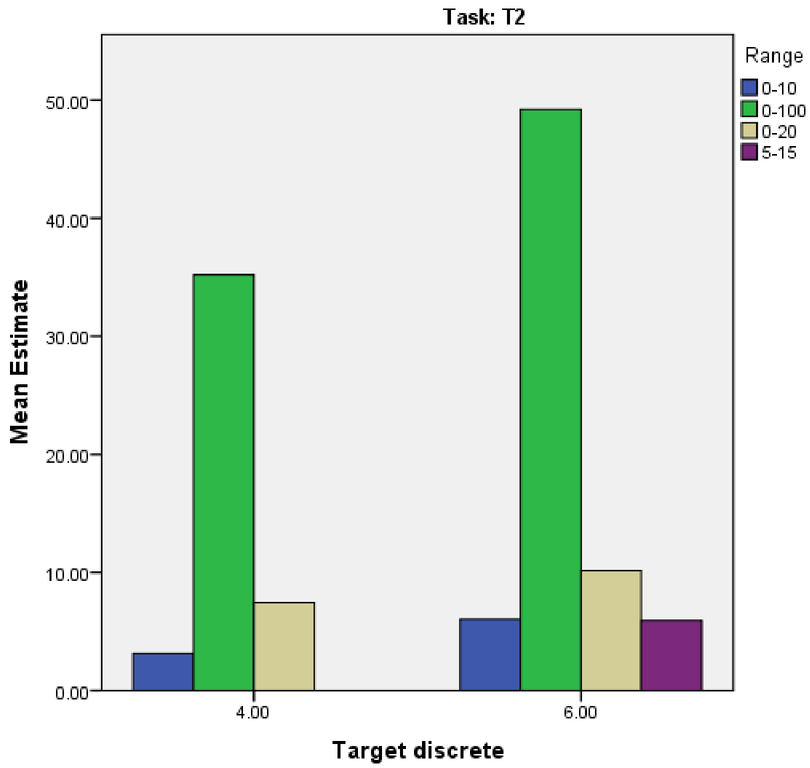


Figure 17 Differences in estimation of same target in different ranges

As hypothesised from previous studies, the same target is estimated very differently in the different ranges, and the main study will investigate how this occurs.

5.6.3 Conclusions and further directions

The findings give good reason to infer that Imogen cognitively represents number in ways which encode significant structural elements, many of which are evident in her interactions with number line estimations. Particularly of note is that she successfully applied counting strategies, commonly regarded as a less sophisticated approach (e.g. White and Szucs, 2012). What is clear from this case study is that the detail of children's interactions must be attended to; whilst Imogen indicated appropriately sized unit jumps, consistent with the evenly spaced numbers in her imagistic representations and her adjustment of unit size depending on scale, this may not be the case among other children. Aspects that may vary are the size of 'jump', whether the child attempts to scale the 'jump', and whether the size of 'jump' is consistent within a trial. Conversely, the inclusion of apparently more sophisticated structure (for example midpoints) may still result in estimation with low linear accuracy, depending on the sophistication and accurate execution of other parts of the strategy.

In terms of the relationship between imagistic representations and number line estimations, structures seen in Imogen's imagistic representation task were clearly demonstrated in the processes of translating representations during estimation tasks. Structures in common were the sequencing of natural numbers, regularity of number spacing, and the grouping of number based on multiplicative relations.

Particular findings from this case will be interesting to follow up in the longitudinal study. An example is the difference between strategies and estimation results in Task 4 (position to number) compared to Task 2 (number to position) that was evident in this case. The full study will also indicate the extent to which other children with comparatively well-developed imagistic representations of number incorporate structures of the number system into their estimations with the frequency that Imogen demonstrates.

5.7 Changes to research design based on the pilot study

This final section outlines the changes made to the research design in light of the pilot study. The main differences between the pilot study interviews and main study interviews were in the phrasing of task questions, the physical design of task materials, and in the detail of variables such as target numbers used.

5.7.1 Task 1 introduction

During the pilot task, children were not explicitly invited to explain their drawings in Task 1, for reasons explained in the Methodology chapter; it is very unclear to what extent children (or indeed adults) are able to accurately account for the representations that they associate with numbers. Nevertheless, during the pilot study many children spontaneously commented on their drawing to the researcher. This enriched the data, so the decision was made to alter the task introduction in the following way: "When you're ready, I want you to try to draw the picture that's in your head on this paper, and tell me about it, if you can." The option *not* to explain was included to avoid overburdening children; it was thus left up to individual children to decide the combination of drawing and explaining according to their preference.

5.7.2 Materials in Tasks 2 and 4

The number lines used in the estimation tasks were printed one per page. In the pilot study, the sheets were fastened as a booklet along the top edge of each landscape page. The thinking behind this was that children might more easily refer back to previous estimations if they chose. However, during the study, over half the participants commented upon the booklet feeling the 'wrong' way round, and commented that it should be 'like a book next time' or similar. A trial booklet with pages joined on the left, as in a picture book, was brought in to show participants (after the pilot study) and children agreed that it was an improvement. This design was therefore used in the main study, since the most important thing was for the materials not to become a hindrance or source of annoyance to participants.

The stickers used in the pilot study were large rockets (see Appendix 5), chosen for ease of handling by the children. Children were asked to 'use the point of the flame' to point to where the target number belonged on the line (Task 2) and correspondingly to 'look at where the flame is pointing' in Task 4. However, only two participants followed this instruction. The majority, as illustrated by the scanned example in Appendix 5, stuck the rocket body itself onto the line, and showed signs that the target number was perceived to 'take up' the width of the entire sticker on the number line. This was clearly problematic in terms of the spatial positioning of numbers. For the main study, the stickers were therefore re-designed (see Appendix 3). In the first place, the rockets themselves were made substantially narrower, and printed onto the narrowest stickers possible (15mm wide). Additionally, to emphasise that the rocket should *point to* a location on the line, rather than 'occupy' or 'take up' a portion of the line, a long thin arrow was added from the rocket flame.

5.7.3 Range order in Task 2 and Task 4

Initially, the order of the number ranges were separately randomised for each child, following the method of Berteletti et al. (2010). However, during the pilot study it was apparent that children had very different levels of confidence with the larger numbers (regardless of their *knowledge* of larger numbers) and that the randomised order of number ranges therefore led to overly different interview experiences for different children. Those who began on smaller number ranges demonstrated a great deal more confidence and were able to build up to the ranges considered more difficult. For the main study, it was therefore decided to maintain the order 0-10, 0-20, 5-15, 0-100 for all children.

In the pilot study, for reasons of time, children only completed Task 4 for two number line ranges each, rather than all four. Half of the children completed the task on 0-10 and 0-100 only, and half of the children completed the task on 0-20 and 5-15 only. The selection of children for each group was random. The findings from this task were sufficiently interesting that it was decided to expand the task to include all four number ranges.

5.7.4 Target numbers

Target numbers were altered to the final lists described in the Methodology chapter. These changes were minor, and mainly focused on how to maximise saturation of target numbers without making the overall interview too long.

Chapter 6 Findings from the group

This chapter presents the first stage of the research findings. It begins with introductory comments about the sample, and brief explanation of some choices and conventions in what follows. It then presents the findings relating to the whole sample of children. The analysis in this chapter focuses mainly on quantitative analysis of children's estimation data, and asks what can be said in response to the research questions from this perspective. This chapter is particularly important for linking the current research to the main body of existing research on number line estimation, given the extent to which this area of research has been dominated by quantitative studies. The second and more important section of the findings consists of three in-depth case studies, which are presented in the following chapters. It is in this second section that the research questions are explored to their full depth, by balancing both qualitative and quantitative analyses of three contrasting cases in fine-grained detail.

6.1 Abbreviations

Throughout the findings, the following abbreviations are used:

- T1, T2, T3 and T4 = Task 1, 2, 3 and 4 respectively, where:
    - T1 asked children to imagine the counting numbers to 100 and invited them to draw and describe what came to mind
    - T2 consisted of number-to-position estimations on ranges 0-10, 0-20, 5-15 and 0-100 in that order
    - T3 asked children to estimate the quantity of small sweets in transparent tubs
    - T4 consisted of position-to-number estimations on the same ranges as T2
- R1, R2, R3, R4, R5 = Rounds 1, 2, 3, 4 and 5 of the task-based interviews, carried out between October 2012 and July 2013.

6.2 Participants

The children who participated in the five rounds of task-based interviews were thirteen Year One pupils from a single class at a local South of England primary school, as described in Chapter 4. Their pseudonyms, and the mathematical attainment groups that the teacher placed them in at the start of the year, are shown in Figure 18. During the school year, children including those in the sample moved between these groups, and the teacher-assessed mathematical attainment groups at the end of the year are shown on the right. For each group of students, the proportion of the group included in the sample is shown. For example, the three sampled children in "blue" group at the start of the year constituted 3/8 of the total blue group.

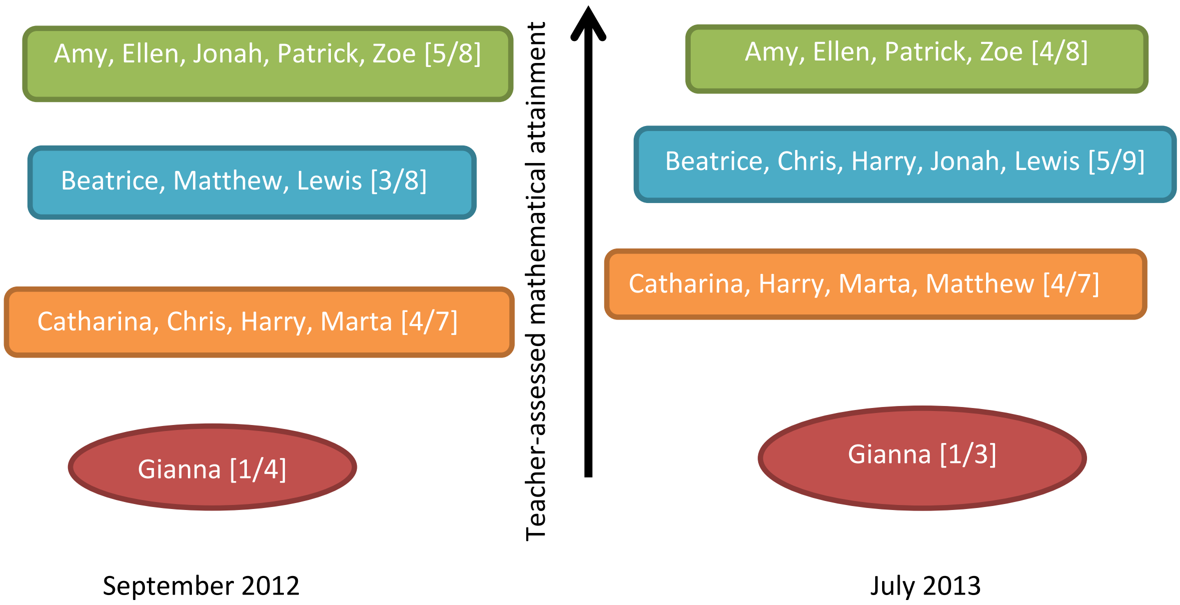


Figure 18 Teacher-assessed mathematical attainment groups

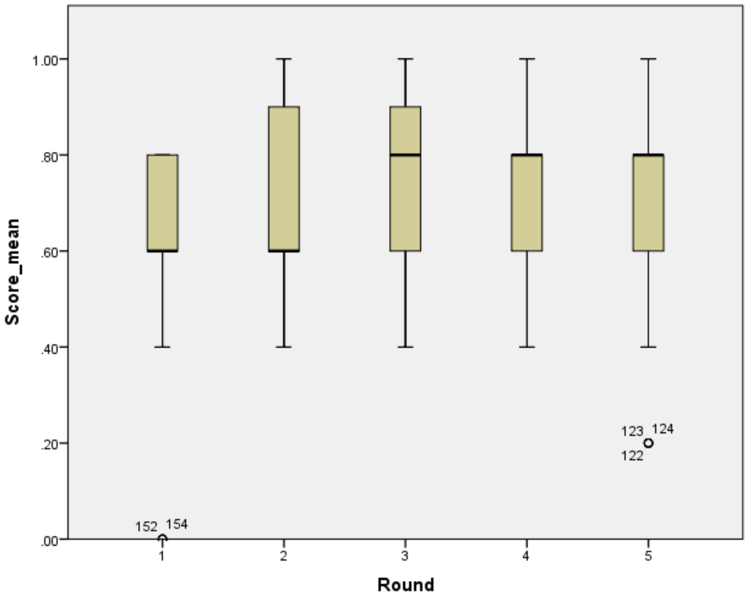
After careful consideration, Gianna's data were excluded from the main data analysis. Gianna's number line estimations were uncorrelated with the presented target number, for all ranges (Spearman's rank correlation, p>.05), and her other interview responses also proved impossible to analyse in mathematical terms. Following the practice of Slusser et al. (2013, p. 197), Gianna's data were therefore excluded from further analysis. Matthew's T4, range 0-10 estimations from R3 were also excluded from analysis. The estimations were uncorrelated with the target numbers, and Matthew's comments to the researcher strongly suggested that he was joking.

6.3 Task 3

Before beginning the main presentation of findings, this final preliminary section gives a brief treatment of interview task 3 (T3). As this section aims to show, the findings from this task were limited, and do not form an integral part of this study's findings. For this reason, T3 is briefly discussed here, and not in the following sections or chapters.

As explained in the methodology chapter, T3 was included for multiple reasons: to look for connections between children's estimation accuracy in different tasks, to provide a different kind of task to interest children, to provide a break between the two substantial number-line estimation tasks, and to examine whether different representations of number occurred during this different type of estimation task. Children were presented with sweets or beads in a transparent container, and asked to choose between two possible answers. Children were randomly assigned one of two sets of target quantities for their first interview: either Set A – [14, 22, 34, 46, 58], or Set B – [18, 24, 32, 48, 60]. In each subsequent round, children were given the alternate target set. For each target number, the alternative answer option was randomly chosen to be either half, one and a half, or double the value of the target number. An example set of answer options using Set A would be [7, 14], [22, 33], [34, 68], [23, 46], and [58, 87].

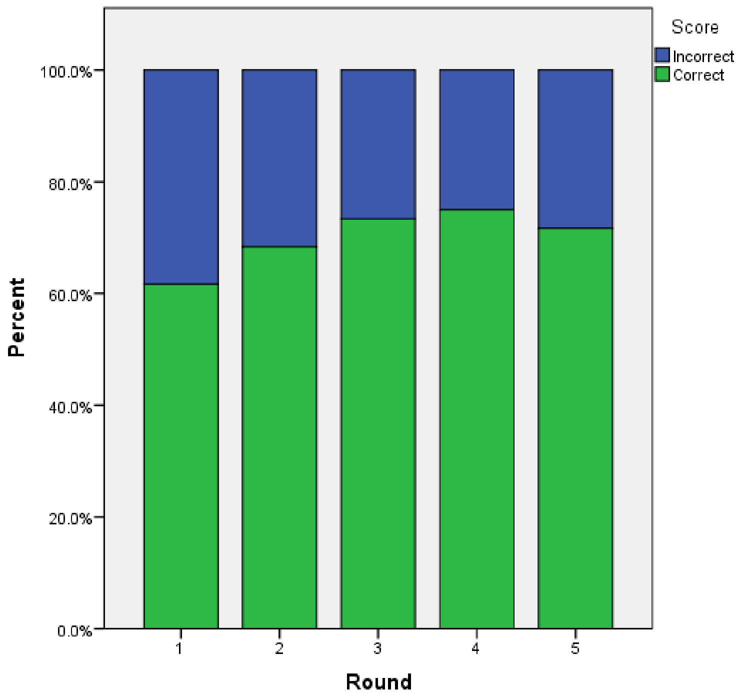
Children completed five T3 estimations per interview, so a score of 80% indicates that a child chose the correct answer option for four out of five quantities in that interview. Children's accuracy of estimation in T3 did not vary a great deal either between participants or over time, as the following boxplots of mean scores demonstrate.



In almost all interviews, children estimated between three and five of the quantities correctly.

Figure 19 T3 estimation accuracy by round

The changes in mean T3 score between rounds are shown more clearly here:



A one-way repeated-measures ANOVA tested children's T3 scores and found no significant change in T3 estimation accuracy with interview round, $F(4,44)= 1.279$, $p=.293$.

Figure 20 T3 estimations by round, all children

Although there was no significant change in children's T3 scores over time, there was a significant association between the ratio of presented answers (ratio of additional answer option to correct answer option) and children's success in estimation, $\chi(2)=22.8$, $p<.001$. There was a strong expectation that children would find trials most difficult, and experience least success, when the ratio of additional answer to correct answer was one and a half, i.e. when the ratio of the two answers was closest to one. The highly surprising finding was that children in fact chose the correct answer least frequently when the ratio of the additional answer to the correct answer was half (see central bar in Figure 21). The order of answer options was random, and clearly for both the "double" and "half" categories the resulting ratio of answer options was two, so this finding is currently very puzzling.

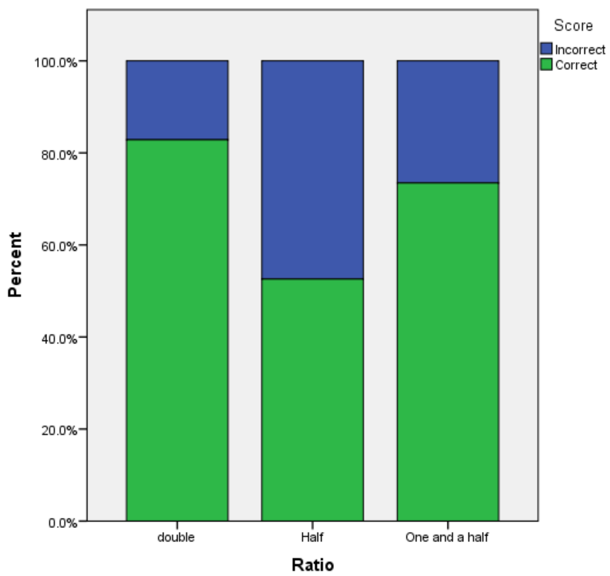


Figure 21 T3 estimations according to ratio of additional answer option to correct answer option

Somewhat disappointingly, children did not tend to represent number during T3 except for stating their chosen answer. A few children offered spontaneous justifications for their choices, and these were all verbal statements indicating that one answer option was too big or small. A typical example of this is Catharina's comment in T3 R4:

J: And … this one here: do you think there are twenty-four or twelve?

C: Twelve. …. Cos there's less.

Since this was the extent of children's representations of number during T3, the decision was made to focus analysis on the other three tasks. For this reason, T3 is not addressed in the main findings or case studies.

The following section presents the first stage of the research findings, containing analysis of the whole sample organised in response to the relevant research questions.

6.4 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research?

6.4.1 What are the modes and component signs used in the representations?

Children represented aspects of number using speech, inscriptions, eye gaze and gesture, and these representations occurred throughout the task-based interviews.

As expected from eye-tracking research literature, the component signs identified in eye gaze were periods of fixed or lingering gaze (on a point) and saccades (brief, rapid, "jerky" movements from one resting position to the next). Counterpart signs to these were also found in gesture: pointing to fixed locations, and saccadic movements of a pointing hand respectively. More generally, gestures are described in terms of the framework outlined in the methodology: in terms of their physical position in space, speed, and hand shape. Many of children's gestures, like their eye movements, were focused on the number lines printed on the page.

In response to T1, most children produced inscriptions with some verbal description. The component signs children used in their T1 representations ranged from pictorial elements to extensive use of notational signs. An example of pictorial component signs can be seen in Beatrice's T1 response from R4:

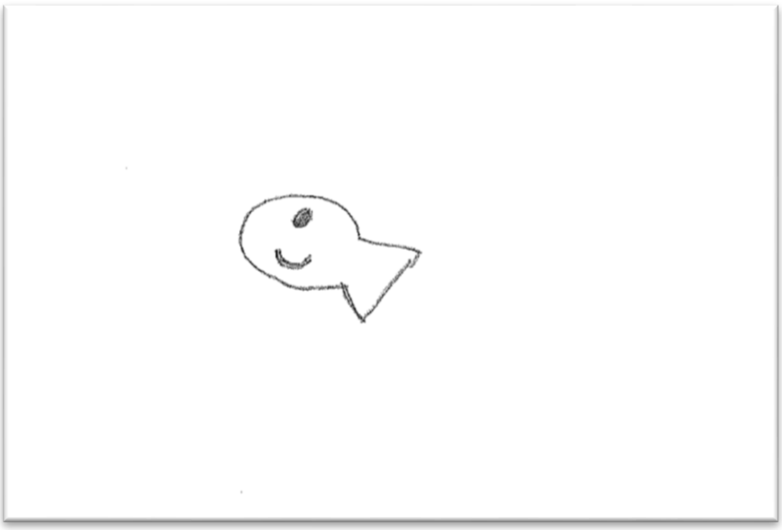


Figure 22 Beatrice R4: "My fishes"

Beatrice's sole comment was the following:

B: My fishes.
J: Yes?
B: [*smiling*]

Beatrice appeared to associate the number one hundred with the image of a familiar numerous set (her fishes). The same image was drawn in R5.

A typical example of notational component signs in T1 can be seen in Harry's R4 response:

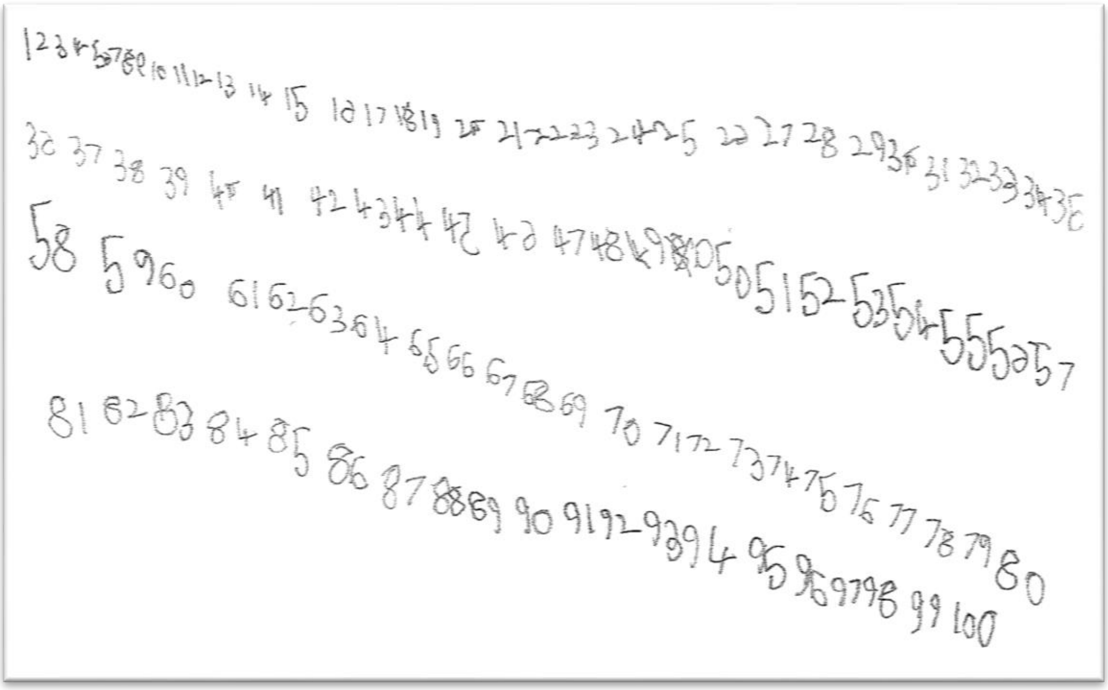


Figure 23 Harry R4

Harry did not make any verbal comment during this task. He wrote energetically – almost hurriedly – except for significant pauses at some changes of decade. The inscription above records some evidence of these pauses in the gaps seen between 60 and 61, and between 69 and 70.

6.4.2 What aspects of number structure are represented?

6.4.2.1 T1 representations

The representations produced in T1 were analysed in terms of the framework developed by Thomas and Mulligan (1995) specifically to describe the structure of children's inscriptions during this task (see section 4.4.1). Their framework identifies the following aspects of number structure in responses to this task:

1. No structure: elements show no apparent relationship to equal groupings or sequence.
2. Linear structure: elements in linear formation (straight or curved), numbers in sequence.
3. Emerging structure: "one hundred represented by equal groups of objects, or linear sequence broken into equal segments" (Thomas and Mulligan, 1995, p. 12).

4. Emerging structure (m): representation shows some aspect of multiplication, such as multiple count and multiplication grid.
5. Partial array structure: elements are in rows and columns, but not a ten-by-ten array.
6. Array structure: elements in a ten-by-ten pattern.

In this study, T1 representations at each of the above levels were observed. The following chart displays the number structure represented by each child in each interview round, with darker shading to indicate representation of number structure considered more sophisticated:

	R1	R2	R3	R4	R5
Patrick C01	Linear	Linear	Array	Emerging (m)	Emerging structure
Amy C02	Array	No structure	Linear	No structure	No structure
Ellen C04	Linear	No structure	No structure	Linear	Emerging (m)
Jonah C13	Emerging	Emerging (m)	Emerging (m)	Partial array	Partial array
Zoe C05	Array	Linear	Linear	Linear	Linear
Beatrice C11	—	—	—	No structure	No structure
Matthew	No structure	No structure	No structure	No structure	Linear
Lewis C09	No structure	—	—	—	—
Catharina C08	No structure	No structure	No structure	No structure	No structure
Chris C12	—	Linear	—	Linear	Linear
Harry C06	Linear	Linear	Linear	Linear	Linear
Marta C10	No structure	No structure	Linear	No structure	No structure

Figure 24 Number structure represented in T1 responses

Repeating this task in five different interview rounds was revealing. The most noticeable finding was the variety of responses that individual children offered, which provides direct evidence for comments in earlier research about how a single instance of the task provides only “partial description” of representational capabilities (Thomas et al., 2002, p. 130). Amy, for example, represented numbers in array structure in the first round, but represented no number structure at all in three of the following rounds. The number structure represented by children in the task did align well with teacher-assessed mathematical attainment. The five children who represented number structure more complex than linear structure in at least one interview (shown in the first five rows of the above table) were also the five children in the highest teacher-assessed mathematical attainment group at the start of the year.

Because this task is open to different interpretations and offers such a partial view of children’s representations, clear developmental trends were not expected. The repetition of T1 in each interview was viewed more just as multiple opportunities for children to produce representations of the numbers one to one hundred. In this sense it functioned well, allowing some children to create widely varying representations of number structure.

The frequency of responses best-described by each level of number structure did not change markedly between R1 and R5:

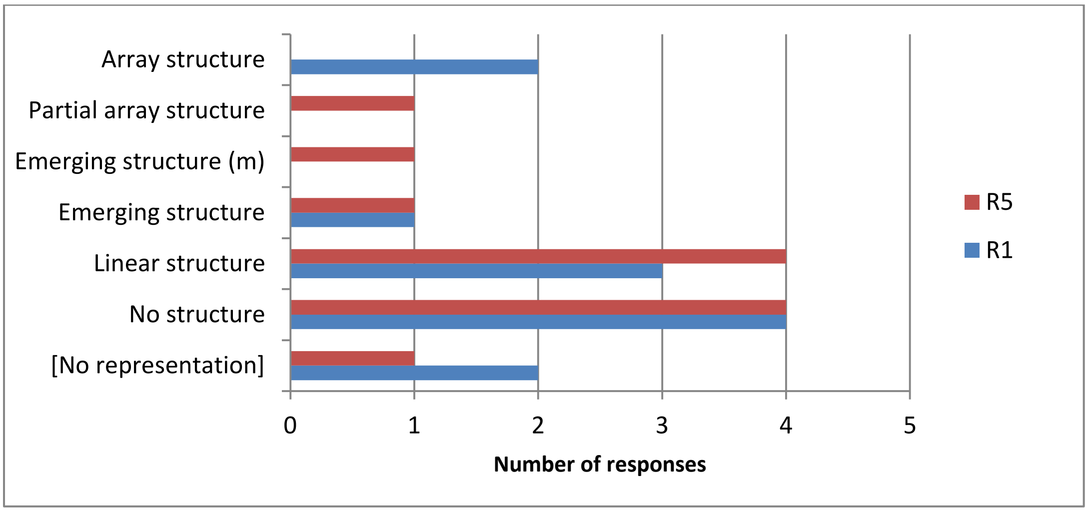


Figure 25 Frequencies of number structure representations in T1 responses

6.4.2.2 T2 and T4

Moving beyond T1 to consider the interviews overall, children's representations during all tasks were analysed according to Resnick's microstage theory of number understanding development (Resnick, 1983). As described in the theory for analysis (see section 4.4.2), Resnick's theory includes mathematical procedures and behaviours that exemplify the particular aspects of number structure at each stage. This aided the identification of number structure in children's task responses. However, the present research was not restricted to analysis of mathematical procedures and behaviours; all cognitive representations were examined for evidence that they represented aspects of number structure. Episodes identified as containing cognitive representation of number sequence, for example, include both episodes featuring task strategies that represent this structure, such as counting-on from the left endpoint, and also episodes where children verbally, gesturally or otherwise represented number sequence outside of particular procedures, for example verbally describing to the researcher why the estimation of target 6 in range 5-15 was "easy".

The following table shows the stages of number structure in Resnick's theory (first column), followed by the elements that make up that stage (second column). The third column provides examples of children's procedures and strategies that include representation of this particular aspect of structure.

Number structure	Detailed stages	Procedures or strategies in which this number structure may be represented
S Sequence	Count sequence from 0 or 1	Count-on from LE.
	Count sequence from other start point	Count-on from non-zero LE. Count-on from estimate to RE for confirmation.
DS Decreasing sequence	Count sequence reversed	Count-back from RE.
RN Relative numerosity (quantity comparison)	Ability to represent (esp visualise) number's magnitude without counting up to it	Applying efficient count strategy, using RN to choose shorter available count. Checking judgements, querying whether an estimate is likely. Use of endpoints as "anchor" points.
PW Part-whole	Partition and recombine numbers (<20)	Use of "anchor" points based on partitioning the range.
	Use of number bonds to ten	Checking estimations using number bonds to ten. Using own midpoint anchor on 0-10 range.
	Multiple partitions of multi-digit numbers	Using own midpoint anchors on 0-20 and 5-15 ranges.
PW10 Part-whole understanding of base ten structure	Numbers as compositions of tens and units	Using own midpoint anchor on 0-100 range. Count-on and count-back by tens. 10x10 array structure of 0-100
	Multiple partitions of larger multi-digit numbers	Count-on or count-back by any given number. Partitioning into quarters, thirds, other.

Figure 26 Aspects of number structure represented according to Resnick's (1983) theory

It is of course not always possible to determine number structure represented, and some strategies can be interpreted as representing different levels of number structure. For example, using the endpoint ten as an anchor for estimating target nine shows, at a minimum, knowledge of the counting sequence and relative numerosity (to choose the efficient use of ten, instead of counting until the count sequence reached nine). However, the same strategy could also arise from a part-whole understanding of ten in which the number bond 9+1 was used to position nine near ten. What Resnick's development account emphasises is that once part-whole relations for a given range begin to be understood by a child, it is not a question of 'either/or': the part-whole schema "allows numbers to be interpreted *both* as positions on the mental number line *and, simultaneously*, as compositions of other numbers" (Resnick, 1983, p. 146).

In terms of coding task responses, attending to representations in all modes provided as much evidence as possible about the number structure involved in the particular task response. Aside from this, coding was carried out conservatively, identifying the *minimum* level of number structure necessary for a representation to be created or strategy to be applied, rather than the number structure *potentially* involved. In terms of interpreting task interactions, the balance between considering single trials and the interview context became critical. The current research specifically set out to examine the estimation of individual targets, hypothesising that features unique to the target might be influencing estimations (see White and Szucs, 2012, p. 9). However, knowledge of the number structure a child represented in preceding trials, or moments, and comments made subsequent to the trial, provided an extremely valuable second layer of evidence for interpreting the estimation process.

6.4.2.3 Numbers-space 'fit': the number line and the idea of scale

In addition to the aspects of number structure addressed in the frameworks already mentioned, children in this study represented another type of structure, which was the structure of how numbers 'fit' into a given space. The representations in this category include many where children express opinions about how many numbers 'should' be represented by a given physical space, and also more mathematically sophisticated statements which begin to express the idea of scaling a linear representation depending on the range and space provided.

This category is not included in either of the theoretical frameworks used to discuss number structure so far (Resnick, 1983; Thomas and Mulligan, 1995). A working hypothesis to explain this is that the explicit number-space association forced by number-line estimation tasks caused children to represent their ideas about number and space more frequently than would be expected.

Representations of number-space 'fit' occurred in interviews with children in each of the teacher-assessed mathematical attainment groups. Not all children made representations in this category, but those who did include both the most and least linearly accurate children in terms of number line estimation.

A distinctive feature of children's number-space 'fit' representations is that they were made at the point of introduction of new number ranges, i.e. when the page turn revealed a number line with a new range, but before the researcher had actually presented a target number to estimate. The representations did not tend to occur within estimations trials themselves, and on the few occasions where they did, the focus was on the relation of the number range to space, and the target number tended to be excluded.

Children's number-space 'fit' representations fell into three distinct sub-categories:

- Not enough space to represent a given range
    - Difficult or unnatural to represent range in given space
- Too much space to represent a given range

    - Difficult or unnatural to represent range in given space
- The need to adjust representation given the range and allocated space for representation

The following examples provide typical examples of these sub-categories.

Not enough space for given range

Patrick, R4, T2, range 0-100:

P: It is very hard because a hundred [*RH makes chopping motion at RE*] … can't fit like a hundred [*RH traces LE to RE*] of those … so have to squeeze them all in …

Zoe, R2, T2, range 0-100:

Z: I wish the zero was there [*points left page edge*] and then the hundred was there [*right page edge*].

Jonah, R5, T2, range 0-10:

J: It - was if it was 100 it should be longer!

Too much space for given range

Catharina, R5, T2, range 0-10:

C: They're quite big for putting all the numbers in! [*chopping motion with both hands, LH at LE and RH at RE, simultaneously. RH palm down then sweeps to and fro between LE and RE.*]

Matthew, R5, T4, range 0-20, target 6:

M: … [*writes answer*] twelve again… It's actually twelve and cos, I know that's there [*points to previous rocket with pencil*] and it's there [*points to equivalent location on current line*]. Cos they've got loads of space [*LH moves quickly to and fro between LE and rocket*].

In this episode, Matthew appears to explain why both the previous and current blank rockets represent the number 12, despite being at different positions on the number line 0-20.

Adjusting representation given the range and allocated space

Patrick, R1, T4, target 6:

P: … seven. …. Because even though it's a bit low, we're going up to twenty …

J: Yes

P: Got to remember the total line.

Matthew, R4, T2, range 0-100:

J: where could 92 go?

M: One hundred … [*LE c/o x3 slowing down and stopping… suddenly points to RE*] - back from one hundred! [*mimes count-back which turns into a wiggly line leftward*]

J: Oh?

M: [*looking at RE*] 100, 99, 98, [*looking at J*] 97, 96, 95, 94, 93 [*looks at own hand again, almost at LE*] -No! That is too big, because look - [*demonstrates RE count-back with exaggeratedly large unevenly-sized jumps, reaching LE*]

J: Ah, yes.

M: [*picks up pencil, hesitates*]

J: It's OK to draw it on the line if you want.

M: 100, [*points RE, looking at own hand*] ... 99 98 97 96 95 94 92. [*counts-back from RE x8 small, careful jumps with pencil point*] Ninety three? [*marks line with pencil then sticks rocket at mark*]

Representations of number-space 'fit' often coincided with representation of relative numerosity and part-whole structure. The following sequence of episodes from Zoe's R4 interview demonstrates some of this overlap. The task is T2, and the range is 0-100:

J: What about three, where does three go?

Z: Here. [*emphatic, sticks rocket near LE*] Cos otherwise there won't be space for every other number [*LH sweeps line from rocket rightwards to RE*]

J: Number six.

Z: [*straight to LE, glides small distance rightwards, sticks rocket*]

J: Good.

Z: Oaah [*points just right of rocket*], there needs to be ninety about, hmmm, there [*points about 2/3 of the way along the line*]

[*J presents 71 rocket*]

Z: Hmmm. [*takes rocket 71 near RE, slowly glides leftward towards MP*] Need to leave QUITE a big space [*sticks rocket*]

J: Yes

Z: Otherwise there won't be space for eighty [*points to right of rocket*] and ninety [*hand jumps further to the right and points again*] and –

In the above episodes, Zoe represents emerging part-whole structure – both vague ("every other number") and more precise (the decades following 71 that must be represented between 71 and the right endpoint) – and this appears to guide her judgment about number-space 'fit'.

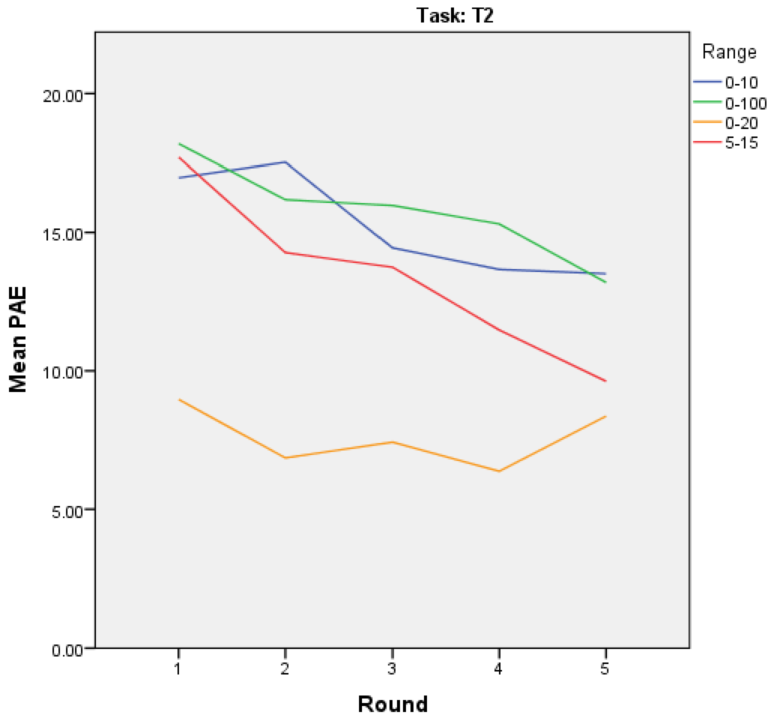
Number-space 'fit' representations are particularly interesting because of their strong link to the precise details of task context, and children's representations in this category introduce new possibilities for understanding number line estimations. One such possibility is an explanation for the high estimation accuracy that children showed for the range 0-20 (see RQ1c for further detail). Contrary to expectations based on previous research, children in this study estimated more accurately in the range 0-20 than in ranges 0-10 or 0-100. Since children expressed both the idea that there was too much space to represent 0-10 and not enough space to represent 0-100, one thought is that the classroom resources in this Year 1 classroom, which overwhelmingly used the length of a piece of A4 paper to represent ranges 0-20 or 0-30 (range 0-10, 0-50 and 0-100 were represented in other formats and shapes within the classroom) had accustomed children to perceiving this scale as 'natural' and to using it as their default scale.

Children's number-space 'fit' representations offer the potential for deeper insight into estimation processes than is provided from identification of strategies. In the final episode of the above examples, observation establishes that Zoe was likely using the right endpoint (one hundred) as a reference or 'anchor' point, and that her estimation was based on judging the target 71 relative to one hundred. Zoe's subsequent representation of how the numbers between the target and right endpoint need to 'fit' provides another level of insight into how the relative judgment was made, and in this case helps to understand why Zoe's eventual placement of the target 71 underestimates its position relative to a linear representation of this range.

6.4.3 What are the notable between-task and within-task connections between representations?

This section of the findings, considering children's results as a group, focuses on between-task and within-task differences in terms of the linearity of children's representations. As explained in the postscript to the literature review, the debate over whether a log-linear or proportional judgment model better describes change in estimation accuracy shows no sign of being resolved. In recent research, some high-profile authors seem to have acknowledged that a more fruitful path is to accept as an observed phenomenon that with age and experience children's number line estimations "more closely approximate a linear function" (Fazio et al., 2014, p. 54) and focus attention on this alone. A good way to do so is through analysis of the percent absolute error (PAE) of individual estimations. Although this is commonly presented as a measure of estimation accuracy, it is more precisely a measure of an estimate's linear accuracy, since the "error" is calculated from an assumed linear distribution.

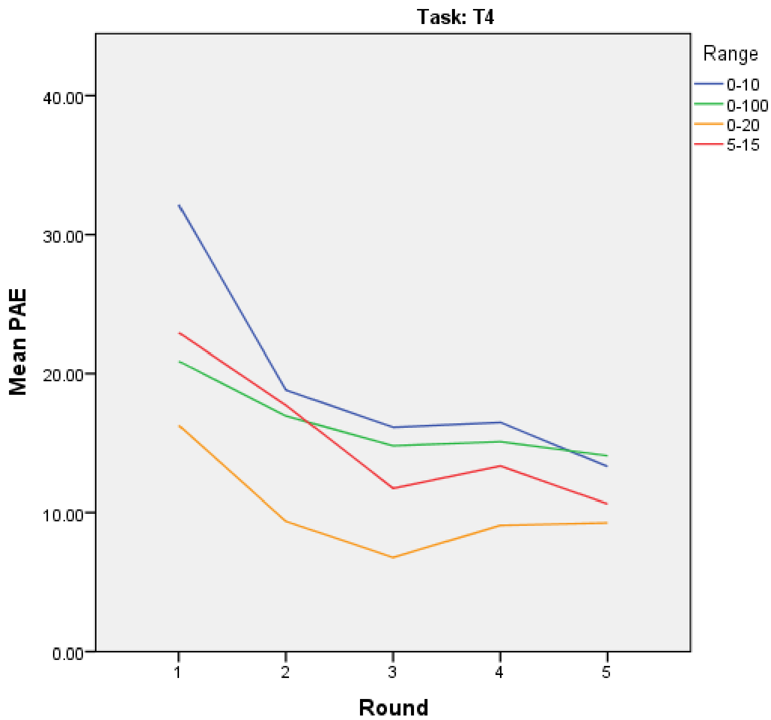
The following two graphs show children's mean PAE in each range, over the five interview rounds, firstly T2 then T4. Although the data points indicate mean PAE in different rounds, the points were joined to form a multiple line graph, as shown.. This practice was adopted following the similar usage of Rouder and Geary (2014) and Ashcraft and Moore (2012). Although it is mathematically incorrect to join the data points in this way, it allows the eye to distinguish changes of PAE in each range, to an extent that was not possible without joining data points.



Children's estimations were most accurate in the range 0-20, and PAE decreased with interview round for estimations in all ranges.

The largest increase in linear accuracy (decrease in PAE) was for estimations in the range 5-15.

Figure 27 Mean PAE of T2 estimates



In T4 as for T2, PAE decreased with interview round for estimations in all ranges. Again, children's estimations were most accurate in the range 0-20.

The largest increase in linear accuracy (decrease in PAE) was for estimations in the range 0-10, particularly between R1 and R2.

Figure 28 Mean PAE of T4 estimates

Individual children's mean PAE for each task condition was calculated, and a three way mixed ANOVA was carried out to further investigate how mean PAE varied with range, task and interview round.

Mauchly's test was used to check the assumption of sphericity (the equality of variances of differences between all pairings of repeated-measures groups), which if violated is often a source of Type I error. Mauchly's test indicated that the assumption of sphericity was violated for both interview round ($\chi^2(9)=81.73$, $p<.05$) and the interaction of interview round and task ($\chi^2(9)=103.48$, $p<.05$), i.e. that the variances of differences between groups were unequal. The degrees of freedom for both were therefore corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon=.442$ and $\epsilon=.358$ respectively). Levene's test of homogeneity of variance between groups (number ranges) was significant at 5%, but not 1%, for two of the ten task x interview round conditions. The significance of results relating to differences between range groups should therefore be treated with some caution. All effects are reported as significant (or not) at 5%.

As the graphs of PAE strongly suggest, there was a significant main effect of interview round on the mean PAE of children's number line estimations, $F(1.77, 54.82)=7.74$. Planned contrasts revealed that this reflected a significant linear trend, $F(1, 31)=14.21$, with mean PAE overall decreasing (linear accuracy of estimation increasing) proportionally with interview round.

There was also a significant main effect of task, $F(1, 31)=8.13$, with the mean PAE of T4 estimates significantly higher than the mean PAE of T2 estimates. Testing also identified a significant main effect of range, $F(3, 31)=7.91$, $p<.001$, although this needs to be regarded with some caution for the reasons noted already. Post-hoc tests indicated that estimations in the range 0-20 were significantly more accurate (had significantly lower mean PAE) than estimations in ranges 0-10 and 0-100.

Testing identified no interaction effects. Neither interview round and range, $F(5.31, 54.82)=.751$, task and range, $F(3, 31)=.932$, interview round and task, $F(1.43, 44.33)=3.36$, nor the interaction of round, task and range simultaneously, $F(4.29, 44.33)=.768$, were significant.

6.5 RQ2: What strategies can be identified in children's interactions with number line estimation tasks?

Strategies in response to estimation trials were inferred from children's representations during T2 and T4 trials. The list of strategies identified was the following:

- Reference to potential anchor point
    - Those included in the task environment (left and right endpoints)

    - Those created or visualised by child
        - Midpoint
        - Quarter-point
        - Three-quarter point
    - Previous trials of the task
        - Within the same range
        - Within a different range
- Counting strategies
    - Count-on from LE to estimate
    - Count-on from estimate to anchor point (e.g. RE)
    - Count-back from RE
    - Count-back from midpoint
- Judgment using relative numerosities: particularly clear when estimate, or further strategies, appeared to be influenced by initial use of relative numerosity of target and task environment.

Task responses that did not themselves constitute strategies were also observed. Those which were recognised in multiple children, or in multiple trials of the same child, with a recognisable pattern of behaviour, were coded in a separate set of "task responses" (i.e. not task responses that did not become strategies). The list of such task responses codes is the following:

- Change of mind
- Immediacy
- Hazard – initial response led child to a recognised mathematical contradiction
- "Easy" – explicit indication that a trial was found easy

Appendix 6 demonstrates a typical example of the behaviours relating to each of the above strategies and codes.

Count-on strategies were each further coded to indicate whether the rocket sticker had been placed at the end of the count (in the case of T2 trials) or the target number written in agreement with and after the count (in the case of T4). There were many cases in which this was not the case, for example where count strategies appeared to be being used as confirmation or checking of estimates, or where children rejected the result of their count strategy and decided to move on to a different tactic.

The sub-questions of this research question concern a detailed view of individual estimation trials, and so are addressed only in the case studies.

6.6 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling?

6.6.1 In what ways does evidence support or not support the log-linear hypothesis?

The first step in answering this question was to carry out the typical analyses of number-line estimation data found in previous research into the log-linear hypothesis, for example Siegler and Opfer (2003, p. 239). This involved fitting linear, logarithmic and exponential models to the median estimates of the sample. This procedure was carried out once for each of 40 task conditions (4 number ranges x 5 interview rounds x 2 tasks). Models that did not reach significance were excluded from further analysis, and the proportion of variance explained (measured by R^2) was calculated for each remaining model. The series of graphs on the following pages shows the calculated R^2 model fits for each model in each task condition. It is followed by a table which summarises the best-fitting model for the participating children's median estimates in each task condition.

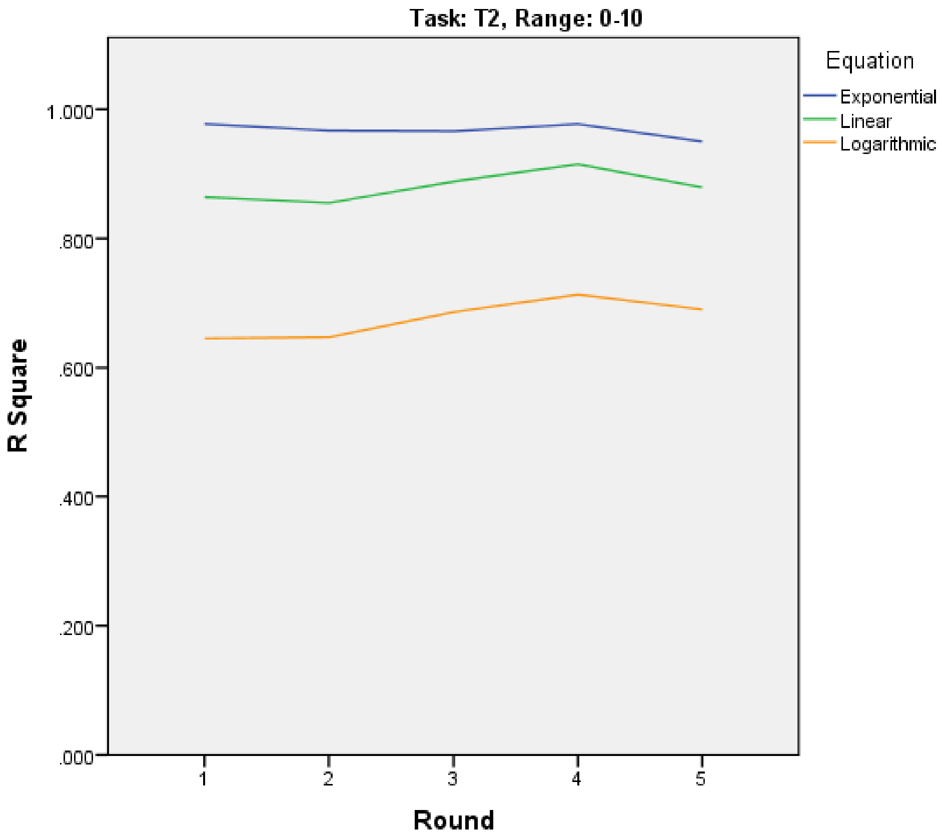


Figure 29

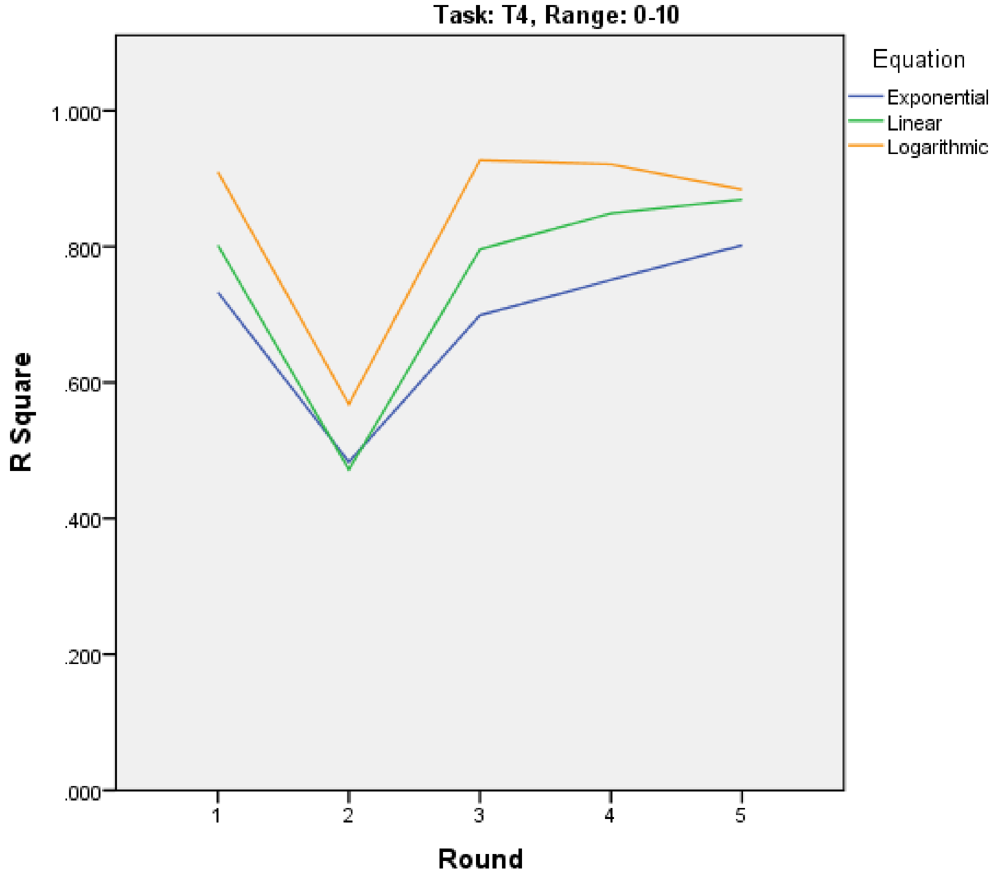


Figure 30

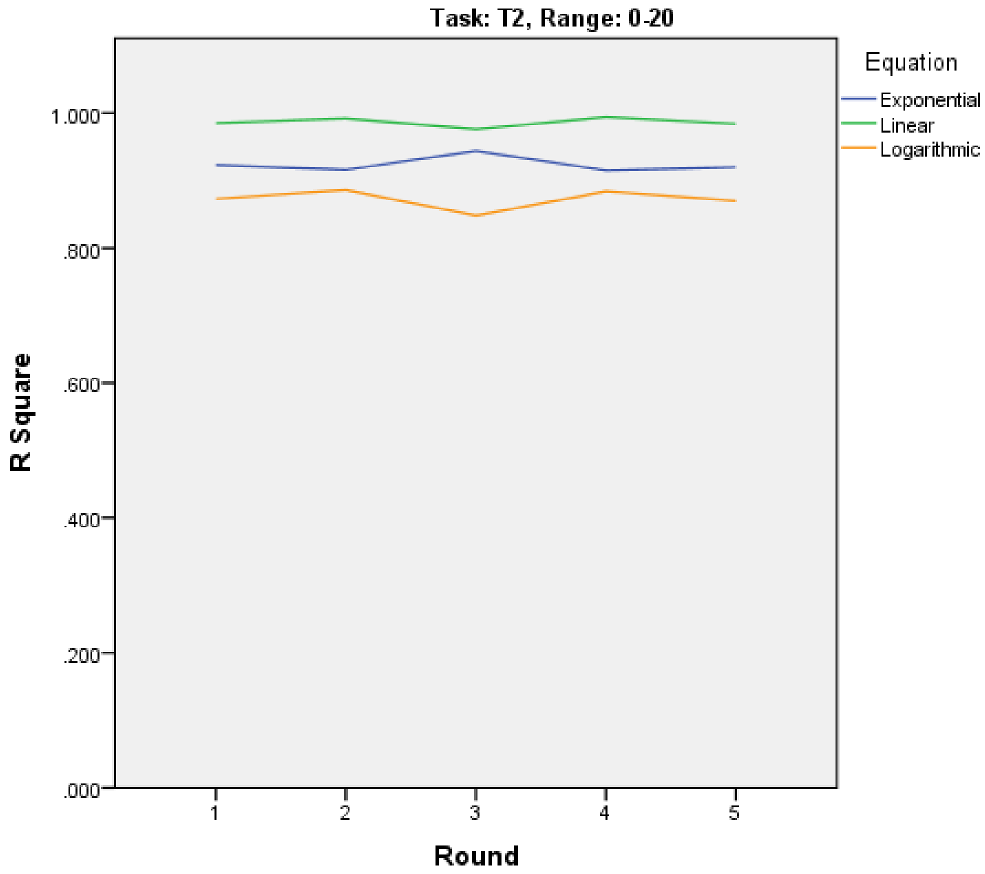


Figure 31

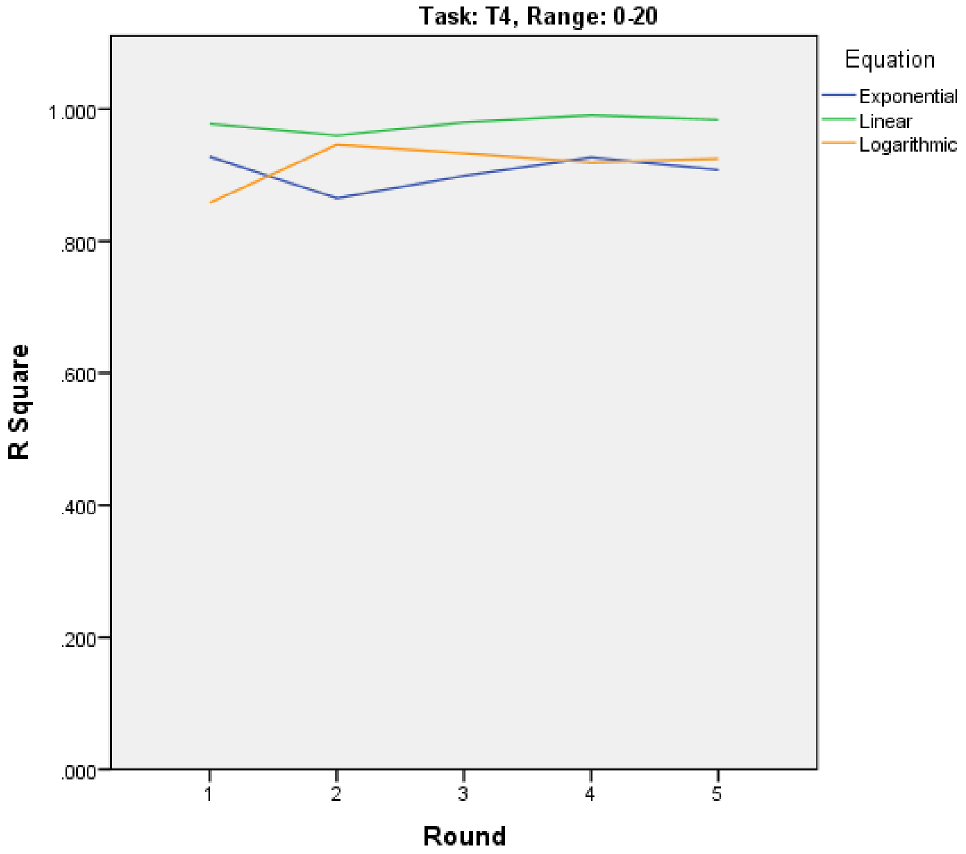


Figure 32

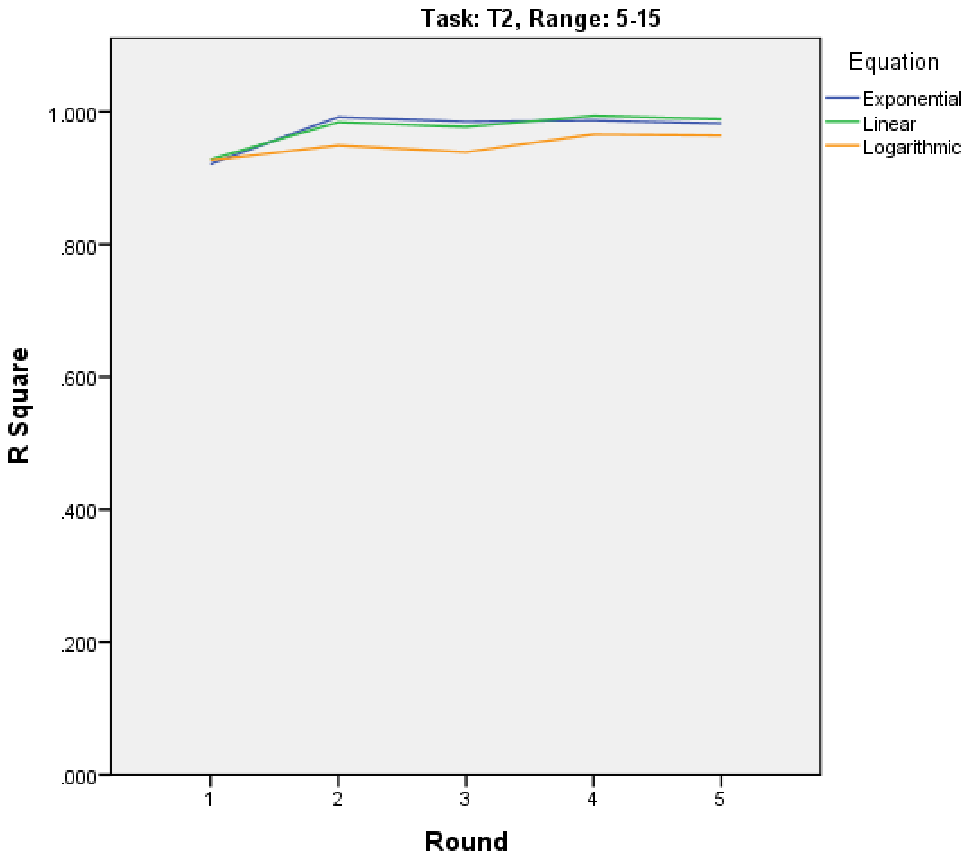


Figure 33

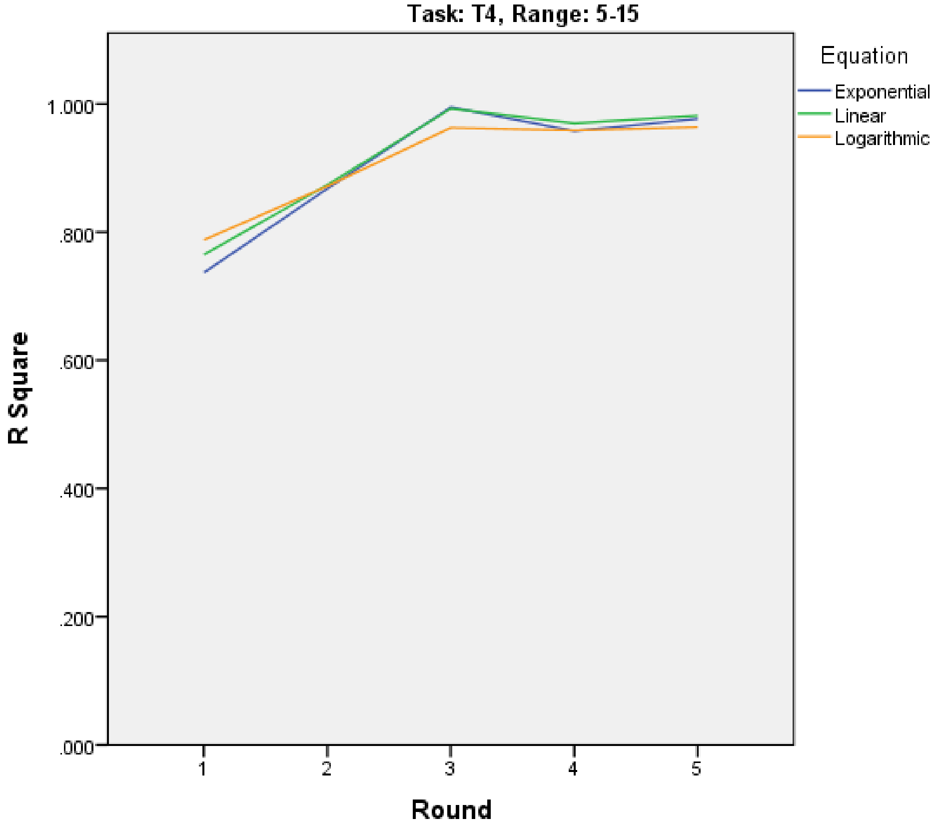


Figure 34

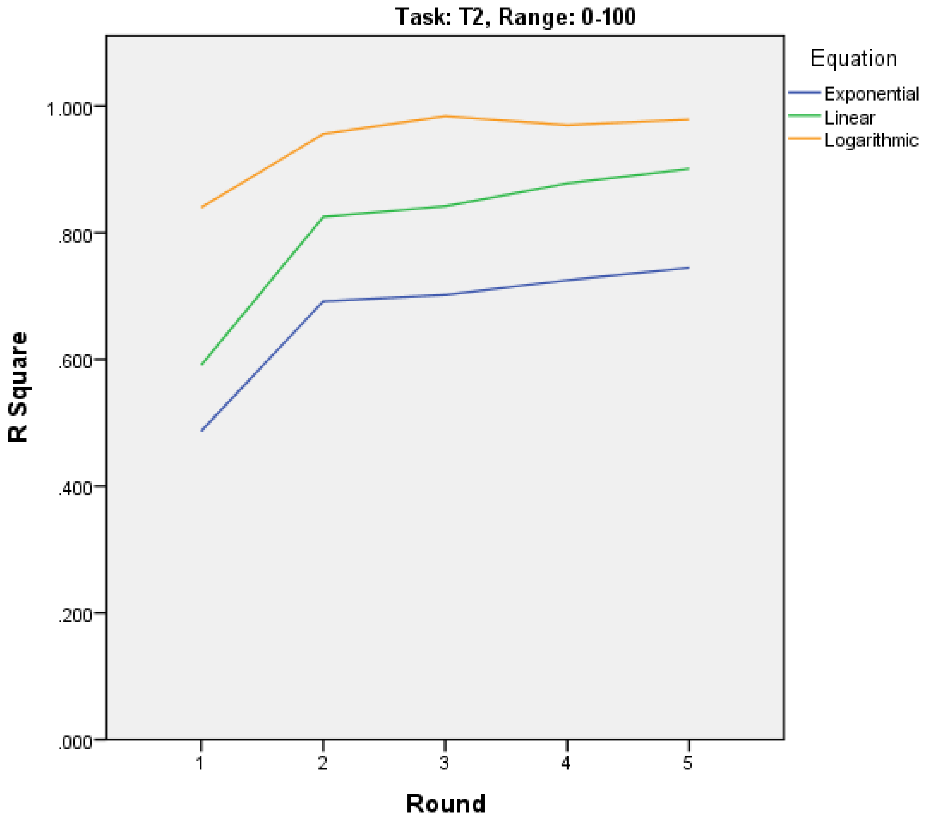


Figure 35

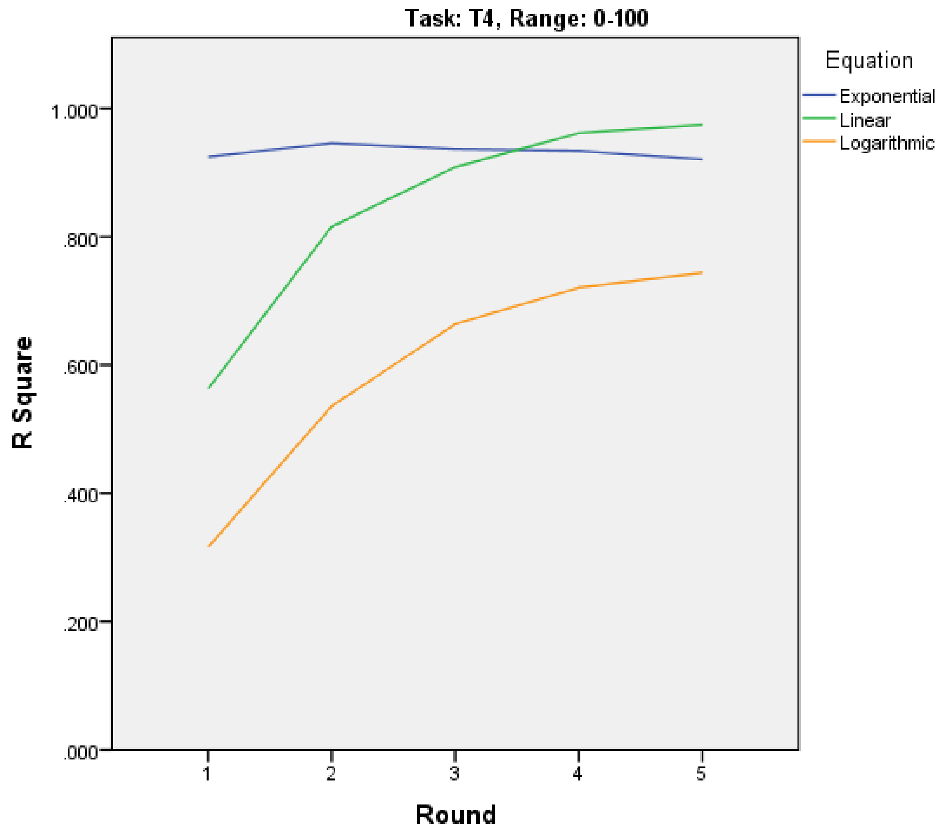


Figure 36

Range	Task	R1	R2	R3	R4	R5	Compatibility with log-linear shift account?
0-10	T2	Exp	Exp	Exp	Exp	Exp	Against
	T4	Log	Log	Log	Log	Log	Against
0-20	T2	Linear	Linear	Linear	Linear	Linear	Neutral/ Compatible
	T4	Linear	Linear	Linear	Linear	Linear	Neutral/ Compatible
5-15	T2	Exp/Lin	Exp/Lin	Exp/Lin	Exp/Lin	Exp/Lin	Against
	T4	Log	---	Exp/Lin	---	---	Against
0-100	T2	Log	Log	Log	Log	Log	Compatible
	T4	Exp	Exp	Exp	Linear	Linear	Compatible

Figure 37 Best-fitting model for median estimates

As this table shows, estimations in the range 0-100 were best fit by the models predicted by the log-linear shift account: in early rounds, T2 estimations were best fit by logarithmic models and T4 estimations best fit by exponential models. In later rounds, T2 estimations remained best fit by logarithmic models, but T4 estimations were now best fit by linear models. If a 'shift' from logarithmic to linear distribution in children's internal representations of number had occurred, which is the mechanism of developmental change that the log-linear account specifies, then it is unexpected that estimations in T4 should have reflected such a 'shift' (between R3 and R4) whilst estimations in T2 did not. Aside from this point, however, children's estimations in the range 0-100 were compatible with a log-linear shift account.

Children's estimations in the range 0-20 were also compatible with a log-linear shift account of number line estimation. No 'shift' is suggested – children's median estimates were best fit by a linear model in each round – but a log-linear shift account of course allows that the 'shift' for these children, and this number range, could have occurred before this research was carried out.

For estimations in the range 0-10, the median estimates followed the opposite pattern to that predicted, with T2 estimations (see Figure 29) well described by exponential models and T4 estimations well described by logarithmic models (see Figure 30). Children's estimations in the range 5-15 did not reflect the predictions of a log-linear shift account either.

The above exploration of children's median estimates is interesting from the point of view of comparison to previous studies. However, it is also interesting to investigate how the changes apparent in the median estimates relate to changes in individuals' estimates. The model-fitting procedure described above was therefore repeated for individual children. Children's estimations in the range 0-100 were of particular interest, since it was for this range that children's median estimates best supported a log-linear shift account. The tables below therefore summarise the results of the model-fitting analysis for range 0-100 only, indicating the best-fitting model for each child's 0-100 estimates in each round and task.

T2	R1	R2	R3	R4	R5
Patrick	Lin	Lin	Lin	Lin	Lin
Amy	Log	Log	Log	Log	Log
Ellen	Log	Lin	Lin	Log	Log
Zoe	Log	Lin	Lin	Log	Lin
Harry	---	Log	Log	Log	Log
Matthew	Log	Log	Log	Lin	Log
Catharina	Log	Log	Log	Log	Log
Lewis	Log	Log	Log	---	Log
Marta	Log	Log	Log	Log	Log
Beatrice	Log	Log	Log	Log	Log
Chris	Log	Log	Log	Log	Log
Jonah	Log	Lin	Log	Lin	Log

Figure 38 Children's best-fitting model for range 0-100, T2

T4	R1	R2	R3	R4	R5
Patrick	Lin	Lin	Lin	Lin	Lin
Amy	Exp	Lin	Exp	Exp	Exp
Ellen	Exp	Exp	Exp	Lin	Lin
Zoe	Exp	Exp	Exp	Lin	Lin
Harry	Exp	---	Exp	Exp	Exp
Matthew	Log	Exp	---	Exp	Exp
Catharina	Lin	Exp	Exp	Exp	Exp
Lewis	Exp	Exp	Exp	Exp	Exp
Marta	Lin	---	Exp	Exp	Exp
Beatrice	Exp	Exp	Lin	Exp	Lin
Chris	---	---	---	---	Lin
Jonah	Lin	Exp	---	Lin	Lin

Figure 39 Children's best-fitting model for range 0-100, T4

These tables provide limited support for the idea of log-linear shifts in individuals' estimations. For T2, seven children's estimations were compatible with a log-linear shift in that they were best fit by either a logarithmic or linear model (with no change) throughout the research, but other children made estimations well fit by linear models and then, in a later round, estimations more closely resembling a logarithmic distribution. For T4, four children's estimations either remained best fit by exponential or linear models, or demonstrated an exponential-to-linear shift, but again other children's estimations did not themselves reflect the expectations of the log-linear shift account.

6.6.2 What is the variability of children's numerical magnitude representations in estimation tasks at different times?

There are several ways to approach assessing the variability of children's numerical magnitude representations. One simple measure is the standard deviation of estimation error (PAE) in different task conditions at different times. The two following graphs, one for T2 and one for T4, show the standard deviation of PAE for all children's estimates in each task condition and round.

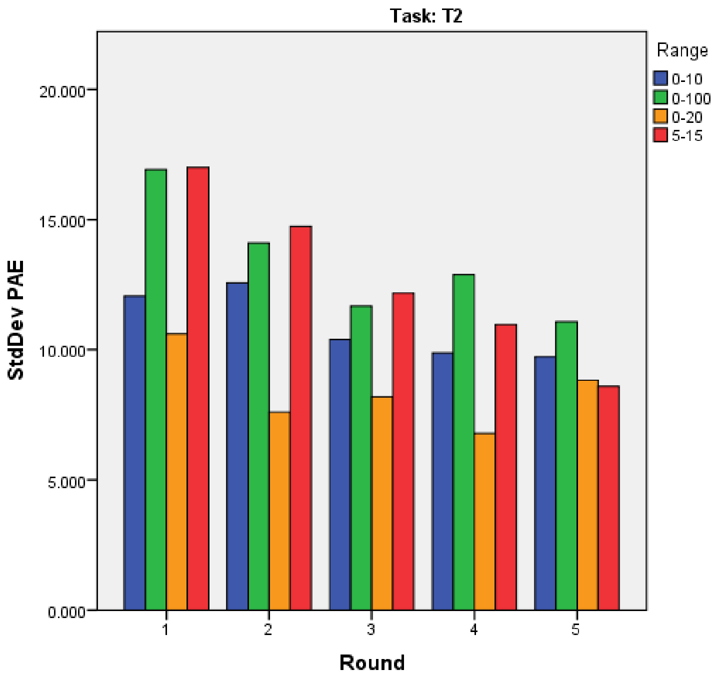


Figure 40 Standard deviation of PAE for T2 estimates by round

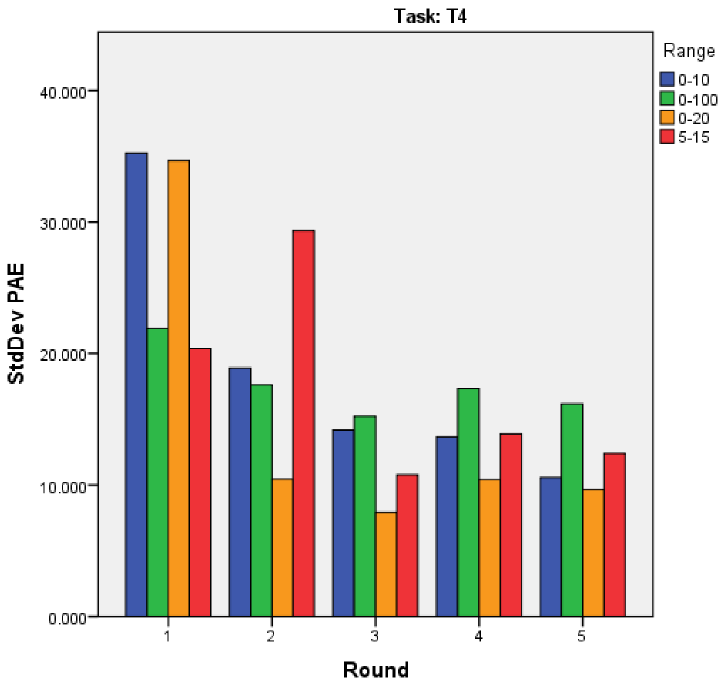


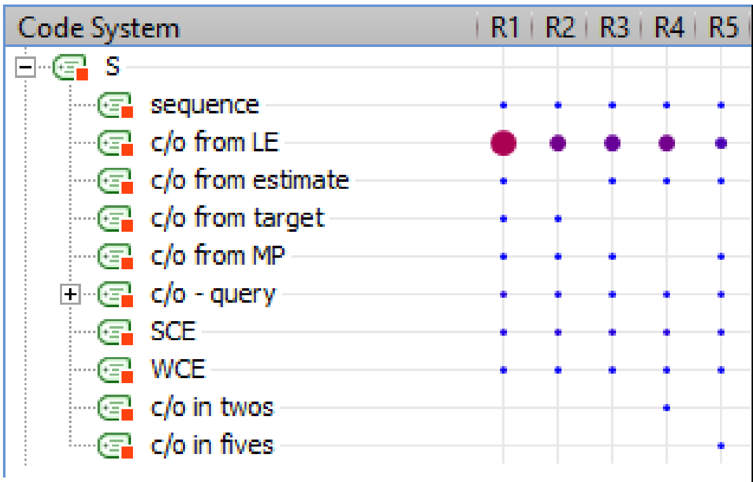
Figure 41 Standard deviation of PAE for T4 estimates by round

For both T2 and T4, the standard deviation of PAE decreased with each interview round. The standard deviation of estimates was overall higher in T2, but decreased more for T4 and by R5 was similar for both tasks. In both tasks, estimates in the range 0-20 showed the lowest standard deviation of PAE, meaning that estimates in the ranges 0-20 were both the most linearly accurate and most consistent.

6.6.3 Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?

At the level of the group findings, the main pattern of change, noted in RQ1c and RQ3, is that children’s number-line estimates more closely resembled a linear distribution and showed less variability with time. The median estimates of the group, explored in RQ3a, did not point towards a sudden ‘shift’ from logarithmic to linear representation of numbers on the number line, and overall the evidence for the log-linear shift account is limited. Discussion of how individual children’s estimates became more linear, and what possible factors might be in the change, involves a much closer examination of children’s responses and so is considered in the three case studies which follow this chapter.

Children’s representations in T1 did not evidence an overall trend, though the nature of the task as a ‘snapshot’ of representational ability means that this is not surprising. The representations that children made in each interview overall are summarised in a series of charts following, with representations grouped according to the five main stages of Resnick (1983)’s developmental theory of number understanding: sequence (S), decreasing sequence (DS), relative numerosity (RN), part-whole (PW), and part-whole base-ten (PW10). Throughout all five interviews, by far the most common structure represented was the target number in relation to one of the number line’s endpoints, a representation categorised within relative numerosity (RN). In order to focus on changes *within* each category of number structure, the following graphical depictions of representation frequency compare the frequencies only within categories, not between. Within each chart, the circles are coloured and sized to represent frequencies relative to the other data points within that chart.



Increasing sequence structure was represented in every interview. Children’s representation of count-on from the left endpoint decreased in frequency with each round, whereas count sequences in twos or in fives appeared only in later rounds.

Figure 42 Frequency of representation of increasing sequence structure (S) by round

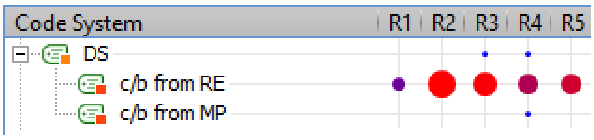


Figure 43 Frequency of representation of decreasing sequence structure (DS) by round

Decreasing sequence structure was also represented in each round. There were few such representations in R1, and a large increase in R2, but after this the frequency remained comparatively stable.

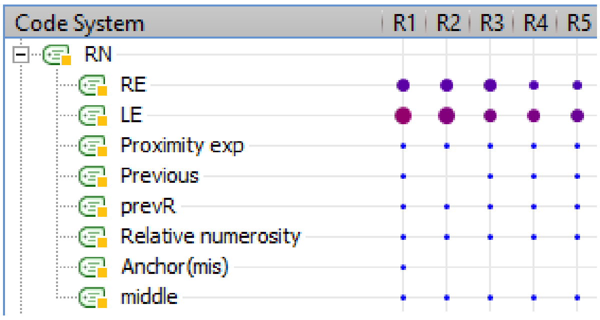
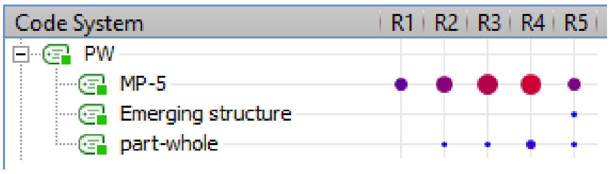


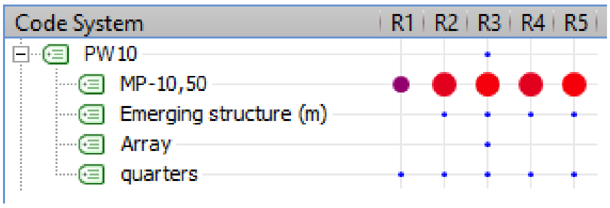
Figure 44 Frequency of representation of relative numerosity (RN) by round

Relative numerosity was also represented in each round. References to the right and left endpoints slowly decreased in frequency with interview round. The frequency of references to 'middle' (less specific than 'midpoint') and to previous trials remained stable.



Representation of part-whole structure increased with each round until R4. In R5, the frequency decreased.

Figure 45 Frequency of representation of part-whole structure (PW) by round



Representation of part-whole structure including base-ten increased with each round. Representation of quarters/three-quarters was infrequent in all rounds.

Figure 46 Frequency of representation of part-whole (including base-10) structure (PW-10) by round

Two trends are demonstrated by the above observations. Firstly, representations within each category of Resnick's framework of number structure were observed in each interview round. Secondly, representations considered less sophisticated (increasing sequence structure, particularly in the form of count-on from left endpoint) decreased in frequency, whilst representations of the more sophisticated number structure (part-whole structure, including base-ten structure) increased in frequency.

This short chapter has presented the first stage of research findings, and drawn out those findings from the group level that have contributed to answering the research questions. As noted at the start of this chapter, the comprehensive treatment of the research questions is enabled by examining children's interview responses at a much smaller grain size. The following chapters present this second stage of the findings, in the form of three contrasting case studies.

Chapter 7 Case study: Patrick

Patrick was selected as one of three case studies because, in the terms of this research, his responses to the task-based interviews were the most mathematically sophisticated: he represented comparatively sophisticated aspects of number structure with greater frequency than other children, and his number line estimations (both number to position and position to number) were highly linearly accurate.

7.1 What does this case say in response to important ideas in the literature?

This case provides strong qualitative evidence for the use of lower and upper endpoints, midpoints, and quarter-points as “anchors” during number line estimations. This supports the broad hypothesis that children use anchors or landmarks – both given (endpoints) and inferred – in some way during their estimations. Such a hypothesis has been advanced in various forms by Siegler and Opfer (2003), Barth and Paladino (2011), White and Szucs (2012), Slusser et al. (2013) and Rouder and Geary (2014). Such a result is not surprising, but with the exception of Sullivan et al. (2011) (who studied only adults’ number line estimations) no recent studies have examined whether behaviour actually supports the hypothesis.

Patrick’s case furthers the debate on whether this estimation behaviour can be inferred from quantitative estimation data. In the studies noted above, the use of anchor points was inferred from patterns of estimation error. In Patrick’s case, highly linear estimates were achieved using a variety of strategies, and frequent cognitive representations of number structure including part-whole structure (a more sophisticated structure). Where estimation data strongly suggested use of a particular anchor point, there was sometimes no qualitative data to support this. Conversely, estimations did not always have the linear accuracy that could be expected from examination of the strategies used.

The linear accuracy of Patrick’s estimations increased somewhat over the school year; however, his estimations were highly linear on all ranges tested even at the start of the year, so the overall increase in linearity was not statistically significant. In the range 5-15, a ‘difficult’ range due to the non-zero starting point, there was a significant increase in the linearity of estimations over the year.

Patrick’s case provides some support for a proportional reasoning account of number line estimation. The best evidence in support of this would be the pattern of over-estimation and under-estimation demonstrated on power-cycle models, and there is some suggestion that this pattern occurred (particularly for the range 0-100). However, important methodological limitations mean that this research is not optimally designed to test the plausibility of the proportional reasoning model: the methods were designed to investigate the log-linear account (see RQ3a and Slusser et al., 2013). The deviation from linearity of Patrick's estimates did not always (or even frequently) follow the distribution predicted by existing accounts.

The data are strongly compatible with the idea found in multiple accounts (Ashcraft and Moore, 2012; Slusser et al., 2013) that increased linear accuracy in number line estimation arises from the incorporation of increasing numbers of anchor points: "The origin of the line is always a region of highly accurate estimates regardless of age and underlying representation (i.e., linear or logarithmic). This point is then joined by accurate estimates at the endpoint of the line and, with increasing knowledge of arithmetic, by the midpoint of the line." (Ashcraft and Moore, 2012, p. 265). I agree with Ashcraft and Moore's further assertion that "this strategy involves not just the perceptual salience of the midpoint of the number line but also the arithmetic knowledge that a hatch mark near the perceptual midpoint must equal a value near 50 given that 50 is arithmetically half the length of the 0–100 number line." (Ashcraft and Moore, 2012, p. 260) However, I do not think that Patrick's case alone is able to offer evidence in support or refutation of this claim, and indeed I struggle to imagine the evidence that would be suitable.

The stage-by-stage inclusion of anchor points is found in overall conflicting accounts, with Ashcraft and Moore (2012) concluding that their data overall better support Siegler et al's view of number line estimation as reflecting the mental representation of number (with stage-by-stage incorporation of anchor points), whilst Slusser et al. (2013) conclude that their data support the view of number line estimation as proportional reasoning. In this second account, the stage-by-stage incorporation of anchor points is one of two factors leading to improved estimation accuracy: the inclusion of further anchor points changes the mathematical model from unbounded, to one-cycle, to two-cycle power function, and a parameter β present in each of these models (see Slusser et al., 2013) indexes an overall bias, which with age and experience approaches 1 ("perfect" estimation). Patrick's case study does not address which of these overarching accounts is the more plausible.

The data strongly support the suggestion of White and Szucs (2012) that "specific numbers could exhibit unique behaviours as a function of the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy" (White and Szucs, 2012, p. 9). Patrick's case suggests strong links between a target's position within the estimation range, the size of the estimation range, and the strategies used.

Finally, this data supports Slusser et al. (2013) in drawing attention to the spatial components of the task itself: "typical analyses of these tasks attribute variations in number-line estimation solely to numerical processing and numerical representations, assuming that the spatial components of the task do not introduce their own variations. This assumption is deeply problematic given a wealth of research on the estimation of spatial position in children and adults" (Slusser et al., 2013, p. 195). Qualitative evidence from Patrick illustrates these difficulties with episodes in which intentions (for example a verbalised comment that "I'll just put it halfway", or an attempt to appropriately scale counting on a particular number line) do not match spatial actions.

In the remainder of this chapter, I present the findings from Patrick's case study in relation to each of the research questions.

7.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research?

7.2.1 What are the modes and component signs used in the representations?

During Task 1 of each round, Patrick produced a drawing on paper, to which he added spoken and gestural explanation supported by gaze direction. During Tasks 2 and 4, Patrick's only written representations were the numerals written into the rockets to answer Task 4 trials. Unlike some other children, Patrick made no additional inscriptions – neither in solving the task for himself, nor seeking to explain to the researcher. Representations of numbers verbally (aloud), in gesture and with gaze were however very frequent during T2 and T4, occurring during almost every trial. During Task 3, Patrick demonstrated few cognitive representations of number, and those that occurred were verbal only.

The component signs of Patrick's drawn T1 representations of number, shown in the following section, were primarily notational. The numerals Patrick used were consistently boxed in each drawing, and in R4 and R5 the squares feature alone, as iconic components. Patrick's verbal and gestural representations indicated that these squares functioned as placeholders for other symbolically-written numbers.

Patrick's representations of number very occasionally included pictorial components, as when he described the possible lengths of number lines in terms of being able to reach to the sky and back. This particular example occurred during R4, a spontaneous comment during T4:

Patrick: … draw a straight line … lines can be any size you want [*RH pencil traces short vertical line mid-air in front of body*].

Researcher, J: Yes, it's true.

P: Cos they're just straight lines. Lines they can be … up to the sky and back again! [*RH reaches full extent to the ceiling, waves*].

7.2.1.1 Structure of T1 representations

In R1, Patrick's drawn representation of the counting numbers featured strong linear structure, but no other structure. The number structure represented (in Resnick's terms) was sequence only.

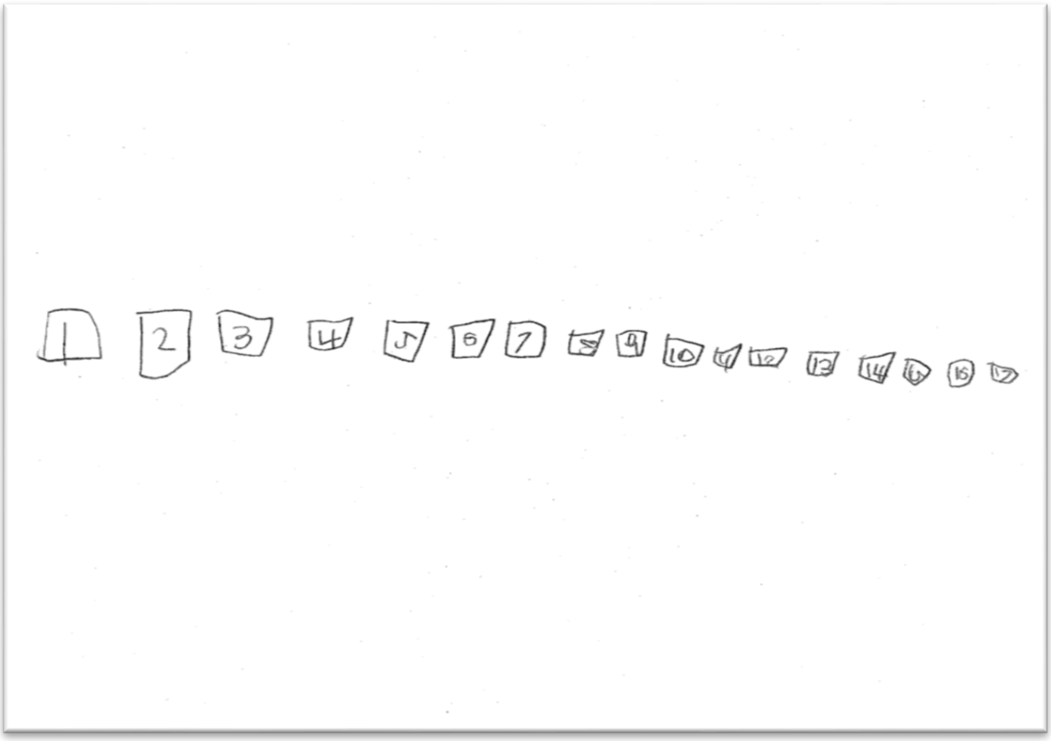


Figure 47 Patrick R1 T1

After drawing, Patrick made a spontaneous comment that was followed up by the researcher:

P: I haven't got space for the rest. [*resting*]

J: So can you tell me about where they'd go if you had all the paper in the world?

P: Well it'd just keep on going in a line. [*RH, holding pencil, sweeps L->R across drawn boxes, continues sweep off the page rightwards, until R arm extended mid-air full stretch to R]*

J: Mmhmm?

P: Until I get to a hundred in boxes. [*resting*]

In R2, Patrick began with a verbal and gestural representation:

P: Sort of like squares… going across. [*right hand sweeps pencil across page L->R, twice. Second time, continues sweep rightward to mid-air, arm reaching to R]*

The researcher then asked if he could show how this might look using pencil and paper, and Patrick drew a very similar sequence of boxed numerals to that in R1, but this time stopped at 13.

P: Written like that.

J: Mmhmm?

P: To… a hundred, like in a line. [*RH sweeps line and continues rightward to mid-air again, arm reaching R]*

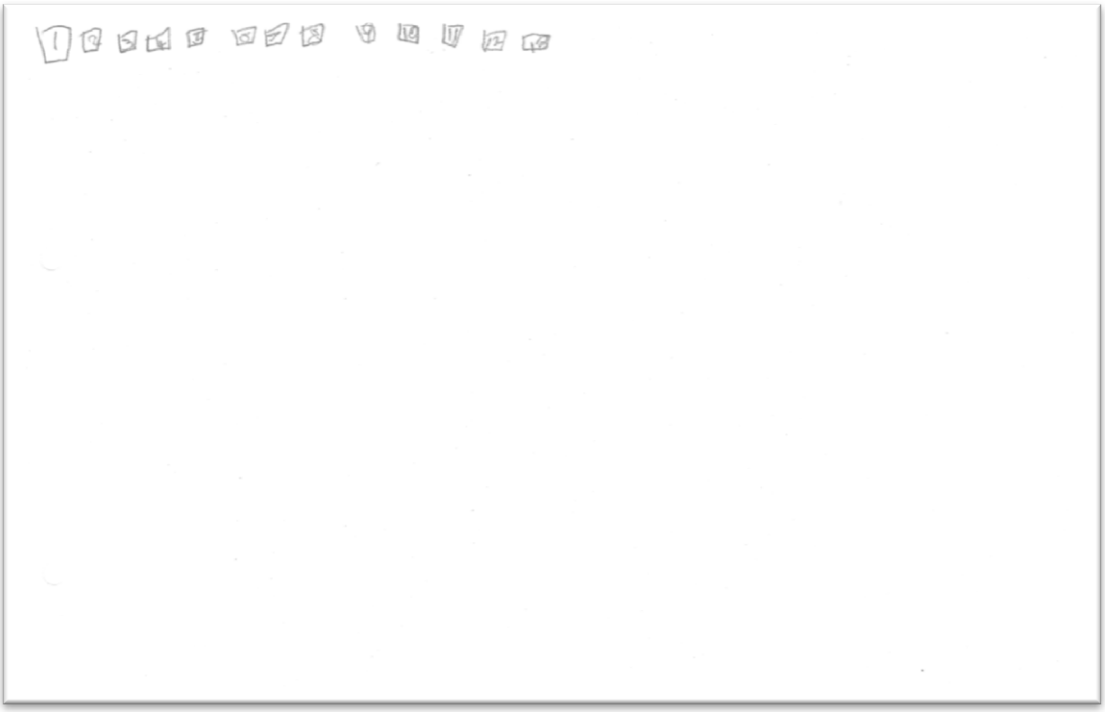


Figure 48 Patrick R2 T1

As in R1, this imagistic representation is classified as having linear structure only.

In R3, Patrick drew and verbally described at the same time. He then continued the verbal description and accompanied it with gesture.

P: Sort of like … got squares. Now it's a bit different, it's going down. [*drawing squares in first column*]

J: Right.

P: Cos it's got bigger number. [*drawing still, then pushes drawing towards J, for her to see*]

P: And it's from like that going down and down and down in lines. [*RH pencil sweeps down first column of drawing, then in further downward sweeps parallel – indicating columns 2, 3, 4 on each "down"*]

P: So if the ones are on that line, all of the twos are on that line, all of the threes are on that line, the fours are on that line… All the fives are on that line, and all of the sixes are on that line. [*RH pencil sweeps L->R rows, beginning with 1, 11 row and then making parallel rows beneath*] …

P: Adding … just like … all of the ones on that side [*RH pencil sweeps L->R from the "1", for 10cm*]

J: Yes

P: And then it goes down to 21, 31 …

J: Ah, right.

P: And so on.

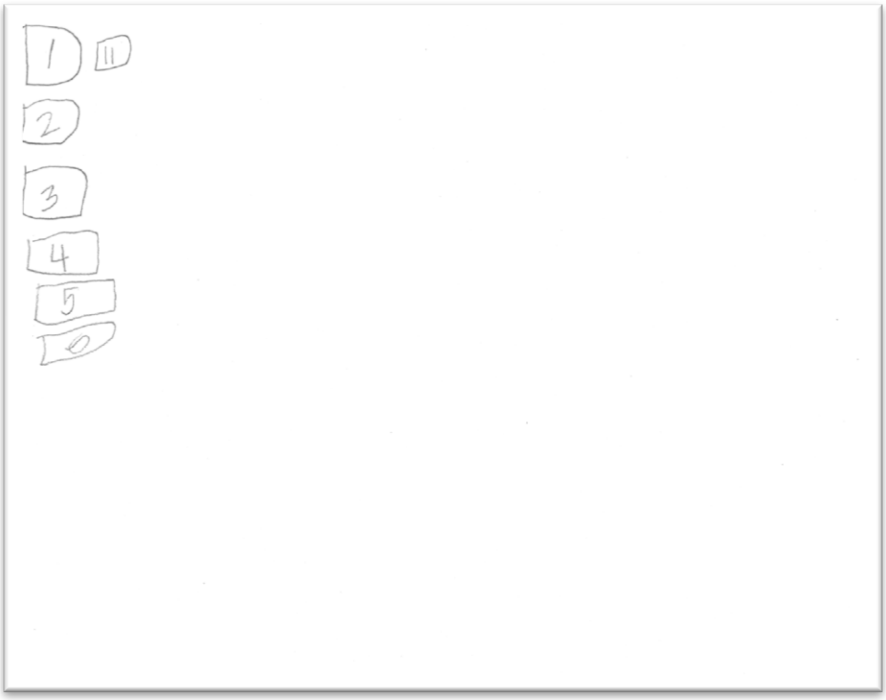


Figure 49 Patrick R3 T1

This drawing, together with its verbal and gestural description, demonstrates array structure, specifically with elements in a 10x10 array, the highest level of structure identified by Thomas and Mulligan (1995). The structure was not simply apparent in Patrick's representation; Patrick chose to describe the representation to the researcher in these terms, foregrounding the structural features; first in terms of the columns (though Patrick did not name them "columns") and then in terms of rows. Thomas and Mulligan (1995) include examples of array structures that seem to have been reproduced holistically, as remembered images, or in Duval's terms *mental images*), by children from their classroom experience of conventional number grids. This does not seem to be the case for Patrick; he has some difficulty articulating the number patterns within the array, but does manage to convey its structure and the fact that that structure is the key feature of this representation. It is not clear from this episode the extent to which Patrick associates with the representation a structure of related numerical magnitudes, in addition to the structure of the pattern in the numerals.

In R4, Patrick began T1 with verbal and gestural description alone:

P: Well … it's like, starting from one in the top corner [*LH points to upper-left corner of room*] – 2, 3, 4, 5, 6, 7, 8 [*pointing jabs in mid-air, descending in diagonal to lower-right corner of room*]
P: and it goes to about … 25 down there [*body turned to R, emphatic point to floor on the right on "there"*] …
P: and another 25 to 50 [*rapid RH point to room upper-right corner, then arm sweeps diagonally down leftward to lower-left corner*] then another 25 to 75 [*L arm reaches forward and left, hand flat, body leans leftward*], and another 25 equals a hundred [*R arm reaches forward and right, hand flat, body leans rightward*]. So it's quite strange.
J: Yes! Could you draw that at all on paper? Roughly – to show me where it goes?

At this point, Patrick drew the following image:

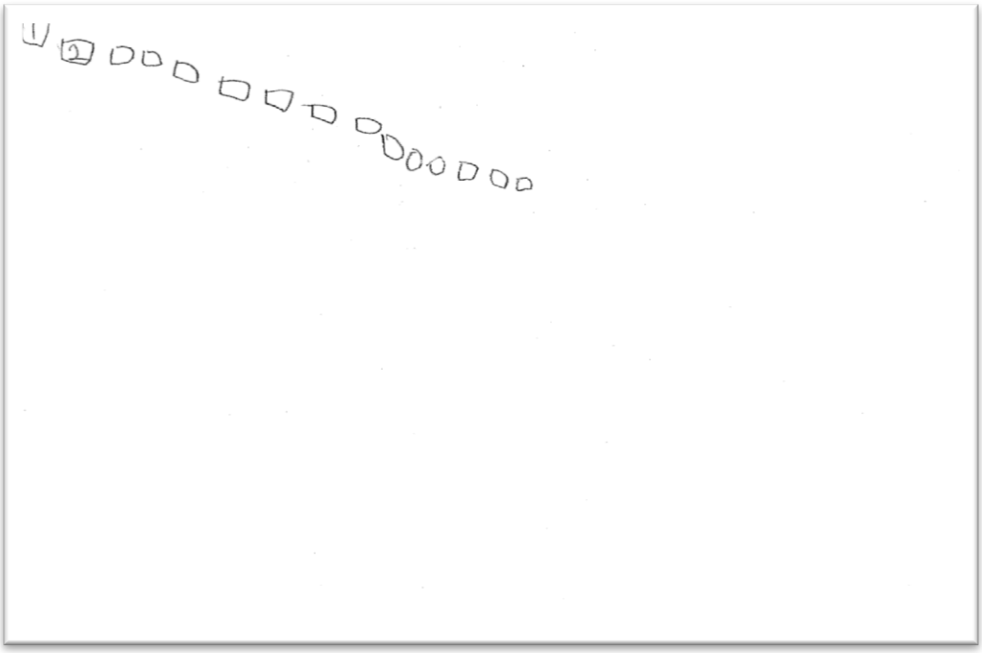


Figure 50 Patrick R4 T1

Patrick continued to verbally explain:

P: From the rest of that it just sails down there like thaaaat [*RH pencil traces page upper-left corner to lower-right corner, then repeats the path faster*] … and then comes back over there [*RH pencil sweeps rapidly from page upper-right corner to lower-left corner*] and then down that bit [*pencil placed vertically at page left edge*] forms two crosses … [*waves RH mid-air briefly, then pencil placed vertically at page left edge again*] then it's just go the numbers. The numbers that are left from the others [*RH pencil mid-air sweeps page upper-left to page lower-right – the path of the drawn boxes]*. And that's like, because the high numbers that were left out of the others.

In the classification of structure by Thomas and Mulligan (1995), this imagistic representation is considered to have emerging multiplicative structure: although the structure visualised by Patrick is not entirely clear, the sequence of numbers is grouped into four equal sections of twenty-five that Patrick explicitly notes add up to one hundred.

In R5, Patrick verbally and gesturally described his imagistic representation at the same time as drawing:

P: Right… it's a bit like … going in order on screen [*drawing boxes from upper-left*], when it gets to there [*traces from drawing to page lower-right corner*] there's a … [*draws box in lower-right corner*], then it goes … goes up there [*RH pencil traces lower-right corner of page to upper-right corner*]
P: then down there [*traces upper-right corner of page to lower-left corner*], then back to where it was [*RH pencil points to page upper-left corner, where drawing began*], and that would be number a hundred there [*points again to page upper-left corner*] …
P: Down, up, to there, down to there, back up to there [*traces above page: upper-left corner to lower-right corner, to upper-right corner, diagonally to lower-left corner, back to upper-left*].
J: OK. Is that always how it looks to you?
P: Yep. Sometimes, sometimes not. Sometimes it looks like that, sometimes [*shrugs*] still don't really know which one I have the most.

Figure 51 Patrick R5 T1

This representation, like that in R4, is classified as showing emerging structure; the linear sequence is broken into approximately equal segments, although this time the numerical structure to the groups is not clear.

For all five rounds, Patrick's gestures and speech contained dynamic elements such as sweeping movements and phrases like "it's going down". However, in each round, the overall impression was that the imagistic representation of numbers Patrick had in mind was itself static, and that the dynamic element came from his care in communication. Patrick appeared to want to guide the researcher through the features of the number representation, and this process took the form of a dynamic journey through his representation.

7.2.2 What aspects of number structure are represented?

In Patrick's case, all stages of number structure understanding in Resnick's developmental theory were represented at some point during the five rounds of interviews. The following table shows the stages of number structure in Resnick's theory (first column), followed by the elements that make up that stage (second column). The third column provides examples of Patrick's procedures and strategies that include representation of this particular aspect of structure.

Number structure	**Detailed stages**	**Procedures or strategies in which this number structure may be represented**
S Sequence	Count sequence from 0 or 1	Count-on from LE.
	Count sequence from other start point	• Count-on from non-zero LE. • Count-on from estimate to RE for confirmation.
DS Decreasing sequence	Count sequence reversed	Count-back from RE.
RN Relative numerosity (quantity comparison)	Ability to represent (esp visualise) number's magnitude without counting up to it	• Applying efficient count strategy, using RN to choose shorter available count. • Checking judgements, querying whether an estimate is likely. • Use of endpoints as "anchor" points.
PW Part-whole	Partition and recombine numbers (<20)	Use of "anchor" points based on partitioning the range.
	Use of number bonds to ten	• Checking estimations using number bonds to ten. • Using own midpoint anchor on 0-10 range.
	Multiple partitions of multi-digit numbers	Using own midpoint anchors on 0-20 and 5-15 ranges.

PW10 Part-whole understanding of base ten structure	Numbers as compositions of tens and units	• Using own midpoint anchor on 0-100 range. • Count-on and count-back by tens. • 10x10 array structure of 0-100
	Multiple partitions of larger multi-digit numbers	• Count-on or count-back by any given number. • Partitioning into quarters, thirds, other.

Figure 52 Aspects of number structure represented by Patrick

Three examples here are used to show the kinds of episodes from Patrick's interviews that were inferred to represent different aspects of number structure from Resnick's theory. The first episode, taken from Patrick's R1 interview, provides examples of representations of increasing sequence, decreasing sequence, and relative numerosity. The episode occurred during Task 2, range 0-20, for target number 19.

J: A nineteen. [*holding out a "19" rocket sticker*]

P: [*eyes to RE; both hands take rocket to RE, then inch left-wards a little*]

P: I think ... it goes there ... [*sticks rocket*]

J: Yes?

P: Because it's really close to the twenty. [*right hand traces from rocket-> to RE; left hand then points to rocket*]

J: Good reason.

P: [quietly] nineteen twenty. [*looking at J; left hand remains on rocket*]

The interpretation of this episode is that Patrick focused upon the proximity of 19 and 20 in the count sequence to position the rocket immediately, using the anchor of the right endpoint (20). From there, he moved the rocket a small amount leftwards, corresponding to a small decrease in numerosity from 20. Patrick then explained or justified the response to the researcher, both verbally and with gesture highlighting the spatial proximity of 19 and 20 in this line representation, and for good measure speaking aloud the relevant short segment of the count sequence.

The two following examples are taken from T2, range 0-100. The first episode occurred during Patrick's R3 interview. It illustrates Patrick's representation of part-whole (partition and combine) structure, as well as a second example of relative numerosity.

J: Where would the number four rocket go?

P: [*eyes immediately to LE*] Very a lot down here. [*eyes remain LE; sticks rocket near LE immediately*]

J: Right.

P: [*eyes on rocket, then jump to RE, then to J*] Cos there's ninety-six separating them.

In this episode, Patrick's initial response (in gaze and verbally) is to use the relative numerosities of zero, four and one-hundred to focus attention upon the left endpoint of the range. Patrick makes clear that four is not merely close to zero in the count sequence, but relative to right endpoint one hundred, which the total line represents, four must in fact be "very a lot down here". After sticking, a glance to the right endpoint appears to check and confirm for Patrick that his response made good sense, and he explains to the researcher with more detail why four must be far from the right endpoint.

The final example presented in this section is from Patrick's R2 interview. This episode demonstrates Patrick's representation of part-whole (base ten) structure.

J: And twenty five.
P: [*eyes immediately to LE, then to mid-line area; takes rocket to left half of line*] …. Quarter. [*sticking rocket on left half of line*]
J: Ah?
P: It's a quarter. I'm putting it there because it's a quarter. Because half would be fifty, and then half, and then half of a hundred is 50, and then … you need four of those. [*points right hand to stuck rocket. Both hands make chopping motion at midpoint on each "half"; right hand points to stuck rocket again on "those"*]
P: Because one's twenty-five, two's fifty, three's seventy-five and four's a hundred. [*right hand points ~25%, glides (wobbly) to midpoint, glides (wobbly) to ~90%, then taps RE; eyes follow pointing finger throughout*]

7.2.3 What are the notable between-task and within-task connections between representations?

In this section, Patrick's representations are compared between tasks and within tasks, asking what varied and what was consistent.

7.2.3.1 Task 1

Patricks' T1 representations were presented in RQ1a above. In order to facilitate direct comparison, the following table summarises Patrick's representations of number during T1, according to the framework of analysis developed by Thomas and Mulligan (1995) for this task.

		Interview round				
		R1	R2	R3	R4	R5
Structure, classification from Thomas and Mulligan (1995)	No structure					
	Linear structure only	X	X			
	Emerging structure					X
	Emerging structure (multiplicative)				X	
	Partial array					
	Array: 10x10			X		
	Component signs	Notational	Notational	Notational	Iconic and notational	Iconic and notational
	Static/dynamic	Static	Static	Static	Static	static

Figure 53 Structure of Patrick's representations of number during Task 1 (Thomas & Mulligan, 1995)

In terms of the stages of number structure understanding in Resnick's (1983) account, the representations in R1 and R2 incorporated sequence structure only; the representation in R3 incorporated sequence structure and part-whole structure including base 10; and the representations in R4 and R5 incorporated sequence structure, and initial part-whole structure (partitioning and recombining, but not base 10 structure explicitly). A more general categorisation would maintain these groupings: the representations of R1 and R2 are extremely similar, the representations of R4 and R5 form another closely related pair, and the representation from R3 stands somewhat alone.

7.2.3.2 Task 2 and Task 4

In this section, representations in different task conditions of the number line estimation task are compared. The series of charts following summarise Patrick's representations in gaze, gesture and speech, and the task conditions in which the representations occurred, in each round of interviews. The representations and strategies coded for structure are grouped according to Resnick's microstage account of number structure understanding (see 4.4.2.1). The size of each circle, together with its colour, reflects the number of episodes in which a particular structure was coded: a large red circle indicates a high frequency relative to the other frequencies in the chart, and a small blue circle indicates a low frequency relative to the rest of the chart.

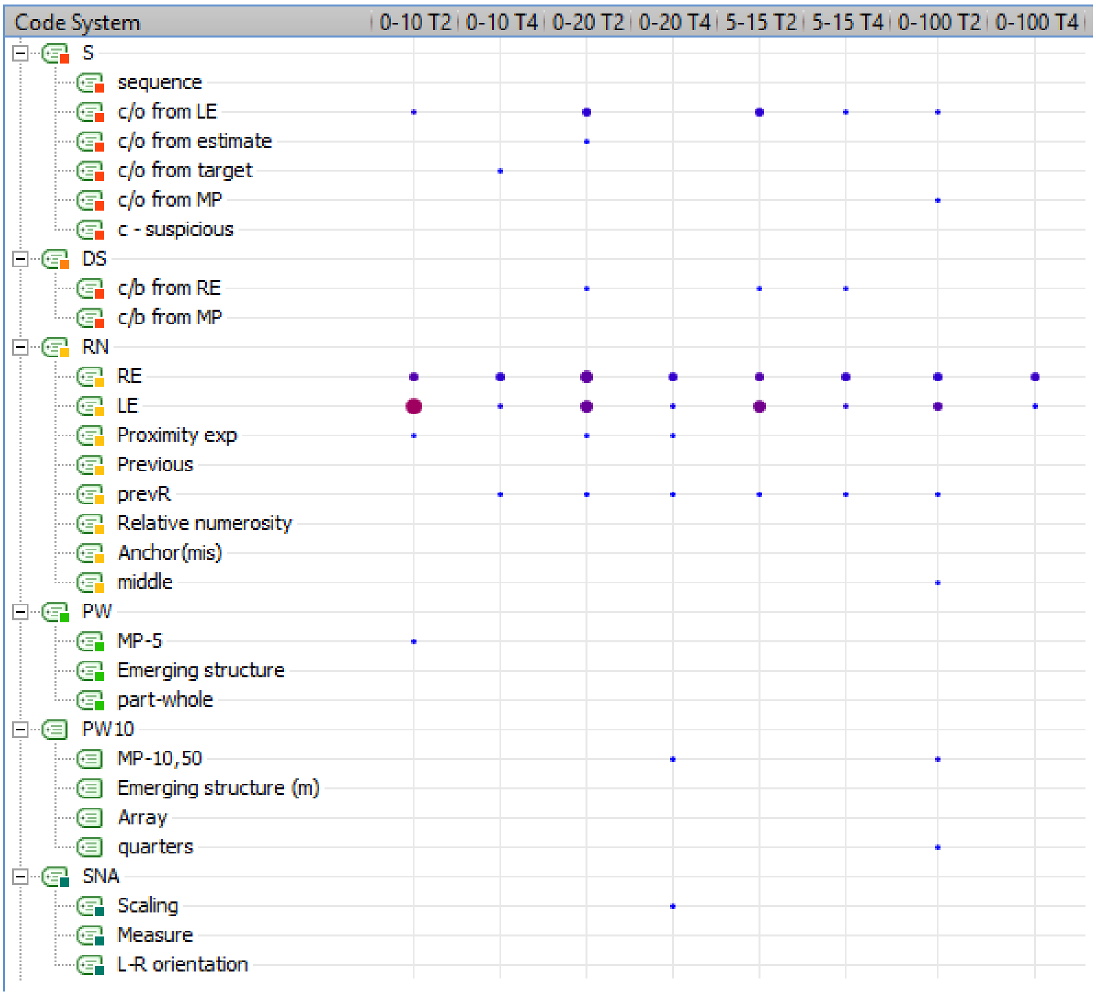


Figure 54 Round 1

In this first round, the representation of number sequence in count-on strategies was seen across all conditions except T4 0-20 and 0-100. References to both endpoints were frequent in all conditions, although less in 0-100 tasks compared to the lower ranges. Few examples of part-whole structure were represented in this round, although midpoint structure in T2 0-10, T4 0-20 and T2 0-100 were exceptions.

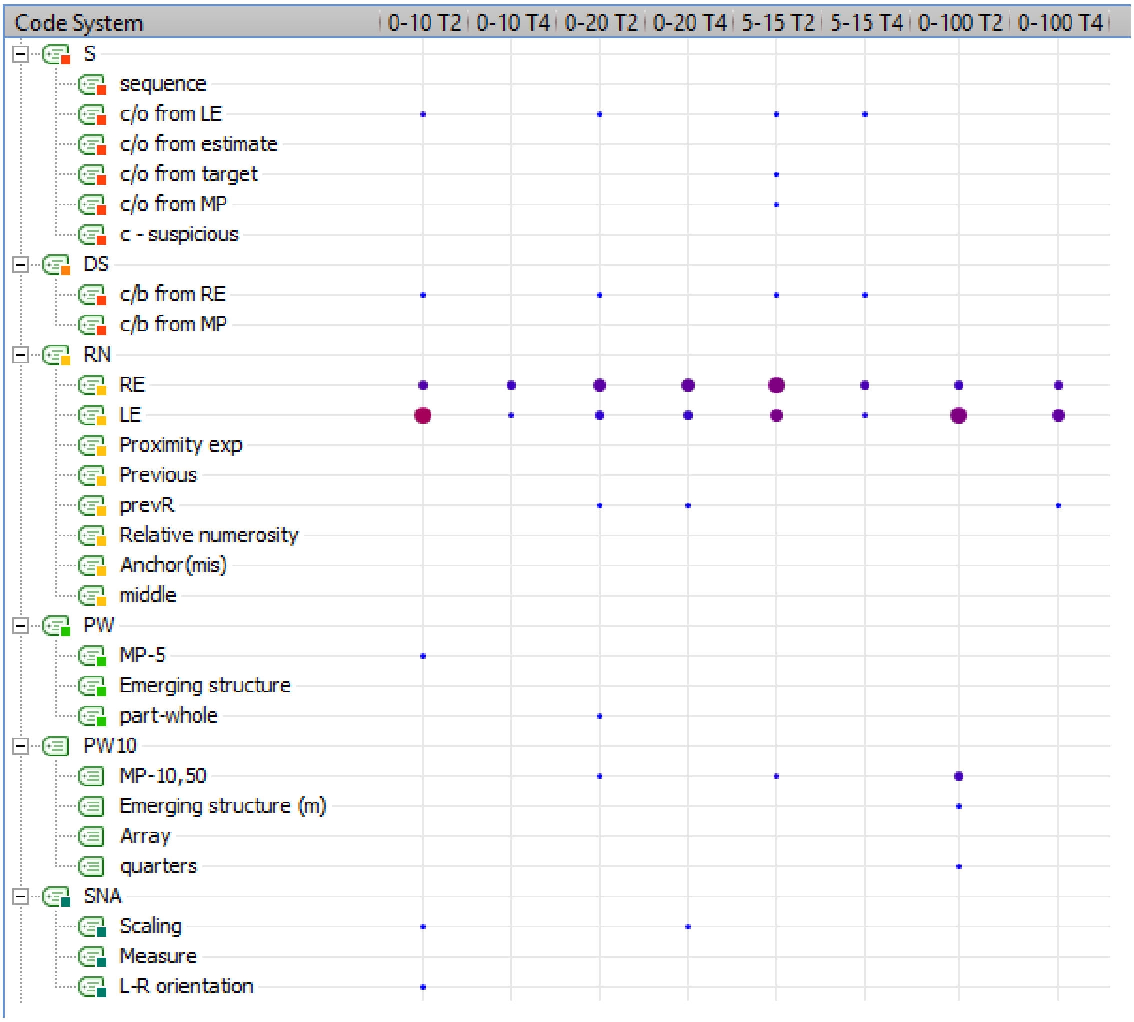


Figure 55 Round 2

This chart shows a decrease in counting strategies in R2 compared to R1. References to both endpoints occur in every task condition once again, and with higher frequency than in R1. Part-whole structure is represented more frequently, particularly the midpoint during T2 0-100 estimations.

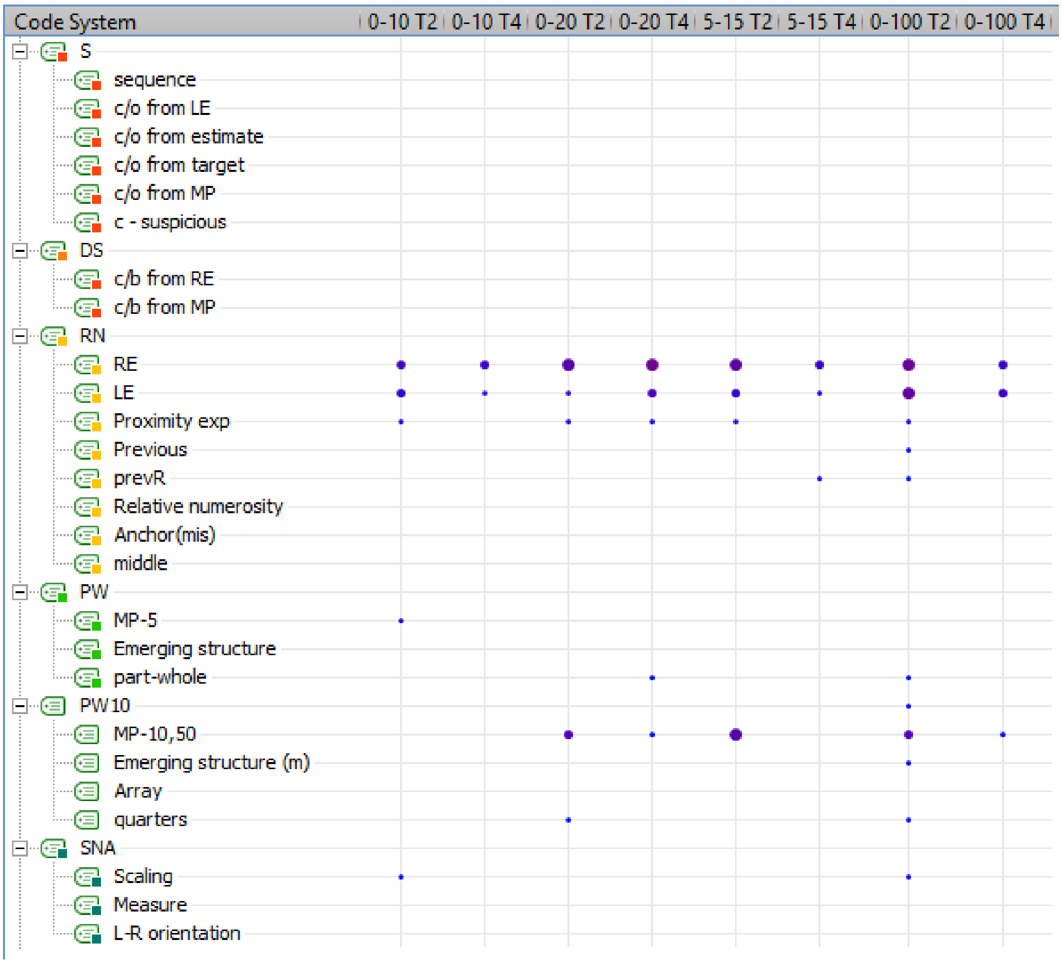


Figure 56 Round 3

In R3, references to both endpoints are again frequent in all task conditions. Compared to R2, there is more representation of part-whole structure, which occurs mainly during T2 estimates on 0-100, but also with increasing frequency in T2 5-15.

What is particularly noticeable here in R3 is that representation of sequence structure (including all count-on strategies) has stopped entirely.

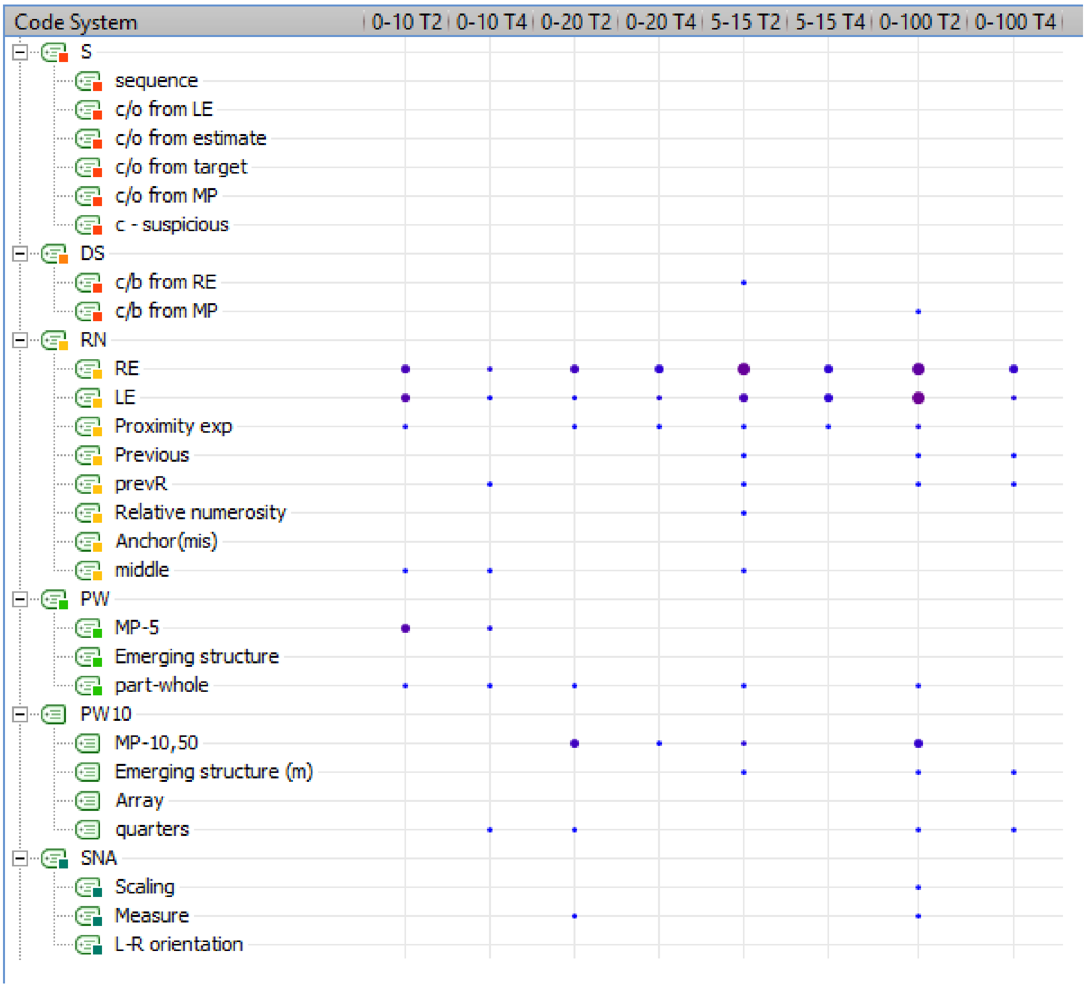


Figure 57 Round 4

In R4, references to both endpoints are again frequent in all task conditions. As in R3, there is no representation of increasing sequence, though decreasing sequence structure is represented. Compared to R3, there is again more representation of part-whole structure, and across a wider range of task conditions. Fewer representations occur during T4 compared to T2.

Representations in R5 (Figure 58, following chart) were very similar to those of R4. A difference is the lack of decreasing sequence representations, and the re-occurrence of several instances of increasing sequence. References to both endpoints are still frequent throughout. The highest frequency of part-whole representation occurs during T2 range 0-100 once again.

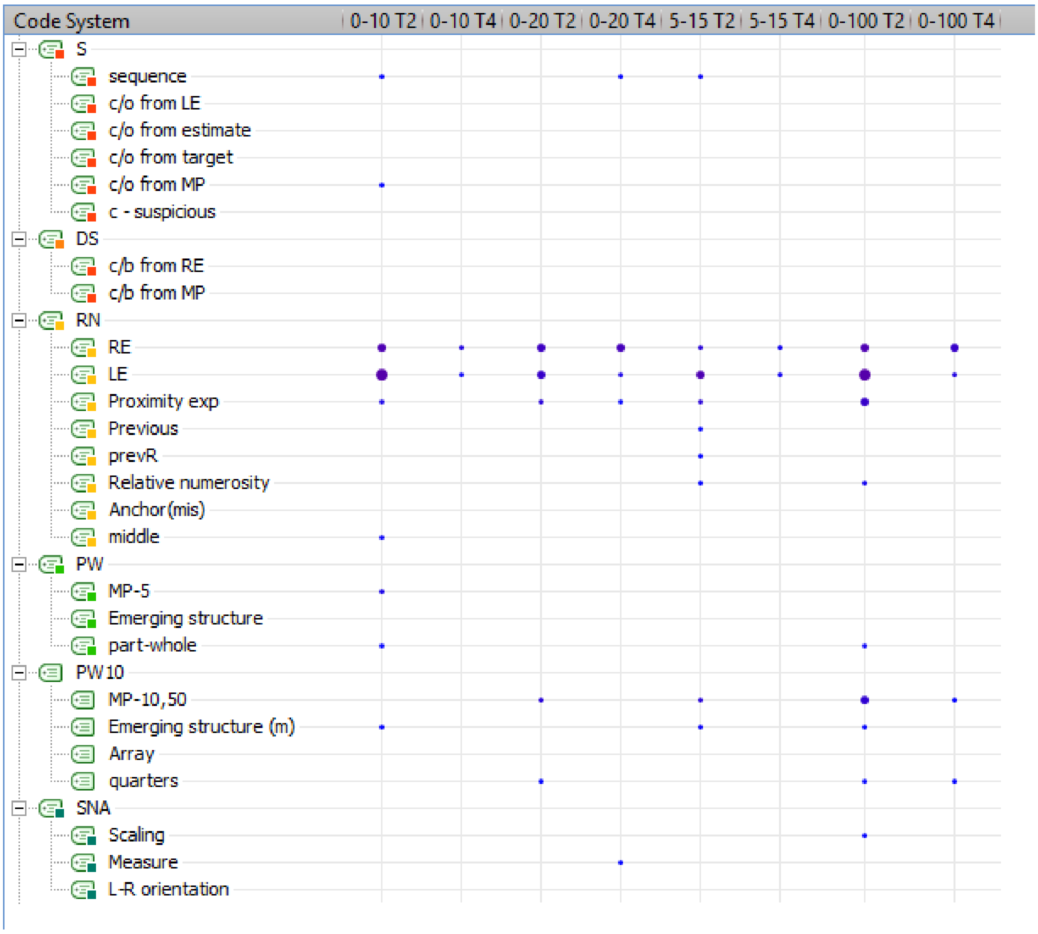


Figure 58 Round 5

One further connection between representations of number not captured by the data above is a particular representation of number sequence starting from “small” and becoming “large”, seen in many children’s interviews with strikingly similar features in each case. In the interview transcripts, this particular representation was named “Sweep to large N”, and it consisted of children making a sweeping gesture from the space in front of them (often on a page) to the right, ending with their arm fully extended to the right. Very typical examples of Patrick using this “Sweep to large N” occurred in both R1 T1 and R2 T1 (see RQ1a).

7.2.3.3 Linear accuracy of representations

Patrick’s representations of number during T2 and T4 also include, of course, his estimations themselves - the representation of target numbers within ranges that each trial requested. In each trial, the range and target were pre-chosen, but Patrick’s estimation choice or decision led to a representation of number-within-range that varied in the extent to which it coincided with a linear representation of number for that range.

In order to compare the extent to which Patrick’s representations of targets coincided with the linear representation for a given range, the percent absolute error (PAE) was calculated.

Inspection of estimation patterns revealed differences in the variability of PAE for estimations within each range. Levene's test confirmed that significant differences existed in the variability of PAE in different range conditions ($F(3,181)=4.763$, p=.003). The following boxplots show the distribution of PAE scores in each range, for T2 and T4 separately.

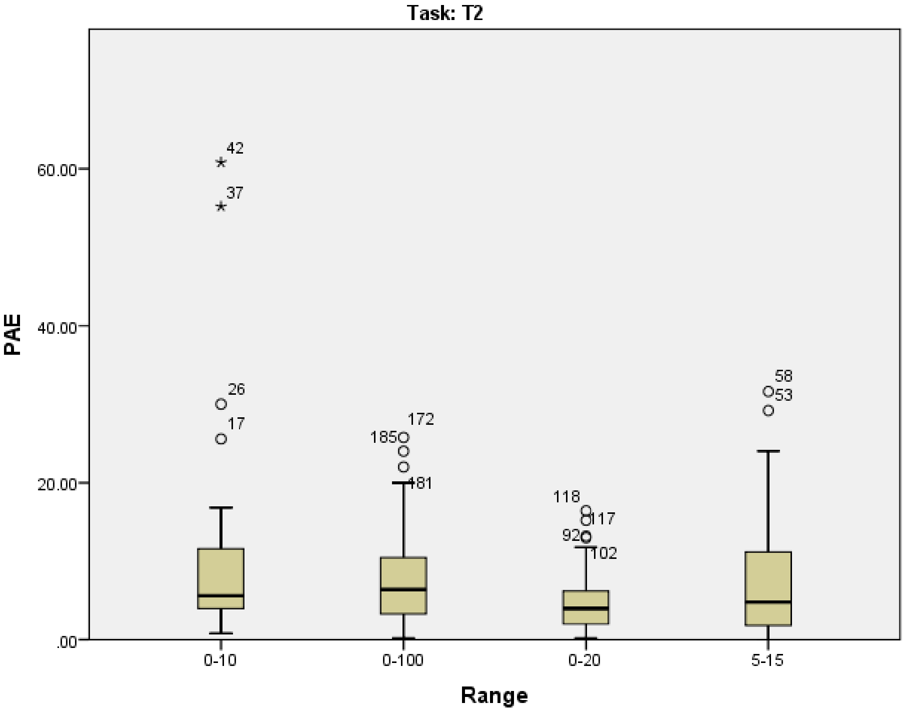


Figure 59 Percent Absolute Error of estimations by range

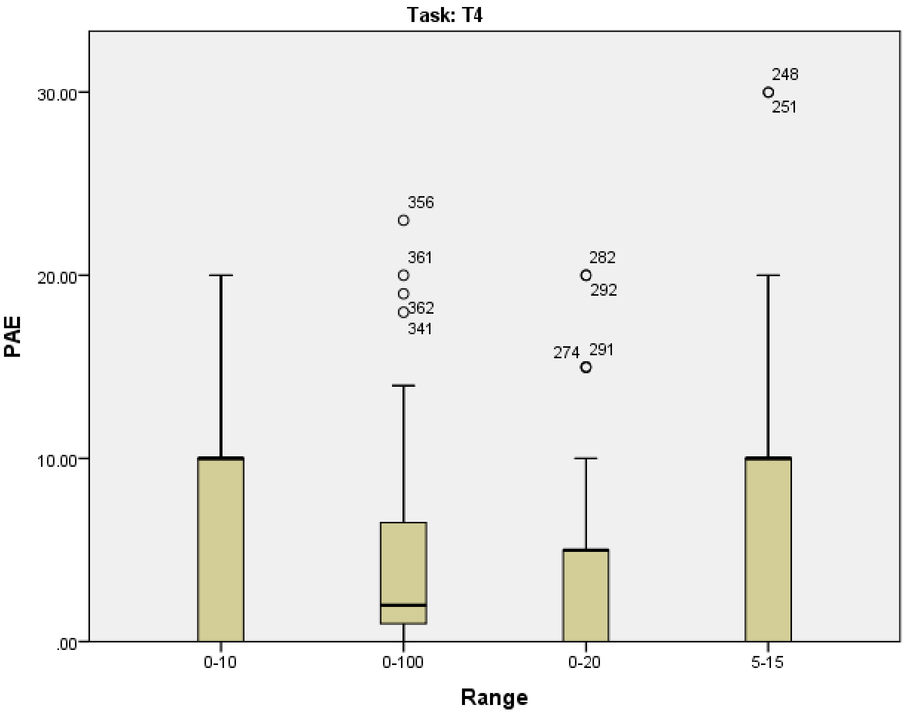


Figure 60 Percent Absolute Error of estimations by range

Significant differences between the variance in the different ranges remained after multiple data transformations, so further analysis of PAE scores was carried out treating the four range groups separately. For each range, a two way repeated measures ANOVA was carried out to investigate the PAE of estimates carried out in each of the round x task conditions. All effects are reported as significant (or non-significant) at p<.05.

7.2.3.3.1 Range 0-10

Mauchly's test revealed that the assumption of sphericity was not met for the main effect of round, $\chi^2(9)=21.94$, $p<.05$, and the interaction effect between round and task, $\chi^2(9)=35.79$, $p<.001$. Degrees of freedom were therefore corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon=.31$ for both the main effect of round and the interaction effect between round and task).

For the range 0-10, there was no significant main effect of round on PAE. There was however a significant main effect of task type on PAE, $F(1,6)=35.64$. Contrasts revealed that T4 estimates were significantly more accurate (significantly lower PAE) than T2 estimates on this range. There was no significant interaction effect between round and task for estimates in the range 0-10.

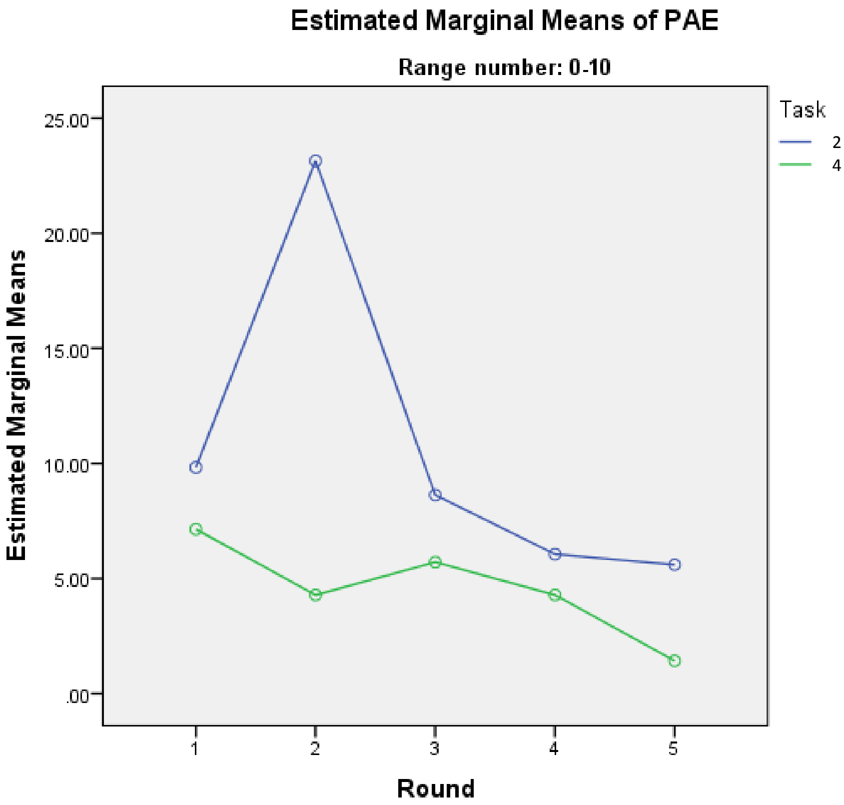


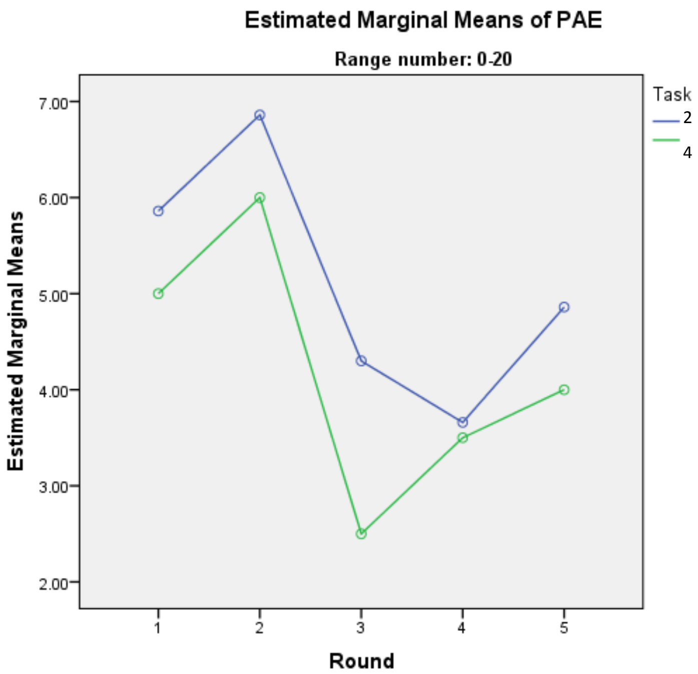
Figure 61 Mean PAE in range 0-10, R1-5

Looking at the plot of estimates by round and task for the range 0-10, for both tasks there was a general trend of increased linear accuracy (decreasing PAE) over the five rounds, but this change did not amount to a significant main effect.

What is particularly noticeable on this plot is the sudden increase in PAE for T2 estimates in R2. The mean PAE for T2 R2 estimates was skewed by unusually inaccurate (for Patrick) estimates of the targets 8 and 9: PAEs of 55.2% (target 8) and 60.8% (target 9). These estimates did not constitute statistical outliers, and both were apparently genuine estimates. The qualitative data gives clear support to particular hypotheses about how these very inaccurate estimates came about. With regard to target 9 for example, Patrick decided to estimate by counting on from the left endpoint using the width of his finger to guide the unit size. This was an unusual strategy for Patrick, and it resulted in a large underestimation of the target's position on the line.

7.2.3.3.2 Range 0-20

For the range 0-20, the assumption of sphericity was met for all effects: round, task and the interaction between round and task. None of these effects reached significance at $p<.05$. The plot of mean PAE by round and task in the range 0-20 is shown below.



The mean PAE for both T2 and T4 increased between rounds 1 and 2, and then fell sharply at round 3.

For T2, PAE then fell again at round 4 before increasing in the final round.

For T4, PAE increased at both rounds 4 and 5.

Figure 62 Mean PAE in range 0-20, R1-5

7.2.3.3.3 Range 5-15

For the range 5-15, the assumption of sphericity was once again met for all effects, and effects were judged significant (or non-significant) at $p<.05$.

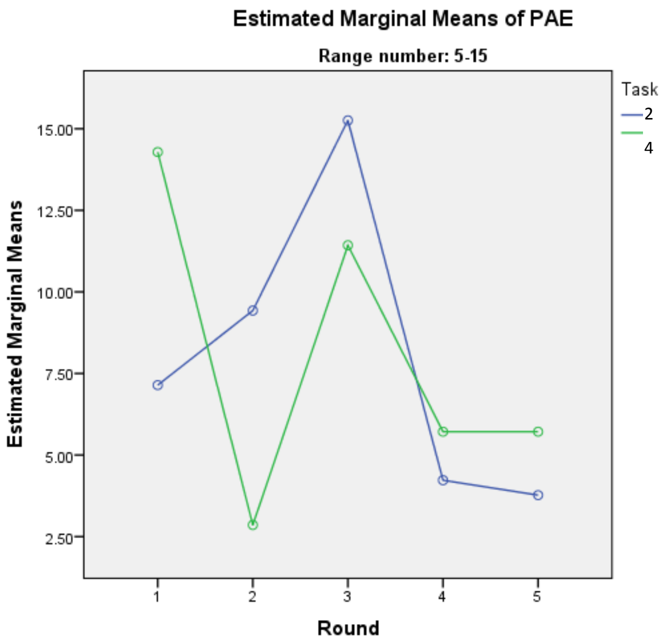


Figure 63 Mean PAE in range 5-15, R1-5

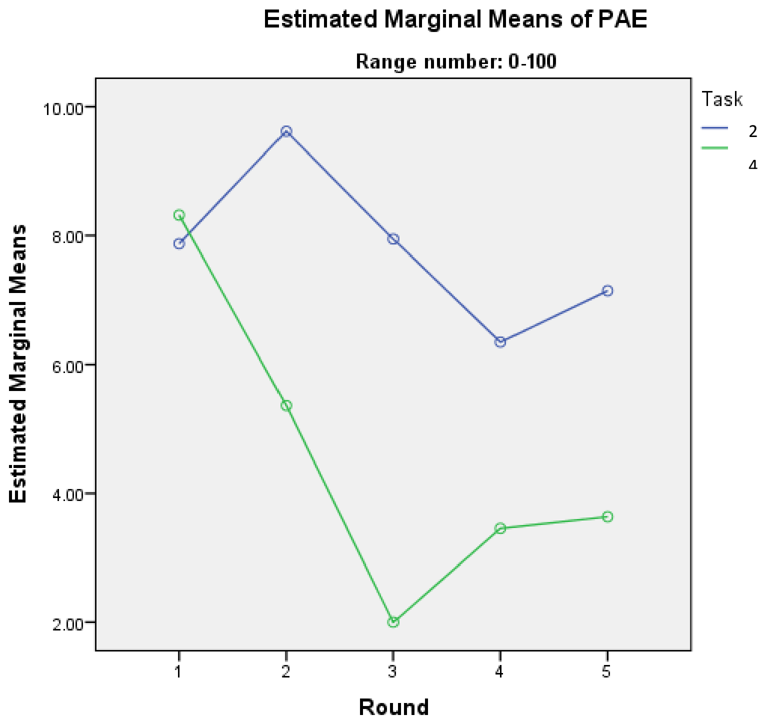
The only effect to reach significance in range 5-15 was the main effect of round on PAE, $F(4,24)=7.91$. Contrasts revealed significant variation in PAE between rounds 2 and 3, $F(1,6)=9.63$, and between round 3 and 4, $F(1,6)=17.39$.

The plot shows that this variation consisted of a significant increase in mean PAE at round 3.

7.2.3.3.4 Range 0-100

Mauchly's test revealed that the assumption of sphericity was not met for the main effect of round, $\chi^2(9)=19.83$, $p<.05$. The degrees of freedom were therefore corrected using the Greenhouse-Geisser estimate of sphericity ($\epsilon=.42$). For the range 0-100, as for the range 0-20, none of the effects (round, task or round and task interaction) reached significance at $p<.05$.

The plot below shows that for T2, PAE increased between the first and second interview rounds. Estimates then became more linearly accurate in the third and fourth round, before a small increase in PAE in round 5. Unlike other ranges, the T4 estimation accuracy on 0-100 followed a different path to that of T2: PAE decreased between the first and second, and second and third rounds, before increasing in both the fourth and fifth rounds.



Although there was no main effect of task on PAE for 0-100 estimates, this plot shows the wider disparity between T2 and T4 mean PAE on this range compared with on other ranges.

Figure 64 Mean PAE in range 0-100, R1-5

This section has highlighted different patterns in the linearity of Patrick's estimates in different task conditions: decreases or increases in PAE did not occur at the same time in different task conditions. The linearity of Patrick's estimates overall decreased very slightly, but this effect was only significant in range 5-15. Patrick's T4 estimations were overall slightly more linearly accurate than his T2 estimations, but there was only a significant difference for estimates in the range 0-10. Examining the graph of mean PAE in this range (Figure 61), it becomes clear that a large part of this difference between tasks is due to the 'spike' in PAE for Patrick's T2 range 0-10 estimates in R2. The following research question (specifically section 7.3.2.1) uses observations of Patrick's representations and strategies to offer a clear explanation for why PAE increased so significantly here.

7.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks?

This section discusses the strategies identified in Patrick's interactions with T2 and T4. It should be emphasised that in many cases the observed strategies were combined: it was frequent, for example, for Patrick to represent some anchor point and then reason aloud about the relative magnitudes of the target and anchor point.

The strategies identified in Patrick's T2 and T4 responses were the following:

- Reference to anchor point
    - Those included in the task environment (endpoints)

        - Those created or visualised by child
            - Midpoint
            - Quarter-point
            - Three-quarter point
        - Previous trials of the task
            - Within the same range
            - Within a different range
- Counting strategies
    - Count-on from LE to estimate
    - Count-on from estimate to anchor point (e.g. RE)
    - Count-back from RE
- Judgment using relative numerosities: particularly clear when estimate, or further strategies, appeared to be influenced by initial use of relative numerosity of target and task environment.

7.3.1 What patterns can be detected in the way children use or do not use these strategies?

Several patterns are evident from Patrick's case. The first is that the less mathematically sophisticated strategy of counting on from the left endpoint until the target number was reached, and then placing the number rocket, was used most frequently in the first round of interviews (when Patrick was at the beginning of the school year, and also unaccustomed to the interview tasks) and for targets close to zero (see RQ1c). Even in the first round of interviews, Patrick did not try to apply this strategy to larger targets in the range 0-100 for example.

To explore the hypothesis that reference points were selected as anchor points, target numbers were coded according to whether they fell nearest to a potential left-endpoint anchor (e.g. target 6 in range 5-15), to a right-endpoint anchor (e.g. target 19 in range 0-20), or to an midpoint or quarter point that a child could infer using part-whole number structure knowledge. Figure 65 plots the frequency of representations and strategies according to target type:

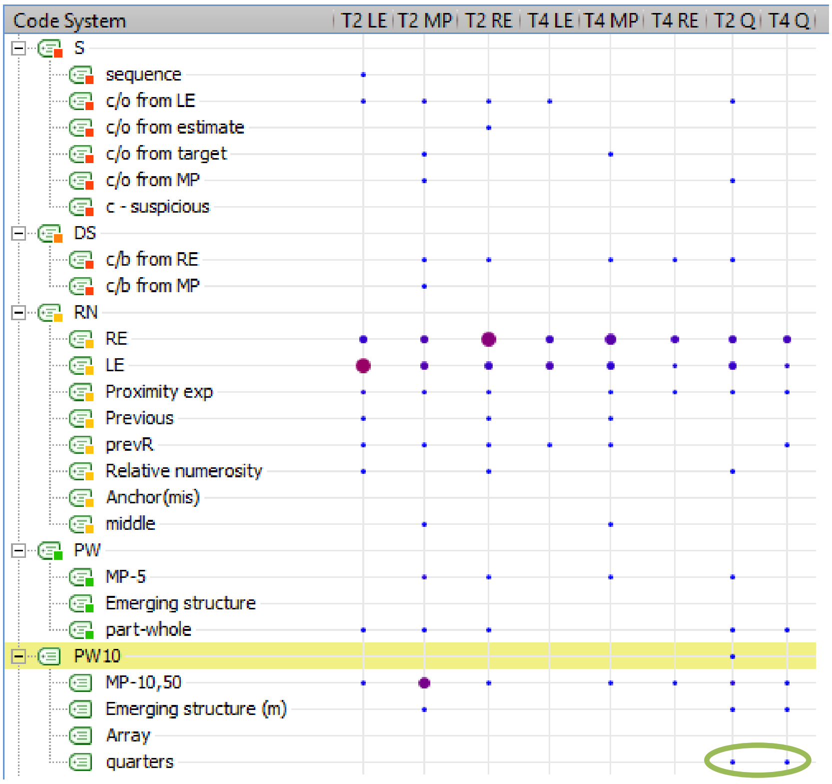


Figure 65 Strategies and representations by target type

For T2 estimates, the frequency of strategies and representations is clearly correlated with target type. Targets close to the left endpoint coincided with more frequent representations involving the left endpoint, and target close to the right endpoint coincided with more frequent representations involving the right endpoint. For targets close to a potential midpoint anchor, there was equal use of left and right endpoints, but a much higher frequency of representations involving the midpoint.

For T4 estimates, the connection is weaker. Targets close to the right endpoint do coincide with more right endpoint representations than left endpoint representations, and the only midpoint representations coincided with targets close to the midpoint. Target close to the right endpoint or midpoint in T4 estimations also recorded no count-on from left endpoint strategies.

The green circle identifies the only strategies and representations to involve quarter structure (including both one quarter and three quarter representations). Overall, these occurred rarely, but as this chart demonstrates, when they did occur it was always during an estimation of a target close to a potential quarter or three-quarter anchor.

7.3.2 How do the strategies used relate to children's estimation results?

In this section, visual plots of estimation accuracy are used to identify trials or rounds in which changes in children's estimation results are evident. After identifying points of investigation, the plots are then compared to charts showing the strategies and representations that occurred during the relevant estimation, to see whether connections can be drawn that may explain some of the estimation results. Three task conditions have been chosen for close examination: T2 0-10, T2 0-100, and T4 0-100.

7.3.2.1 T2, range 0-10

The following chart shows Patrick's estimates for T2 range 0-10, with the estimates for each round colour-coded.

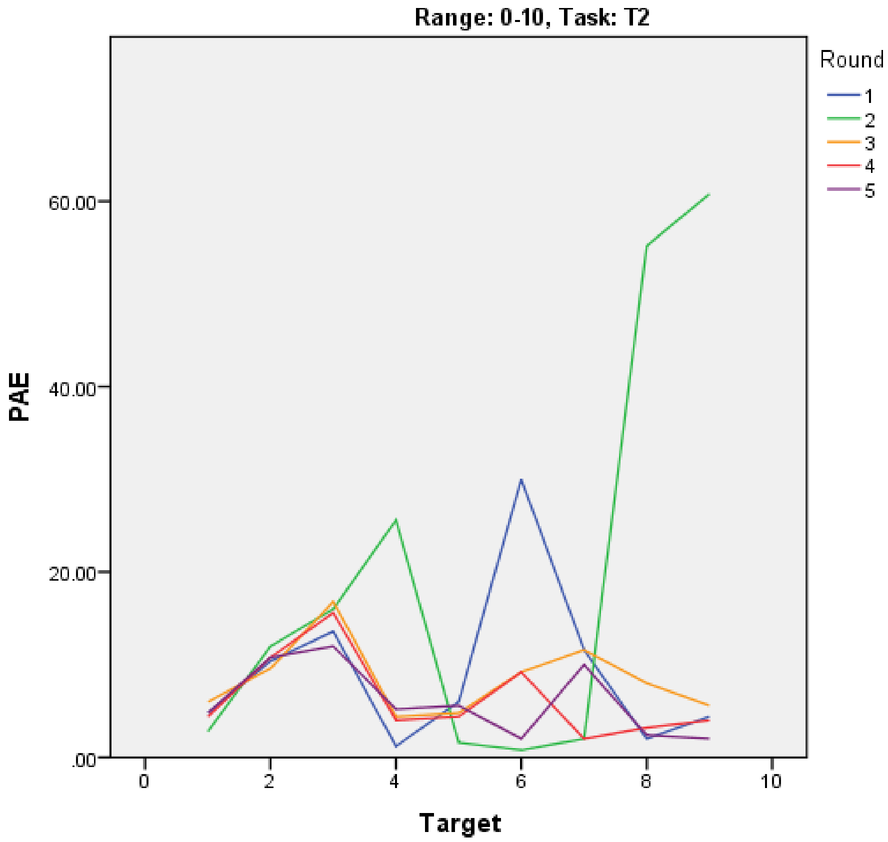


Figure 66 PAE of target estimates, T2 0-10

The initial observations for T2 estimates on 0-10 are that estimations were, with one exception, highly linearly accurate for targets 1 and 9, near each endpoint, and target 5, at the midpoint. The questions arising to be answered by the qualitative data are:

- What strategies were used for the highly accurate 1, 5, and 9 estimates?
- What is different in R2 for target 4?
- What is different in R1 for target 6?
- What is different for targets 8 and 9 in R2?

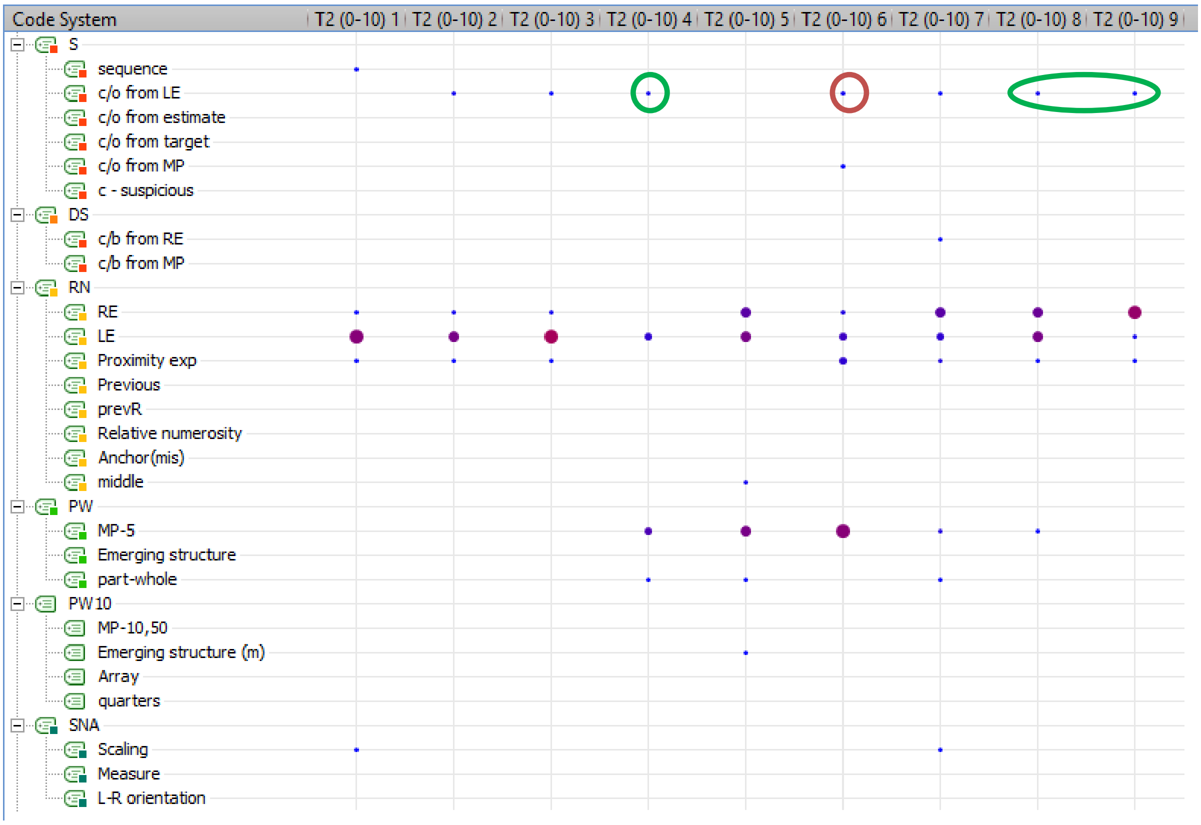


Figure 67 T2 0-10 estimates

The accurately-estimated target 1 coincided with frequent representations of the LE, the closest anchor point, and similarly the accurately-estimated target 9 coincided with frequent representations of the RE, its closest anchor point. Target 5 coincided with frequent references to an inferred midpoint, as the graph of estimation accuracy suggested.

Targets 3, 8 and 6 also coincided with frequent references to their nearest endpoints (LE, RE and midpoint respectively). However, the estimation of these targets was less accurate, pointing to the difficulty when estimating targets even slightly farther from the same anchor points. Target 6 was particularly inaccurately estimated in R1 (the blue peak on the graph), and examination of the transcript reveals here that this was the first trial of the first interview, and that Patrick counted up from the left endpoint (with no representation of the midpoint) – this instance is circled in red.

The R2 transcript was examined for targets 4, 8 and 9 (the green PAE peaks on the graph). For all three of these targets (circled green above), Patrick counted from the LE, and made no reference to any other anchor than the LE. In the context of the rest of Patrick's interview, and other round interviews, this was unusual, and it seems reasonable to believe it contributed strongly to unusually inaccurate estimations. Patrick in fact explained his strategy verbally on the first of these trials (target 9):

P: Cos I was like counting, I put my fingers to – to see how long it was, it was two fingers … along, so I counted that nine. [*puts left hand two fingers together to make a unit, and demonstrates measuring along the line rightwards using this unit*]

The gesture of counting along the line in this precise manner was repeated for target 8 and for target 4, although the verbal explanation was not.

7.3.2.2 T2, range 0-100

This chart shows the estimations made in T2 range 0-100, once again for all five rounds.

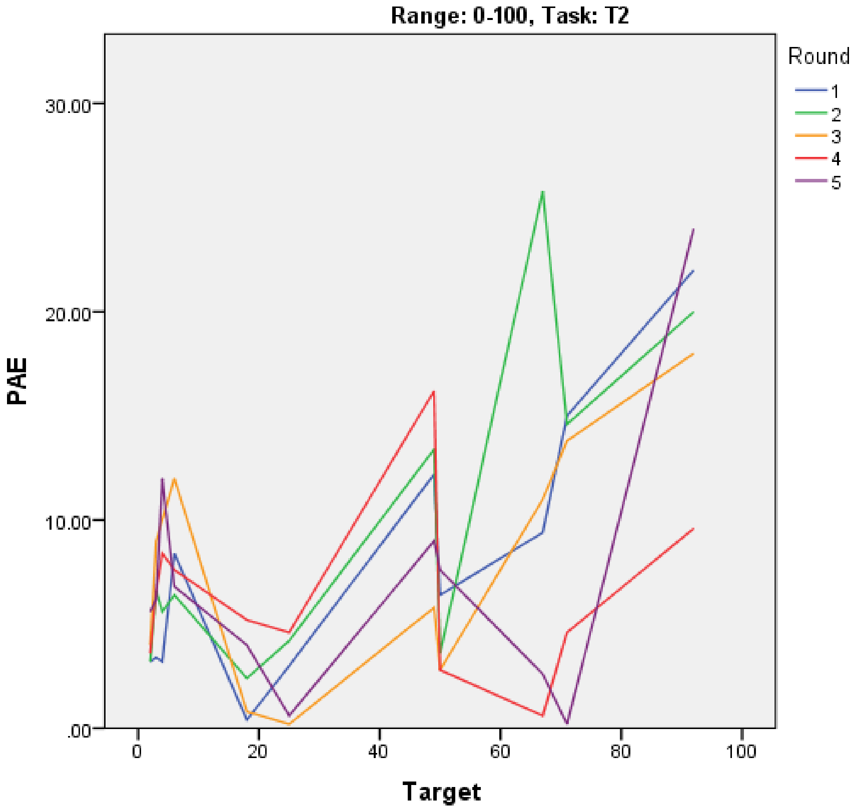


Figure 68 PAE of target estimates, T2 0-100

This chart demonstrates that T2 estimates for the range 0-100 were for the most part highly correlated between rounds. Considering all rounds together, accurate estimations occur for the very lowest targets near the left endpoint, for targets 18 and 25, and target 50. Looking at the following chart (Figure 69) which plots the strategies and representations for individual targets, the lowest targets featured mainly references to the left endpoint (circled orange), although several count-on strategies and representations of part-whole were also used. These strategies proved accurate for target 2, but increasingly less accurate by targets 4 and 6. No representations are recorded for target 18 other than references to the endpoints, but for target 25 part-whole structure is represented (circled blue) including the quarter structure, as the low PAE shown above might suggest. The low PAE at target 50 does indeed coincide with representations of the midpoint (circled green). However, midpoint structure was also represented frequently for target 49, and as the graph above clearly illustrates, these estimates showed high PAE. This is an interesting example of Patrick consistently representing relevant numerical structure for a target close to the endpoint, yet still recording high PAE.

For target 92, Patrick referred only to the right endpoint, showing use of relative numerosity. His linear accuracy for this target was however not high, with the exception of in round 4.

Once again, points of disparity between rounds are interesting. The questions that arise are:

- Why was the estimation of 67 so much less accurate in R2 than R1 and R3?
- What changed in Patrick's estimations of 67 and 72 in R4 and R5, where the estimates become suddenly much more accurate?
- How did Patrick estimate target 92 in R4, with significantly higher accuracy than in all other rounds?

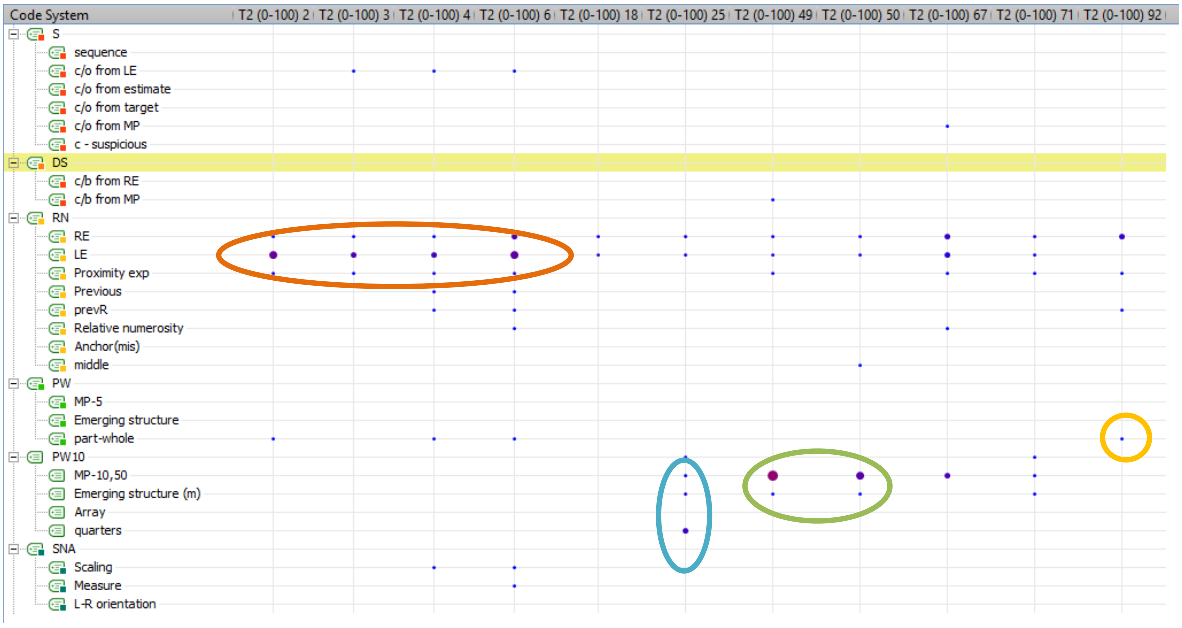


Figure 69 0-100 estimates, T2, all rounds

Looking at the transcript for the estimation in R2 of target 67, Patrick referred only to the midpoint and placed the rocket sticker very quickly. It is not clear why he did so (the chart above shows the estimation of 67 coinciding with representation of other structure in other rounds), but this strategy (or lack of) seems a likely contribution to the low accuracy of the estimation.

In R4, the transcript shows Patrick giving a good deal of attention to the estimation of 67, one potential reason for high accuracy:

P: Let's see … 67 … that goes there … so probably about therrre. [*takes rocket towards MP, hovers, then glides further rightwards. Left hand chops at midpoint, and rocket is stuck ~ 60%*]

J: lovely

P: I'm using the arrow to see where the rocket's going … tiny bit far from …

However, the estimation of target 71 occurred several trials later, and here Patrick made no comments or representations, sticking the rocket almost immediately. Target 92 (estimated most accurately of all rounds in R4) was similarly positioned without pausing, commenting or gesturing in any way. These trials do not provide any evidence in themselves as to how Patrick achieved accurate estimation. In R5, target 92 was estimated with the comment " … 8 off, so it's why it should be …" – indicating accurate part-whole knowledge of the relation between 92 and 100. However, this number structure was not enough to enable an accurate positioning of 92 – in fact this round's estimation underestimated the position of 92 more than any other round. A potential cause of this could be that Patrick's focus was solely upon the difference of 8 between 92 and 100, rather than the relative numerosities of 0, 92 and 100 as a set, and so over-estimated the 8 (causing underestimation of 92).

Target 67 in R5 was estimated very rapidly. Patrick's only comment was "Sixty seven … I think it could go there", taking the rocket immediately to its final position and sticking. Target 72 was similarly rapidly estimated, with Patrick commenting "Seventy one … there, cos it's a bit close, but not really close". This comment seems to refer to a judgment of relative magnitudes, although it is not clear to which other magnitudes Patrick was referring (no representation of endpoint or midpoint occurred in gaze or gesture).

The following graph (Figure 70) shows PE (percentage error), instead of PAE, to allow examination of over- and under-estimation patterns for T2 estimates on 0-100.

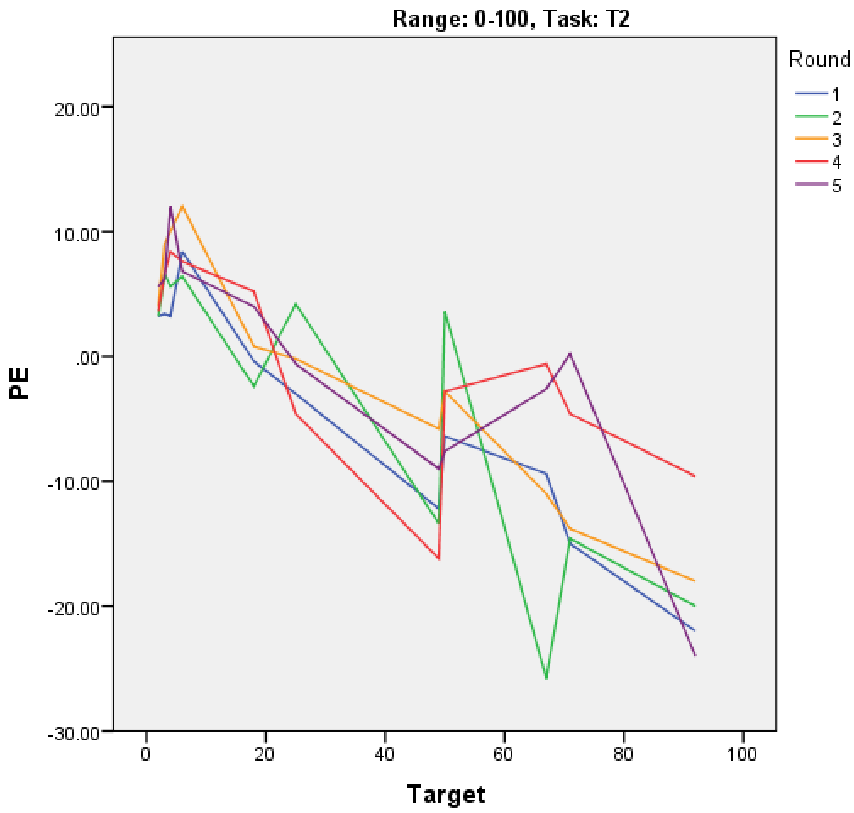


Figure 70 PE of target estimates, T2 0-100

In all rounds, there is over-estimation for targets under 25, followed by high accuracy around target 25. Estimations in all rounds then under-estimate targets before becoming more accurate again at target 50 (though the estimates for 50 still err on the side of underestimation). Estimates then increasingly underestimate, even for target 92, the closest target to the right endpoint. Although Patrick used the right endpoint as an anchor in estimating 92 (see Figure 69), it seems that the proximity of 92 to 100 was not close enough for Patrick to use the anchor to gain accuracy of estimation.

The patterns described above are somewhat consistent with the patterns of over- and under-estimation predicted by the psychological models of proportion judgment that are modelled using power functions (e.g. Slusser et al., 2013). These would predict S-shaped curves of over- followed by under-estimation with accuracy at anchor points (so a 1-cycle S-shaped curve for estimations using the two endpoints, plus midpoint). Patrick's estimates do not quite follow this distribution, but there is indeed a pattern of over-estimation followed by return to accuracy (at 25), then under-estimation, then a move back *towards* accuracy (target 50). After target 50, there is not over-estimation as the model predicts, but there is reduced under-estimation.

7.3.2.3 T4, range 0-100

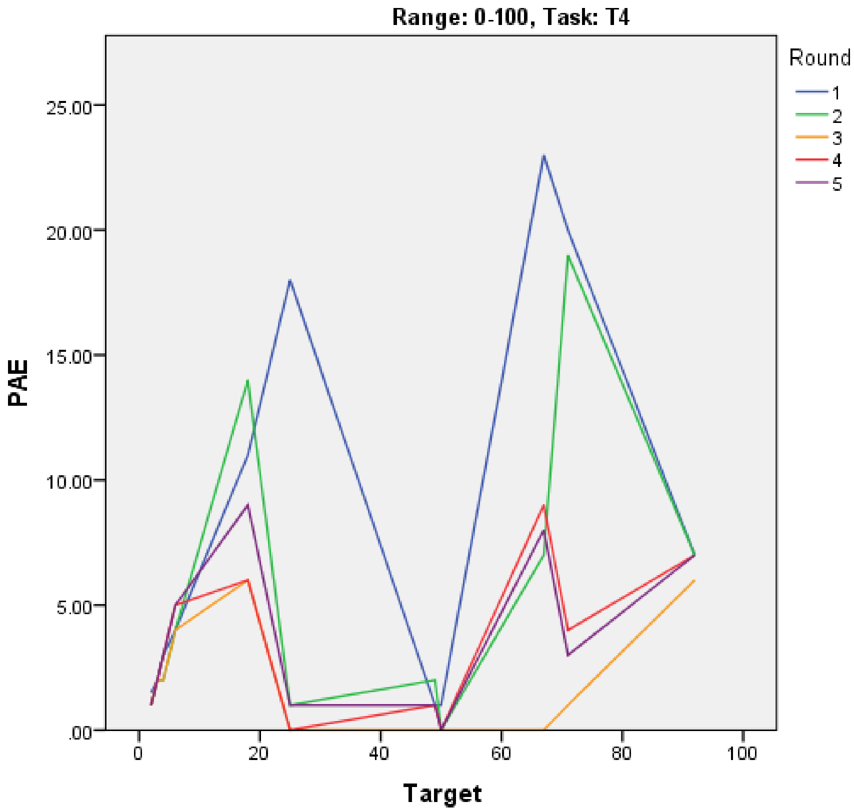


Figure 71 PAE of target estimates, T4 0-100

This plot of PAE for T4 estimates on 0-100 shows high accuracy for targets very close to the left endpoint, and very close to the midpoint. For selected rounds, there is also high accuracy at target 25 and 71. The questions arising for the qualitative data are:

- What was different about the estimation of 25 in R1? Is there any sign of quarter point representation?
- Is there qualitative evidence for the midpoint anchor?

As in T2, targets 67 and 71 are interesting. In R1 and R2, both 67 and 71 are noticeably less accurate than in all three later rounds. More specifically, in R4 and R5, accuracy suddenly increases. In R3, estimation accuracy is very high throughout.

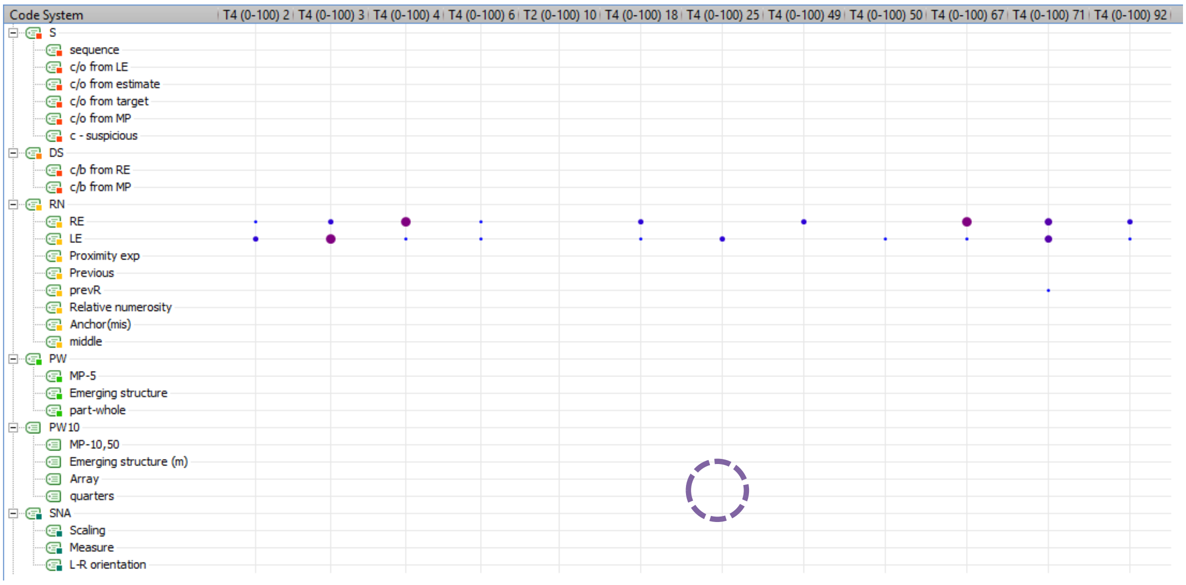


Figure 72 T4 0-100 estimates, R1 and R2

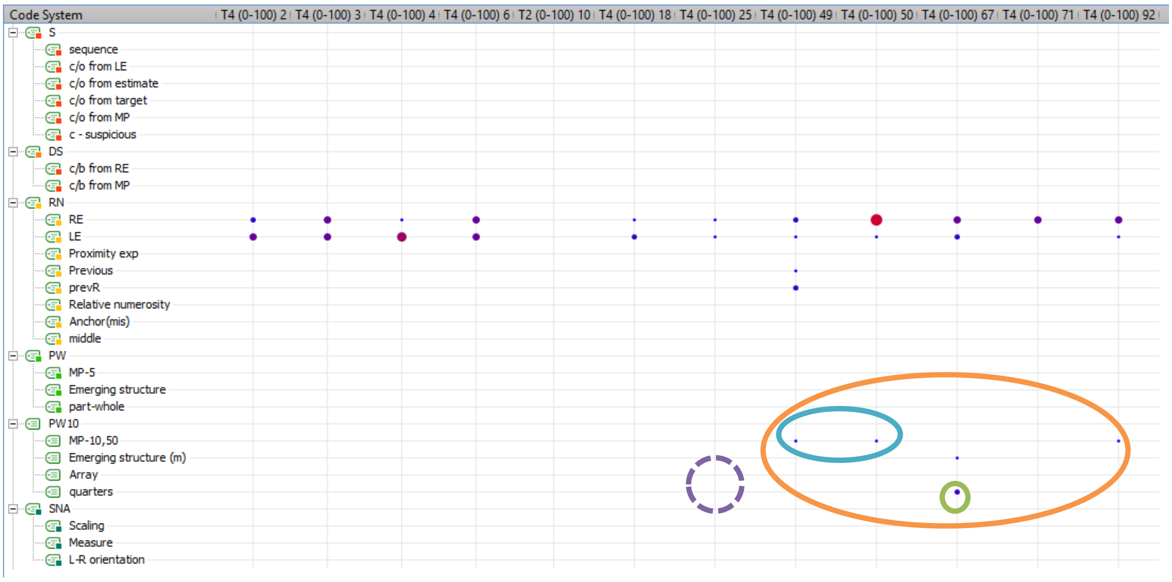


Figure 73 T4 0-100 estimates, R3 R4 and R5

These two charts show visually the differences between number structure represented in R1 and R2, compared to in R3, R4 and R5 (for T4 0-100 estimates). The obvious difference (circled orange) is the representation of part-whole structure including base ten structure in R3, R4 and R5, which is entirely absent in R1 and R2.

The graph of linear estimation accuracy highlighted a fairly sudden decrease of PAE for estimates of 67 and 71 in the later rounds (compared to R1 and R2), and this coincides with representation of the quarters of 100 (circled green) that did not occur in R1 and R2. Looking to the interview transcripts, Patrick's application of the structure in answering the task seems clear. In R4, for example:

P: See to here… it's … just over three-quarters so I'm going to go for 27 – 77! [*writes answer 77*]

Whilst target 67 shows a sizable decrease in PAE for later rounds, the estimation accuracy is not in fact as low as for target 71, which also shows a decrease in PAE to an even more accurate level. As discussed elsewhere, the context of the individual target estimations provides an important second layer of evidence. Thus, although quarters of 100 were not explicitly represented for target 71, since this structure was represented by Patrick for a close target in the same set of estimations within the interview, it is a reasonable hypothesis that knowledge of this structure could have helped the linear accuracy of his estimation of 71 as well as of 67.

The two charts of number structure representation also show representation of the midpoint of 100 (circled blue) for targets 49 and 50 in R3, R4 and R5, not present in R1 and R2. Going back to the graph of PAE, these targets were already accurately estimated in R1 and R2, but in later rounds they are indeed even more accurately estimated.

A final point in the consideration of T4 0-100 estimates, which underlines again the importance of considering the interview context and not only representations made in the moment of each individual trial, is notable absence of representation of quarter structure during estimation of target 25 (circled purple). This target was very accurately estimated during all rounds except R1, and in no round was the quarter structure of 100 represented during its estimation. From the representation of number structure seen in the rest of the interviews, it would be reasonable to infer that Patrick knew that 25 was a quarter of 100 (and indeed comments during T2 say this explicitly), but there is no direct evidence of the target being represented with this structure. This is an example of the limitations of the methods used in this research (and other research): despite efforts to capture as much data as possible about Patrick's interaction with the task, these methods are unable to provide evidence as to whether Patrick was using one quarter as an anchor point in this situation.

Looking at the graph of percentage error (PE) as opposed to PAE, an additional trend can be seen: of underestimation on the lower half of the range, accuracy at the midpoint, and overestimation on the upper half of the range. This is as expected from the literature on proportion judgment (e.g. Rouder and Geary, 2014). As in T2, the patterns of over- and under-estimation do not follow very precisely what would be predicted by a one- or two-cycle power function model. However, particularly in R4 and R5, a pattern of under-estimation, followed by increased accuracy, followed by over-estimation can to a certain extent be seen.

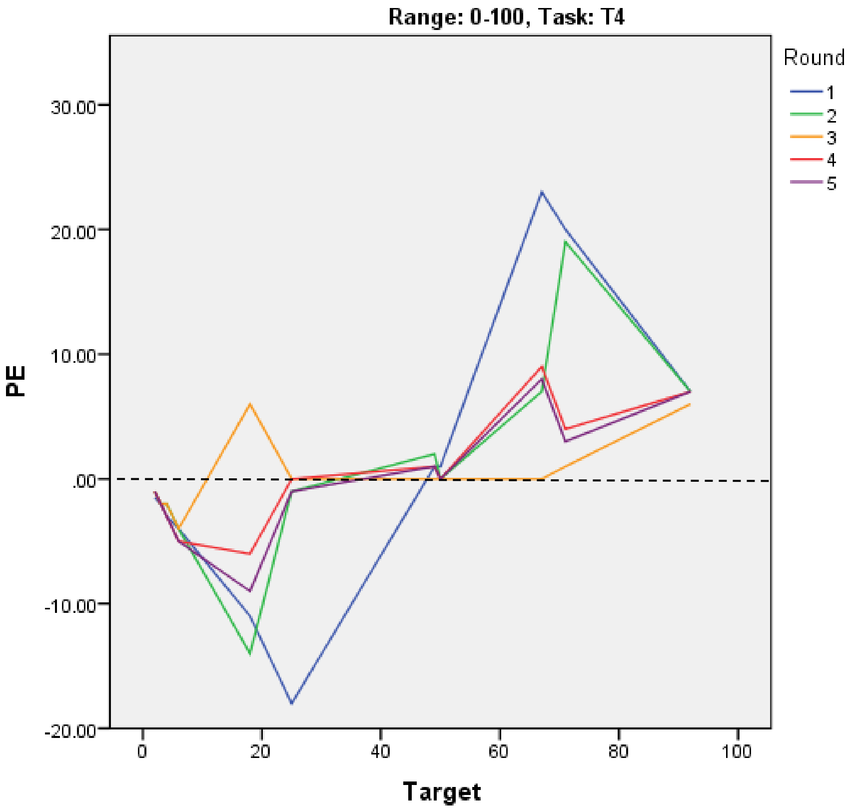


Figure 74 PE of target estimates, T4 0-100

This section has demonstrated that many patterns in Patrick's estimation results can be connected to observations of the representations and strategies during individual estimation trials. There are numerous examples of trials with particularly high PAE which can be plausibly explained by specific observations of representations and strategies, for example count-on from left endpoint with an inappropriately scaled unit size. Accuracy of estimation near the left endpoint is well supported by qualitative evidence for the use of a left endpoint anchor, and where targets close to the right endpoint were accurately estimated, there is good evidence for the use of a right endpoint anchor.

In comparison to left and right endpoint anchors, midpoint anchors appeared to be used less frequently, as expected. There are consequently fewer trials to examine, but where the estimation error of targets close to the midpoint was low, representation of the midpoint was often observed. The reverse relationship was less clear, with representation of the midpoint certainly not always associated with more linearly accurate estimation. The representation of quarter and three-quarter points was associated with more accurate estimates.

Although there are many examples where difference in PAE was associated with different observed representations and strategies, the relationship is not simple, and the representation of more sophisticated number structure did not always result in a more accurate estimation.

The findings in this section provide limited support for a proportion judgement account of number line estimation, but only really for estimates in the range 0-100. Estimates in this range, for both T2 and T4, display several of the patterns of over- and under-estimation predicted by proportion judgement accounts.

7.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling?

This question involves examining the changes evidenced by both the qualitative and quantitative estimation data collected from Patrick's interviews. Before considering changes in estimation results, and the bearing they have on rival accounts of number line estimation in the research literature, Figure 75 summarises the aspects of number structure that Patrick represented in each round of interviews. As in previous sections, the representations and strategies coded have been grouped according to the five main stages of Resnick's (1983) developmental theory of number understanding: sequence (S), decreasing sequence (DS), relative numerosity (RN), part-whole (PW), and part-whole base-ten (PW10). As before, the size and colour of circles indicates relative frequency. However, whilst in previous charts frequencies were compared to all other frequencies in the same chart, in this case they are compared only to the frequencies of the same category of representation in other rounds. The advantage of this is to better highlight trends within individual categories of representation, which were partially obscured when frequencies were compared throughout the whole chart (due to the overall dominance of certain categories). The disadvantage is that comparisons cannot be made *between* categories. A large red circle indicates a high frequency of a representation (compared to its frequency in other interview rounds), and a small blue circle indicates a relatively low frequency.

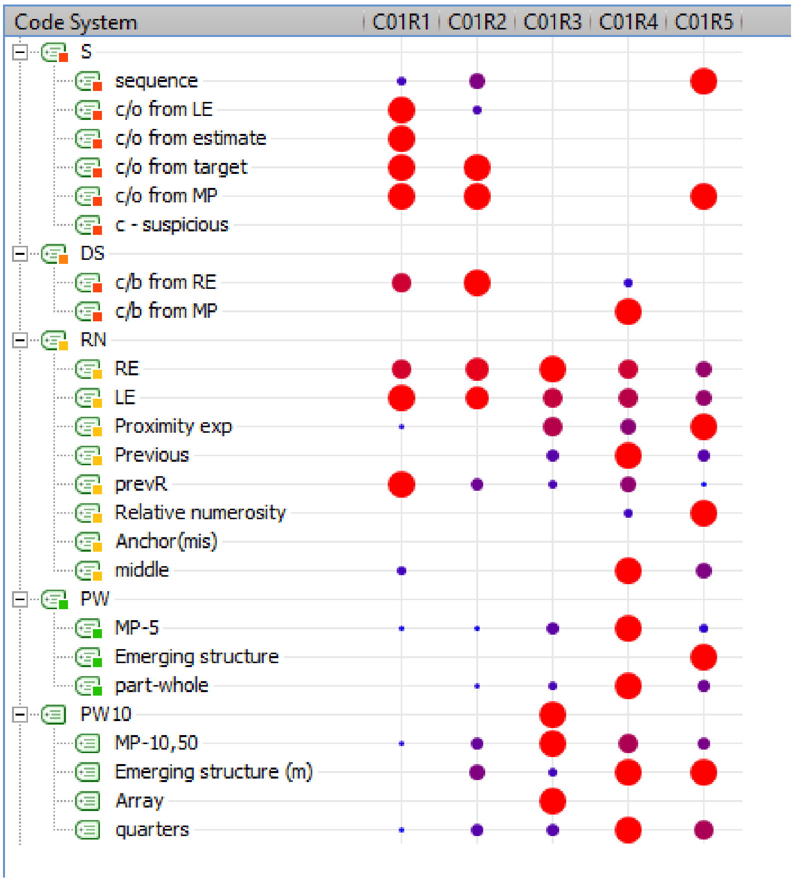


Figure 75 Number structure represented by round

There are two main points to be drawn out of this summary data. Firstly, that Patrick represented all five main stages of number structure in almost every round. The only exceptions to this are that sequence structure was not directly represented in R3 or R4, and decreasing sequence structure was not represented in R3 or R5. It must be emphasised that this does not mean that Patrick did not use his knowledge of the number sequence or decreasing number sequence in these interviews. The current research did not set out to answer this question – to examine qualitative data for evidence that particular number knowledge was used or implied. Rather, it only means that in these rounds Patrick did not directly represent that structure or a mathematical strategy (e.g. count procedure) that directly involved that structure. What is more interesting than the absences is the positive presence of representations of the more sophisticated structures in each round.

The second main point is that the diagram illustrates a general trend towards more representation of part-whole structure with time, and less representation of sequence structure. In terms of Resnick's account of number understanding development, the trend is one of representing increasingly sophisticated number structure throughout the school year, as we might expect. The fact that the less sophisticated aspects of number structure were represented less and less frequently does not follow necessarily; one could represent increasingly sophisticated number structure whilst retaining the representation of less sophisticated elements. However, Patrick did not do this - consciously or unconsciously he increasingly chose not to represent number sequences. A reasonable hypothesis to explain this could be that as he gained confidence and fluency with more sophisticated number structure, he no longer found it efficient to represent number sequence as frequently as he did at the start of the year. Tied in with this hypothesis is the fact that the children in this research gained repeated experience with the interview tasks used ; this stability afforded plenty of opportunities for children such as Patrick to develop their task responses over time.

7.4.1 In what ways does evidence support or not support the log-linear hypothesis?

Patrick's number line estimations were highly linearly accurate during both T2 and T4 (number to position estimation, and position to number estimation) and across all five interview rounds.

For the majority of task conditions (a given range and given task) in each round, Patrick's estimates were better fitted by a linear model than by a logarithmic model or exponential model. The following sequence of charts shows the R^2 model fits of the three models fitted: linear, exponential and logarithmic, for each task condition, across the five interviews rounds.

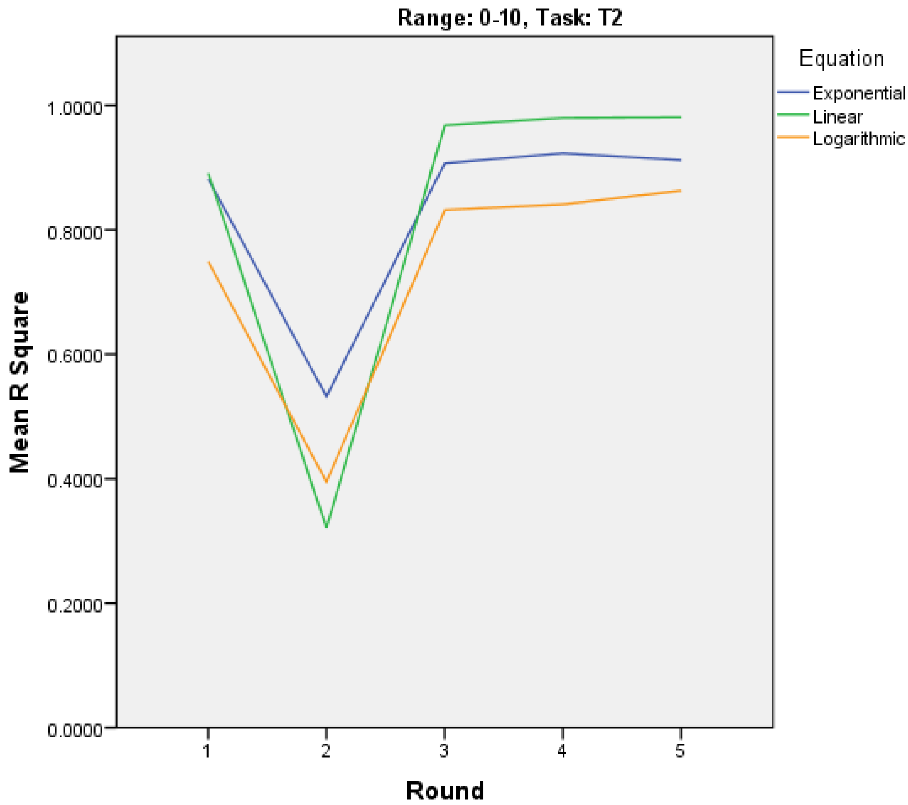


Figure 76

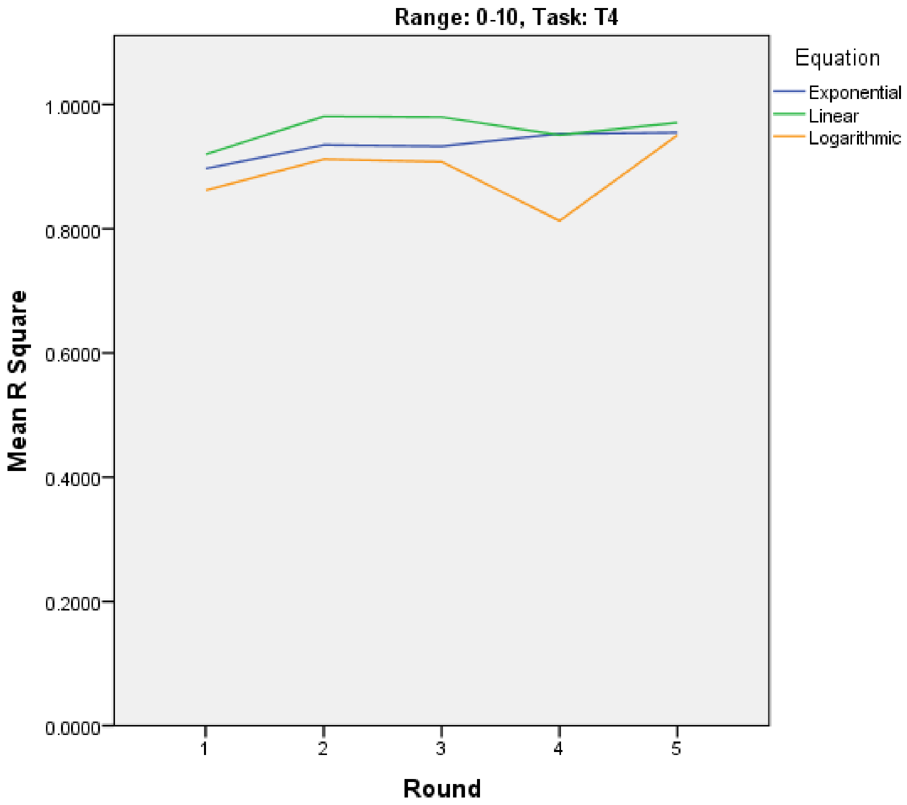


Figure 77

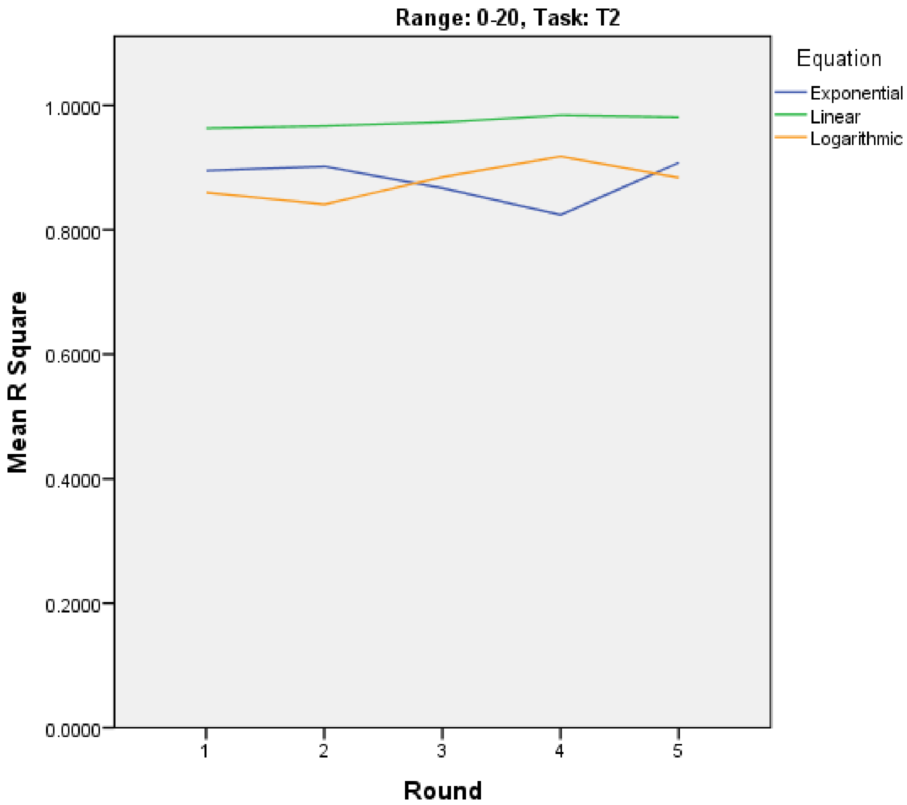


Figure 78

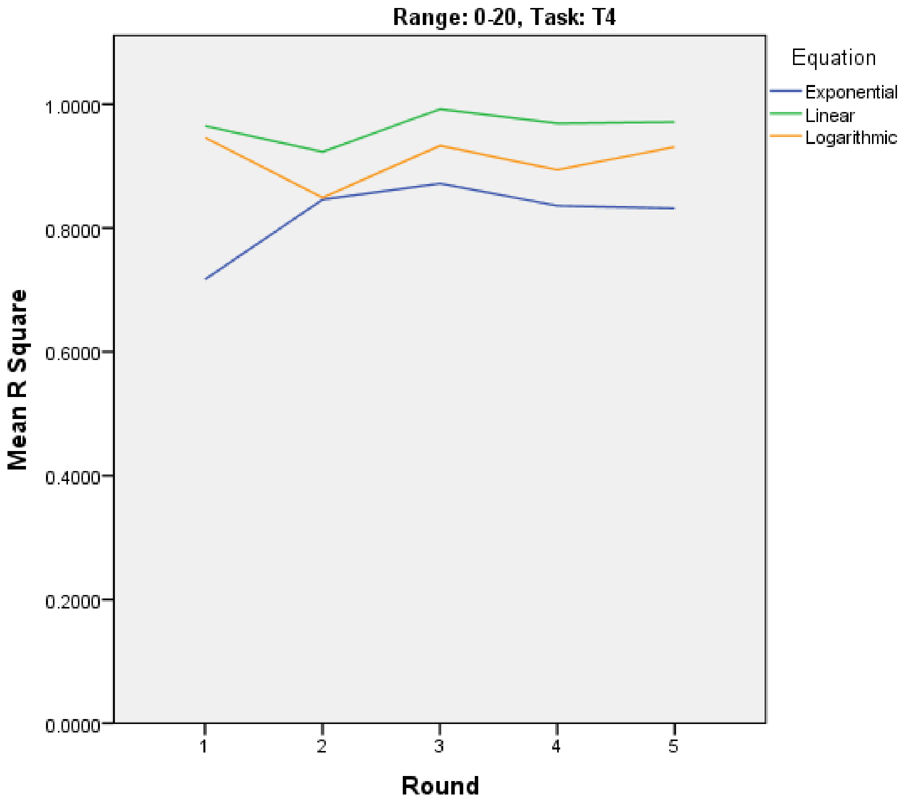


Figure 79

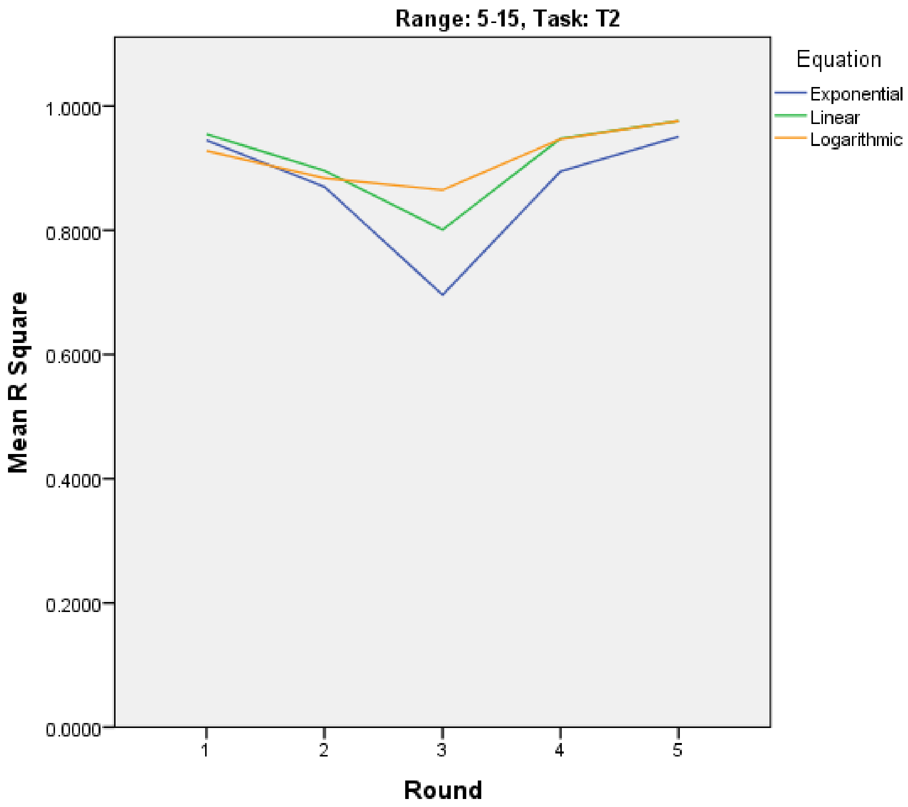


Figure 80

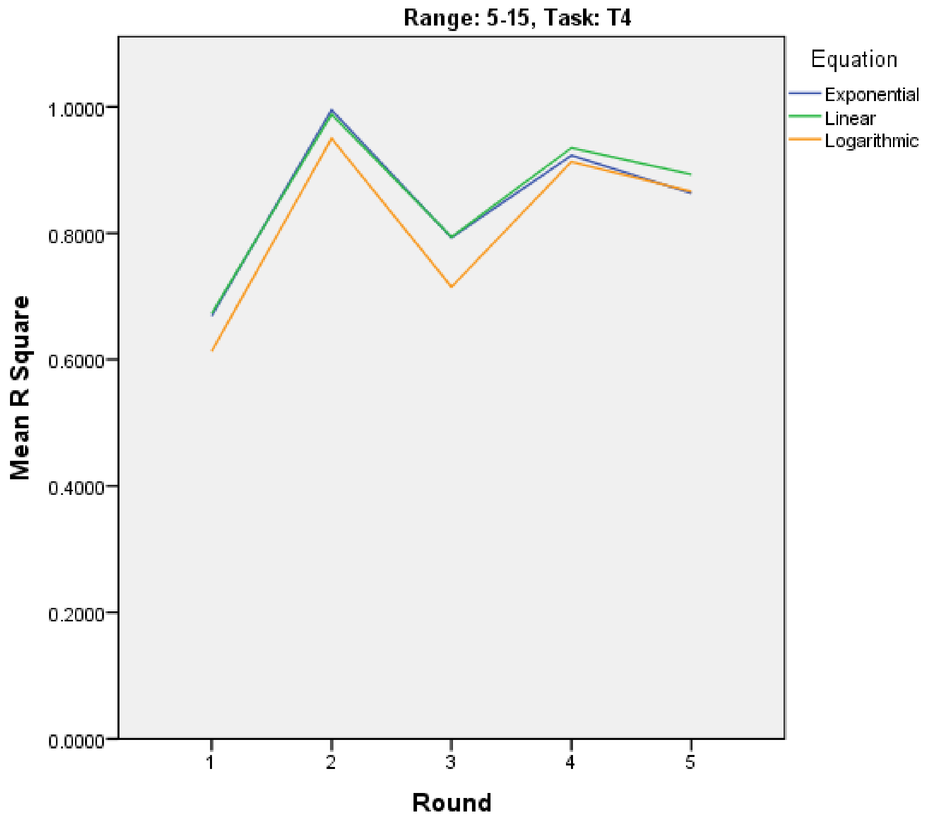


Figure 81

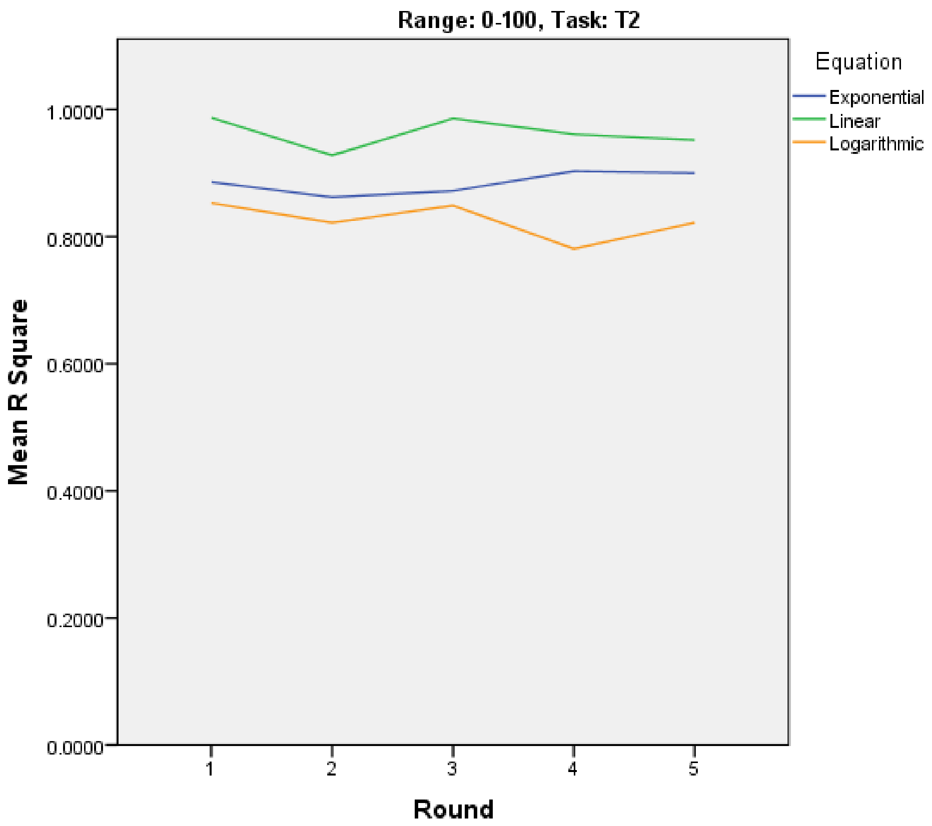


Figure 82

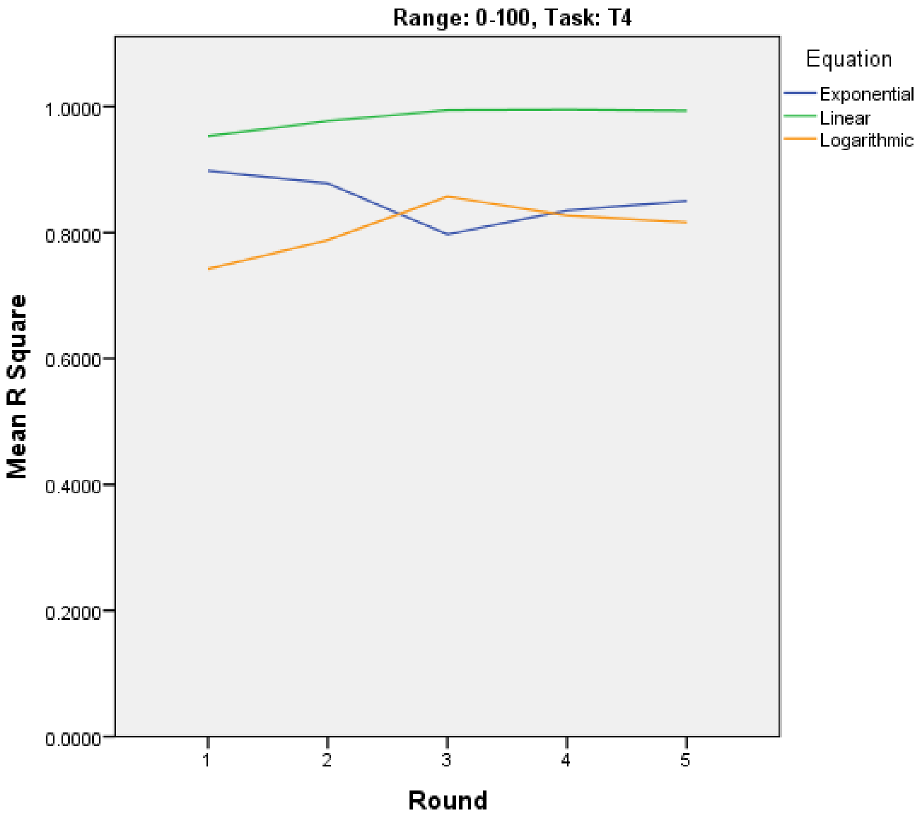


Figure 83

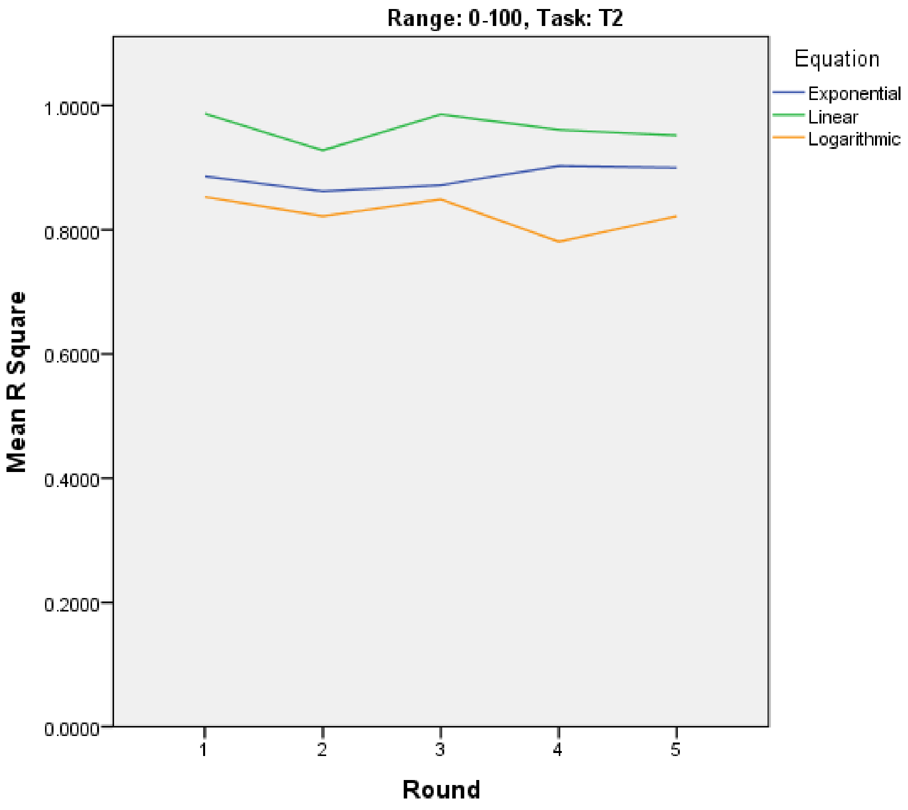


Figure 84

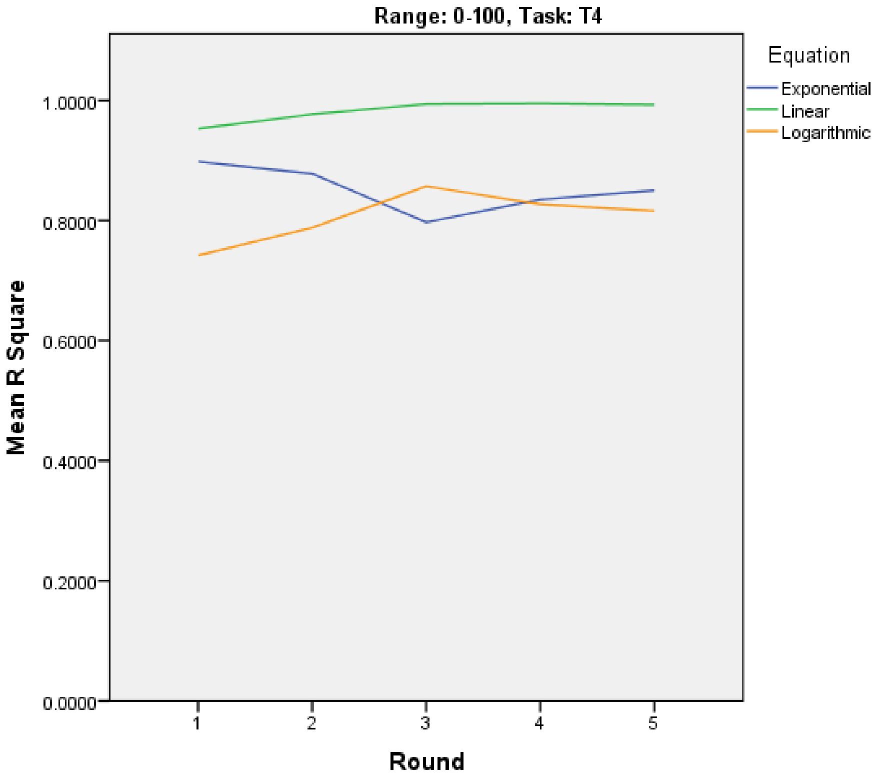


Figure 85

These charts reveal good model fits for linear models compared to logarithmic and exponential models. In some conditions the advantage of the linear model (green) is clear; in other conditions (range 5-15) the difference is marginal. Importantly, on only two occasions do the fits of non-linear models 'overtake' the linear fit: for Task 2, range 0-10, the linear model performs worse than either alternative model in R2 (only), and in Task 2, range 5-15, the linear model performs worse than the logarithmic model in R3 (only). The suddenly lower linear accuracy of estimations in T2, 0-10 R 2 has already been discussed (see RQ1c, Range 0-10).

To see more clearly how well the linear models fit, the following table indicates the R^2 figures of the best fitting linear model for each task condition, with shading to show where the linear model explained over 89% of variance.

Range	Task	R1	R2	R3	R4	R5
0-10	T2	0.89	0.32	0.97	0.98	0.98
0-10	T4	0.92	0.98	0.98	0.95	0.97
0-20	T2	0.96	0.97	0.97	0.98	0.98
0-20	T4	0.97	0.92	0.99	0.97	0.97
5-15	T2	0.96	0.90	0.80	0.95	0.98
5-15	T4	0.67	0.99	0.79	0.94	0.89
0-100	T2	0.99	0.93	0.99	0.96	0.95
0-100	T4	0.95	0.98	0.99	1.00	0.99

Figure 86 Table showing R^2lin values for each task condition, by interview round

The ANOVA analysis of linear accuracy (PAE) reported in RQ1 revealed that overall, the accuracy of Patrick's estimations were very stable. The only estimation range for which there was a significant effect of round, or round and task interaction, was the range 5-15, for which there was a main effect of round (see RQ1c).

Patrick's case does not therefore provide direct evidence with regard to an observed logarithmic-to-linear shift of estimates on any given range, since such a shift did not appear to occur.

7.4.1.1 Alternative accounts

Patrick's case is additionally relevant to the debate *surrounding* the log-linear shift account and its rivals, since it provides qualitative data on the estimation of individual target numbers. These can be compared against specific hypotheses proposed in the literature, particularly as alternatives to the log-linear shift account, as outlined in the Plan of Analysis section (4.5.4).

Plotting PAE by target number is a powerful way to visually assess the hypotheses of proportional judgment account that estimation accuracy should be highest near to anchor points and lowest at the farthest point from anchor points. This can be seen in RQ2b; which looked at changes in estimation accuracy to see whether qualitative evidence supported the idea that the changes were linked to developmental change in task response.

A very important limitation when considering the plausibility of the proportion judgment account (modelled with cyclical power functions) is that target numbers in this research were not sampled evenly in all ranges. In ranges 0-10, 0-20 and 5-15, targets were evenly sampled from across the range. However, for range 0-100, in common with much previous number line estimation research, targets were sampled specifically to compare the fits of logarithmic and linear models, i.e. sampling more heavily from the lower part of the range (e.g. Ashcraft and Moore, 2012; Siegler and Opfer, 2003; Thompson and Opfer, 2010). The impact of this is well explained by Slusser and colleagues: "This practice focuses on participants' propensity to overestimate small numbers, but yields little data to reveal the details of underestimation patterns for larger numbers." (Slusser et al., 2013, p. 196)

7.4.2 What is the intra-child variability of children's numerical magnitude representations in estimation tasks at different times?

There are several ways to approach assessing the variability of children's numerical magnitude representations. One simple measure is the variance of estimation error (PAE) in different task conditions at different times.

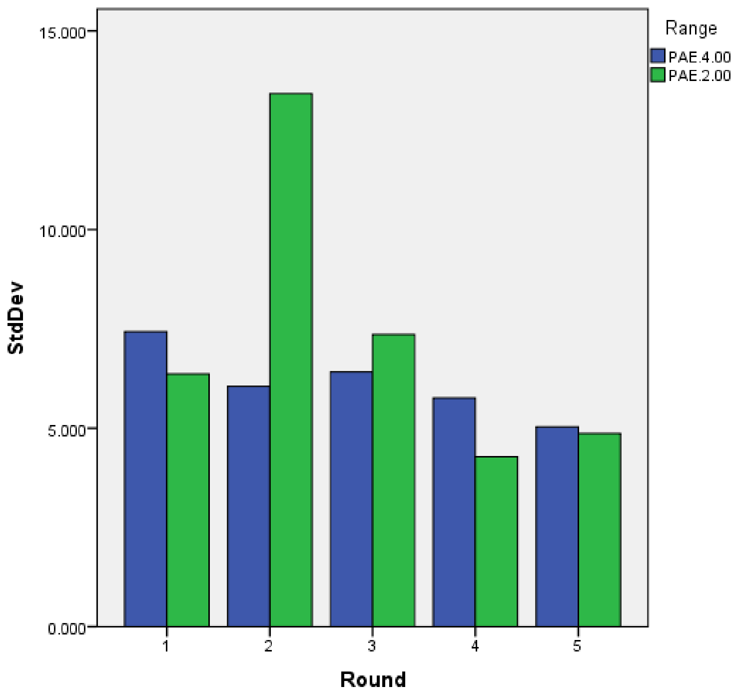


Figure 87 Variance in PAE for T2 and T4 by round

This chart shows overall small differences between the variability of T2 and T4 estimation accuracy, except in R2, where the standard deviation of estimation accuracy in T2 is considerably larger than in T4, and considerably larger than in either task in other rounds. Over the course of the five interview rounds, variability of estimation in both tasks overall decreased.

As described in RQ1c, the variability of estimates within different ranges was not the same, but no strong overall trend emerged. In R1, the standard deviation of PAE was comparable for ranges 0-10, 5-15 and 0-100 (between 4% and 10% for both T2 and T4) and lowest for estimates in range 0-20: 3.4% (T2) and 4.7% (T4). In R2, the standard deviation of PAE lay between 4.8% and 8.1% for all task conditions except for 0-10 T2 estimates, where the standard deviation was 23.3%, due to Patrick's unusual estimations for targets 8 and 9. In R3, the standard deviation of PAE fell to between 2.3% and 5.7% for T2 and T4 on 0-10, 0-20 and 0-100, but was noticeably higher in range 5-15: 11.5% (T2) and 10.7% (T4). In R4, variability was comparable across ranges with the highest standard deviation of PAE occurring in T2 0-10 estimates (8.3%) and the lowest in 0-20 T4 estimates (2.0%). Finally, in R5, standard deviation of PAE was between 3.1% and 5.2% for all conditions except T2 estimates on 0-100 (6.6%) and T4 estimates on 5-15 (7.9%).

7.4.3 Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?

Patrick's case illustrates two expected overall trends. Firstly, that over the course of the school year, he represented more sophisticated aspects of number structure more frequently and represented less complex aspects of number structure less frequently (see Figure 75). From the data collected, there is no evidence that this change occurred suddenly; it appeared gradually between each round of interviews. Task 1, in which Patrick drew a representation of numbers to 100 in each round, provides a salutary reminder that this task captures only a partial glimpse of children's representation of numbers. The array structure represented in R3 appeared "suddenly" after R2, and then "disappeared" in R4 and R5, but within the context of the rest of the interviews – coded for number structure representation throughout – that it can be seen as just one example of Patrick's increasingly frequent representation of part-whole number structure, including base ten structure. Seen in this light, the array representation did not appear 'suddenly' and did not disappear afterwards. The factors influencing the choice of a particular drawn representation on each interview day are of course numerous, and including not least which representations of number children have been working with in classroom maths lessons that week, but that is not to diminish the point that Patrick's option to represent number on paper in a particular way occurred as part of a larger trend in his representations of number.

The second main trend illustrated by Patrick's case is a small overall increase in the linear accuracy of number line estimations, a result clearly predicted by previous studies. Patrick was unusual in producing highly linear estimates from the very first round (on all ranges), and so no statistically significant increase in linearity was measured overall. It did however amount to a significant main effect for estimations in the range 5-15, which the majority of children found to be a 'difficult' range because of the non-zero start point.

Chapter 8 Case study: Marta

Marta was selected as the second of the three in-depth case studies because, in the terms of this research, her responses to the task-based interviews were among the least mathematically sophisticated: she represented few aspects of number structure compared with other children, and her number line estimations (both number to position and position to number) were the least linearly accurate of all the children interviewed.

8.1 What does this case say in response to important ideas in the literature?

Marta's case provides an in-depth look at the number representations of a child who showed less mathematical sophistication than her peers. The number structure Marta represented confirmed the developmental trajectory outlined by Resnick (1983): Marta's representations did not change dramatically over the course of the research interviews, but with time she represented more number structures identified by Resnick as more mathematically advanced, and decreased her reliance on the less sophisticated number representations.

Marta's case provides good qualitative evidence for the use of lower and upper endpoints, and occasionally midpoints, as 'anchors' during number line estimations, as hypothesised by many previous researchers from non-qualitative data.

The case provides little support for the logarithmic-linear shift hypothesis. Although Marta's number line estimations had initially low linear accuracy, they were not well fit by logarithmic models (or exponential for the inverse task T4) either, and in no task condition was there a 'shift' to a linear model. Instead, as for Patrick, Marta's case more convincingly illustrates simply an increase in the linearity of estimation over the course of the school year. The increase in linearity was not the same in all task conditions, and the largest change occurred between R1 (the first interview) and subsequent interviews.

A particularly useful contribution of Marta's case is good evidence for the hypothesis that younger children's linearly inaccurate number line estimations may often be attributed to a basic strategy that only takes account of the left endpoint of number lines. Marta's initial representations and strategies were dominated by references to the left endpoint and by 'count-on from left endpoint' strategies. Marta's case provides clear examples of trials in which, as Slusser et al. (2013) for example suggest, the number line task as completed by the child is effectively open-ended, since the right endpoint is either noted and dismissed, or not considered at all.

With respect to the connection between increasing linear accuracy and strategy, Marta's case, like others, provides evidence for the importance of anchor points. In many task conditions, striking increases in linear accuracy occurred between rounds for key target numbers: those very close to the right endpoint or to the midpoint. This kind of evidence has been interpreted by previous authors (e.g. Ashcraft and Moore, 2012) as evidence for the stepwise inclusion of additional 'anchor' points in children's estimations. Matching Marta's estimations to observed representations and strategies provided some clear examples of representation of new (compared to the same trial in previous rounds) number structure that coincided with substantially increased estimation accuracy, for example, a new reference to the midpoint accompanied by far higher accuracy for target 49 than in previous rounds.

In other episodes, however, new representation of more sophisticated number structure did *not* coincide with increased estimation accuracy, and vice versa. This reinforces two important points. First is the importance of considering the rest of the trials on a given range in a given interview, since the number structure that a child is able to represent for a given number range may not be represented in each trial, even whilst that structural knowledge still informs other trials. This study was of course not able to directly assess the knowledge that was used in a particular estimation. The best understanding of a single estimation trial in this study was gained from combining estimation patterns (such as suddenly increased accuracy for a key target), number representation during the single trial, and number representation during other trials of the same task condition. The second important point is that increased estimation accuracy may derive from other changes in addition to the representation of more number structure. This could relate to the parameter indexing bias (β) in proportional reasoning models, although as for Patrick's case, Marta's case was not designed to test this theory specifically.

In the remainder of this chapter, I present the findings from Marta's case study in relation to each of the research questions.

8.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research?

8.2.1 What are the modes and component signs used in the representations?

In four of the five interview rounds, Marta produced a drawing on paper during Task 1. In R3, Marta gestured and verbally described but produced no inscriptions. The drawings Marta produced in T1 are shown in Figure 88, and in order to facilitate side-by-side comparison of the drawings, the white space surrounding each drawing has been removed. Each drawing originally occupied an approximately 40mm x 40mm square in the centre of the A4 page provided, with the paper in landscape orientation.

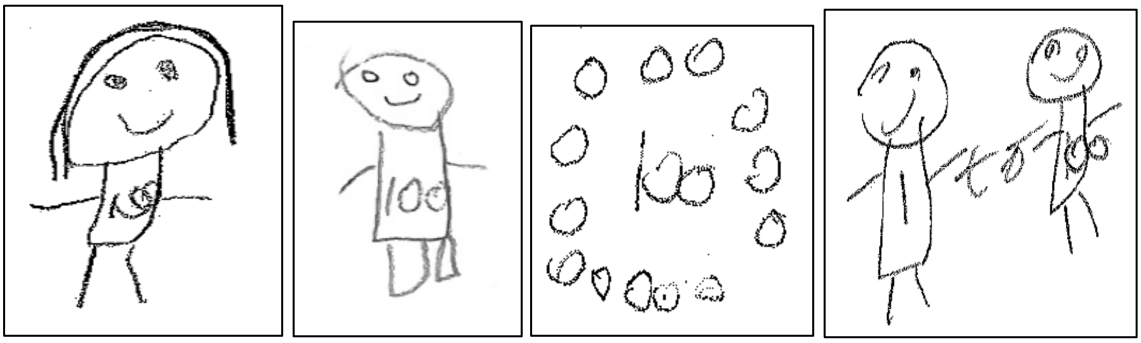


Figure 88 Marta's T1 inscriptions. Left to right: R1, R2, R4 and R5.

The component signs of Marta's T1 inscriptions are notational (featured in each drawing), pictorial (R1, R2 and R5) and iconic (R4).

During R3, Marta did not make any inscription on paper for T1. The exchange with the interviewer went as follows:

J: What do you see when you imagine the counting numbers?

M: Mm.... all of the the numbers going... [*right hand held behind head*]

J: Yes? Do they go in a particular way?

M: [*nodding*]

J: Can you show me with your hands, or tell me how?

M: [H*olds up LH mid-air in a fist, fingers away from body. Raises up thumb first, then index, then middle finger, then retracts hand*.]

J: Yes?

M: [*nodding*]

…

J: Can you draw it, or is it too difficult to draw?

M: Too difficult.

In this episode, Marta used her fingers to demonstrate an increasing count sequence. It is possible that the invitation to "show me with your hands or tell me" prompted the use of fingers, however, Marta's initial response ("all of the the numbers going…") contains already some dynamic element. This could already be progression through the count sequence (with any mode of representation), although Marta could also mean a sense of watching notational numbers move in another way.

During Tasks 2 and 4, Marta's only written representations were the numerals written into the rockets to answer T4 trials. Gestural representation was common, and almost always within some form of task strategy such as count-on. Other gestural representation and verbal representation occurred infrequently.

8.2.2 What aspects of number structure are represented?

8.2.2.1 T1 representations

According to the framework of Thomas and Mulligan (1995), none of Marta's inscriptions in T1 could be identified as representing number structure. Her purely verbal and gestural response to T1 in R3, however, represented number sequence. In terms of Resnick's microstage theory of number understanding development (Resnick, 1983), the inscriptions included no representation of number structure, whereas the R3 response represented (increasing) sequence structure.

Marta's representations during all tasks were analysed according to Resnick's microstage theory. The following table contains the aspects of number structure in Resnick's theory matched to procedures and strategies. The aspects that Marta represented are shown in normal text, and grey text is used for the aspects that she did not represent.

Number structure	Detailed stages	Procedures or strategies in which this number structure may be represented
S Sequence	Count sequence from 0 or 1	Count-on from LE.
	Count sequence from other start point	• Count-on from non-zero LE. • Count-on from estimate to RE for confirmation.
DS Decreasing sequence	Count sequence reversed	Count-back from RE.
RN Relative numerosity (quantity comparison)	Ability to represent (esp visualise) number's magnitude without counting up to it	• Applying efficient count strategy, using RN to choose shorter available count. • Checking judgements, querying whether an estimate is likely. • Use of endpoints as "anchor" points.
PW Part-whole	Partition and recombine numbers (<20)	Use of "anchor" points based on partitioning the range.
	Use of number bonds to ten	• Checking estimations using number bonds to ten. • Using own midpoint anchor on 0-10 range.
	Multiple partitions of multi-digit numbers	Using own midpoint anchors on 0-20 and 5-15 ranges.

PW10 Part-whole understanding of base ten structure	Numbers as compositions of tens and units	• Using own midpoint anchor on 0-100 range. • Count-on and count-back by tens. • 10x10 array structure of 0-100
	Multiple partitions of larger multi-digit numbers	• Count-on or count-back by any given number. • Partitioning into quarters, thirds, other.

Figure 89 Aspects of number structure in Resnick's theory represented by Marta

Although Marta represented aspects of each category of number structure at least once during the five interviews overall, her representations were overwhelmingly focused on the left endpoint of the number lines, and on increasing number sequence. This is illustrated in Figure 90, which shows the relative frequencies with which Marta represented each category of number structure across the five interviews. The circles are coloured and sized to represent the varying frequencies with which representations occurred, relative to all the other frequencies in the chart.

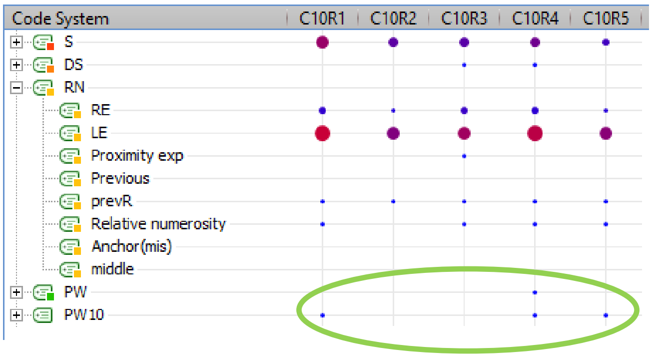


Figure 90 Frequency of number structure representations by Marta

The dominance of references to the left endpoint and increasing number sequence is clear, and in fact the points indicating representation of part-whole structure of any kind (circled together in green) each represent less than five occurrences (across all five interview rounds).

An example here is used to show the kind of episodes from Marta's interviews that were inferred to represent different aspects of number structure. The extract is taken from Marta's R1 interview, and demonstrates representation of increasing number sequence, reference to both left and right endpoints, and reference to an inferred midpoint (part-whole structure of 100). The episode occurred during T2, and Marta was asked to estimate the position of the target 49 on the line 0-100.

J: Forty nine. [*offers rocket sticker to M*]
M: [*grins. Looks at rocket then LE. Takes rocket and counts on from LE; reaches RE with 20 unit hops. Stops and takes hand away from line, leans on elbow and sighs, looking at J.*]
…
M: [*looks at J, then rocket in hand. Smiles and takes rocket to midpoint of line, sticks it down*]

The interpretation of this episode is that Marta initially turned to the left endpoint and a count-on strategy in order to represent the target 49 as '49 units on from zero'. When her count sequence reached the right endpoint after only 20 units, Marta was forced to conclude that she would not be able to complete the procedure. Furthermore, she rejected the option of sticking the rocket down where the aborted count ended (adjacent to the right endpoint), suggesting representation of the relative numerosities of 49 and 100 – which would make such placement impossible. Marta's querying look to the researcher and sigh suggested disappointment at the failure of the first strategy. In contrast, the smile and direct movement to the midpoint, after a pause resting on her elbow, are interpreted as Marta's use of a new representation of 49, as near to half of 100.

8.2.3 What are the notable between-task and within-task connections between representations?

8.2.3.1 Task 1

Marta's T1 representations were presented in RQ1a, and the following table summarises them according to Thomas and Mulligan (1995)'s framework in order to facilitate comparison.

		Interview round				
		R1	R2	R3	R4	R5
Structure, classification from Thomas and Mulligan (1995)	No structure	X	X		X	X
	Linear structure only			X		
	Emerging structure					
	Emerging structure (multiplicative)					
	Partial array					
	Array: 10x10					
	Component signs	Pictorial and notational	Pictorial and notational	Iconic	Iconic and notational	Pictorial and notational
	Static/dynamic	Static	Static	Dynamic	Static	Static

Figure 91 Structure of Marta's representations of number during Task 1 (Thomas & Mulligan, 1995)

In terms of the number structure represented, Marta's T1 representations were highly consistent: the inscriptions represent no number structure in any round, and the verbal and gestural representation of R3 represents sequence structure, which is represented very frequently indeed in the other interview tasks of every round.

Marta's representations in T1 were also very consistent in component signs. For each of R1, R2 and R5 (interviews spread over 9 months) the inscription Marta produced showed striking similarity to the inscriptions of other rounds. In each there is pictorial representation of one or two people, with notational representation of “100” on the torso. In the R4 inscription, there is no pictorial representation of a person, but the notational representation of 100 (only) is again in common with the other inscriptions.

The principal connection evident between the T1 representations and T2 and T4 representations is the low level of representation of number structure (compared with the representations produced by all children interviewed). In T1, the only number structure represented was increasing sequence, and in T2 and T4 this structure predominates, along with reference to the left endpoint (see Figure 90).

8.2.3.2 Task 2 and Task 4

The previous sections considered connections within T1, and between T1 and the number line estimation tasks (T2 and T4 together). In this section, T2 and T4 are explored in more depth. Comparison between representations in different task conditions reveals the within-task connections that occurred during number line estimations. In R2, Marta did not want to complete T4, so for the purposes of analysing the linear accuracy of target estimations, T2 and T4 were considered separately to allow for consideration of this missing data.

Firstly, Figure 92 summarises the representations in each task condition for all five interview rounds considered together, including Marta's representations whether in gaze, gesture or speech. The representations and strategies coded for structure are once again grouped according to Resnick’s microstage account of number structure understanding (see 4.4.2.1).

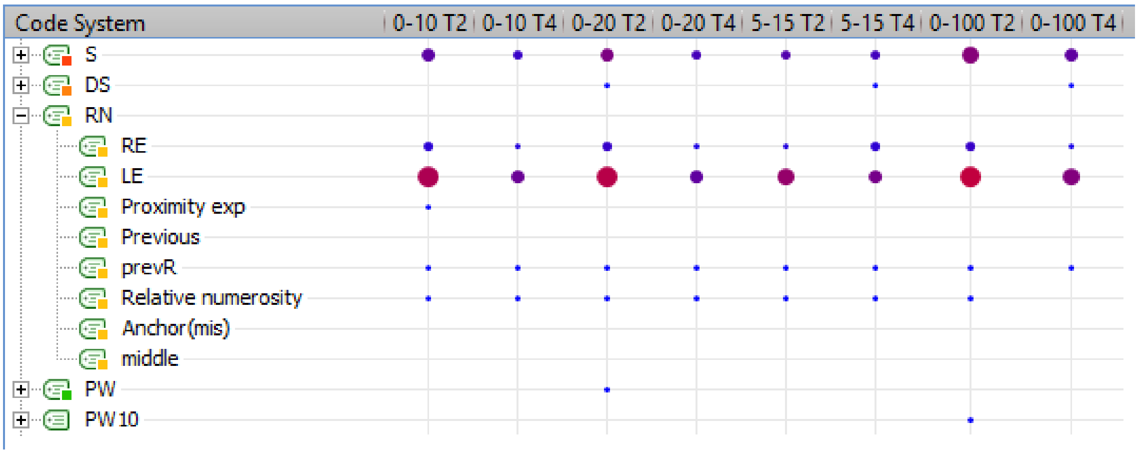


Figure 92 Number structure by task part, all five rounds

Overall, this diagram re-emphasises the finding already discussed in RQ1b: Marta’s representations of number structure were predominantly of the left endpoint and increasing number sequence. These number structures, together with references to the previous trial, and to the right endpoint, were represented in all task conditions. The main difference *between* task conditions that this chart highlights is that for each number range, representations of number structure were more frequent overall during T2 estimations than during T4 estimations. The only task conditions featuring representation of part-whole structure were 0-20 T2 and 0-100 T2.

The charts below summarise Marta’s representations by task condition for each round separately. This was done in order to investigate whether the same within-task and between-task connections were apparent in each interview. As in all three case studies, and noted above, the chart includes Marta's representations from gaze, gesture and speech.

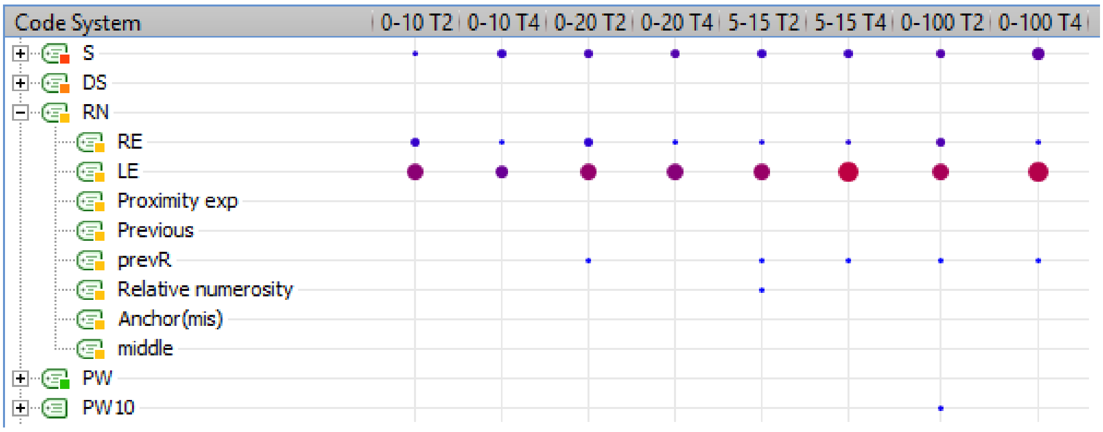


Figure 93 Round 1

In this first round, reference to the left endpoint occurred frequently in every task condition, with no overall difference between T2 and T4 conditions. Reference to the right endpoint occurred less frequently, and more commonly during T2 estimations than T4 estimations. Reference to the previous trial did not occur at all in 0-10 estimations, but at least once in both T2 and T4 for ranges 5-15 and 0-100. Representation of increasing sequence structure occurred in every task condition, but least frequently in T2 0-10, and most frequently in T4 0-100. Representation of part-whole structure occurred only infrequently, and only in T2 0-100.

For R2 (below), the chart omits T4 since Marta completed only four (of 37) T4 trials in this round.

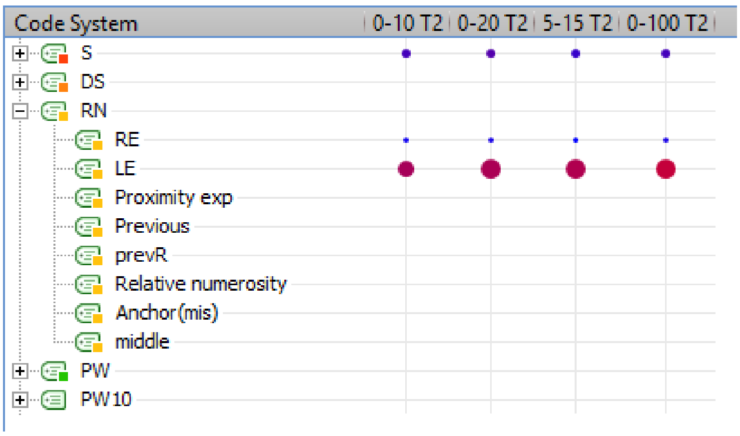


Figure 94 Round 2

Once again, increasing sequence structure is represented in each condition. The only other structure represented is the left endpoint, which is represented with high frequency in each range, and the right endpoint, which is represented with considerably lower frequency. The number structures are represented with almost identical frequency across the four ranges in R2.

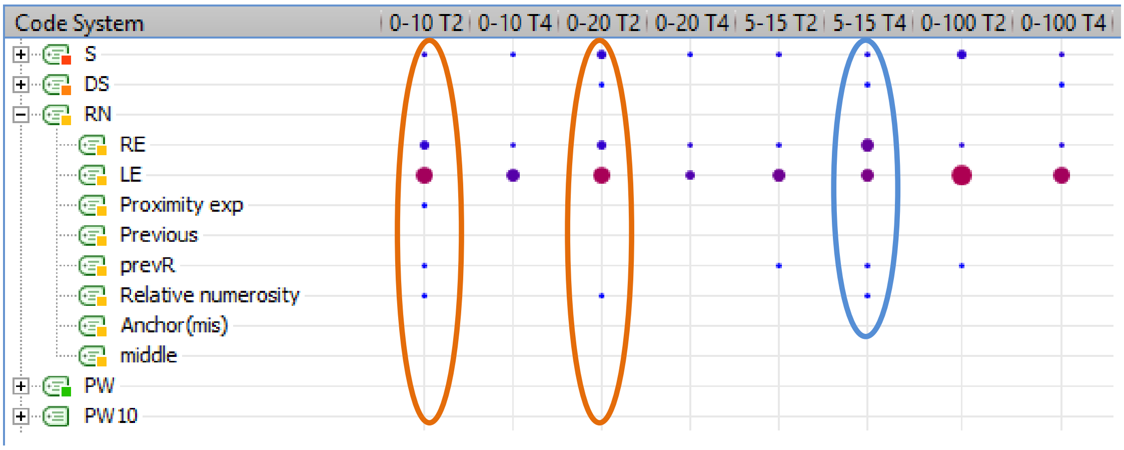


Figure 95 Round 3

In R3, once again representations of right and left endpoints, and increasing number structure, are demonstrated in each task condition. References to the left endpoint are again the most common number structure represented. For T2 estimations in ranges 0-10 and 0-20 (orange), the endpoints are referenced more frequently in T2 than in T4, and representation of previous trials and relative numerosity also occurs in T2 and not T4. For range 5-15, this pattern is somewhat reversed: T4 (blue) features more frequent representations of the endpoints, and representation of more types of number structure, than T2.

Round 3 is the first round in which decreasing number sequence is represented, and the representations occur in T2 0-20, and T4 5-15 and 0-100.

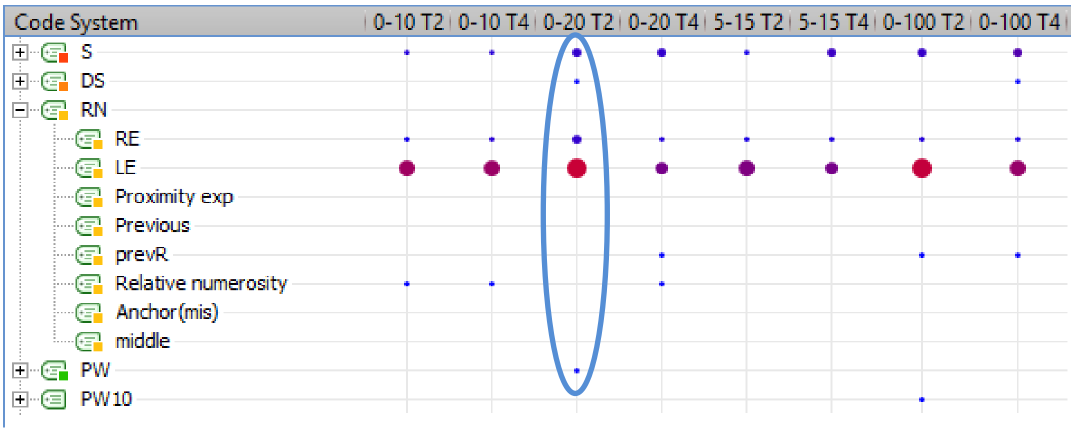


Figure 96 Round 4

Again, in R4 representations of both endpoints and increasing number structure occur in each task condition. References to the left endpoint are again the most common number structure represented. There is no striking difference between task conditions in this round. Representation of the line's endpoints occurs slightly more frequently in T2 estimations compared to T4 estimations, as in R1 and R3.

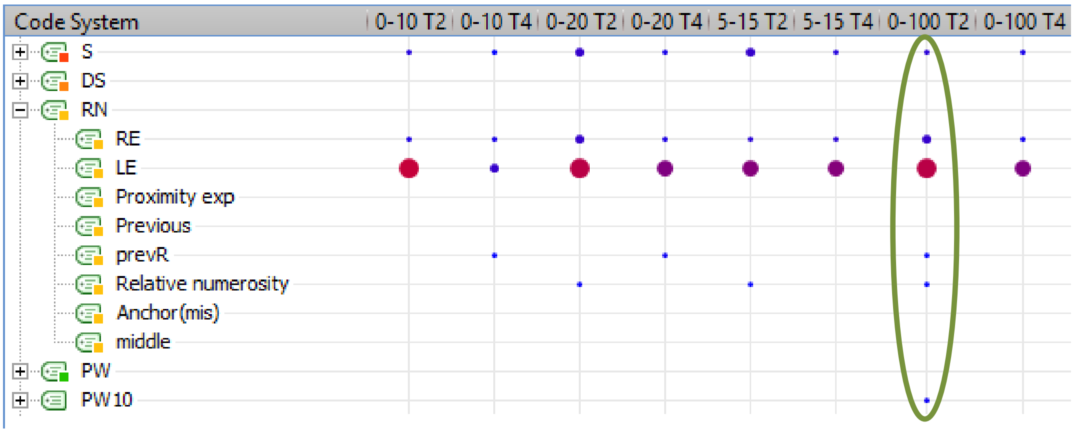


Figure 97 Round 5

In R5, exactly as before, representations of both endpoints and increasing number structure occur in each task condition, are the most frequently represented, and the left endpoint is by far the most frequently represented structure.

As in all previous rounds where T2 and T4 were compared, both endpoints are represented more frequently during T2 estimations than T4 estimations. The task condition featuring the most representation of number structure in R5 was T2 0-100 (circled green).

In summary, these charts show that the distribution of number structure representations between task conditions in Marta's interviews considered all together (Figure 92) was replicated in the individual rounds. Left endpoint and increasing number sequence representations were dominant in each task condition of each round. In rounds where Marta completed both T2 and T4, a slightly higher frequency of number structure representations was observed in T2 compared to T4.

8.2.3.3 *Linear accuracy of representations*

Marta's representations of number during T2 and T4 include the estimations themselves - the representation of target numbers within ranges that each trial requested. The extent to which Marta's representations of targets coincided with the linear representation of number on each given range was measured using the percent absolute error (PAE).

8.2.3.3.1 Task 2

To explore the within-task variation of Marta's T2 estimates, a two way mixed ANOVA was carried out. This compared the PAE of T2 estimates carried out in each of the round x range (5x4) conditions. Levene's test for homogeneity of variances revealed significant differences between the PAE of estimates in different ranges. In order to proceed, data were therefore transformed using a square-root transformation, after which the condition of homogeneity of variances between groups (ranges) was met. All effects in the following analysis are reported as significant or non-significant at $p<.05$.

There was a significant main effect of the interview round on the linear accuracy of Marta's T2 estimates, $F(4,132)=7.14$. Planned contrasts revealed a significant linear trend, $F(1,33)=25.08$, $p<.001$, indicating that the linear accuracy of Marta's T2 estimates increased (PAE decreased) proportionately with each round, as illustrated in Figure 98.

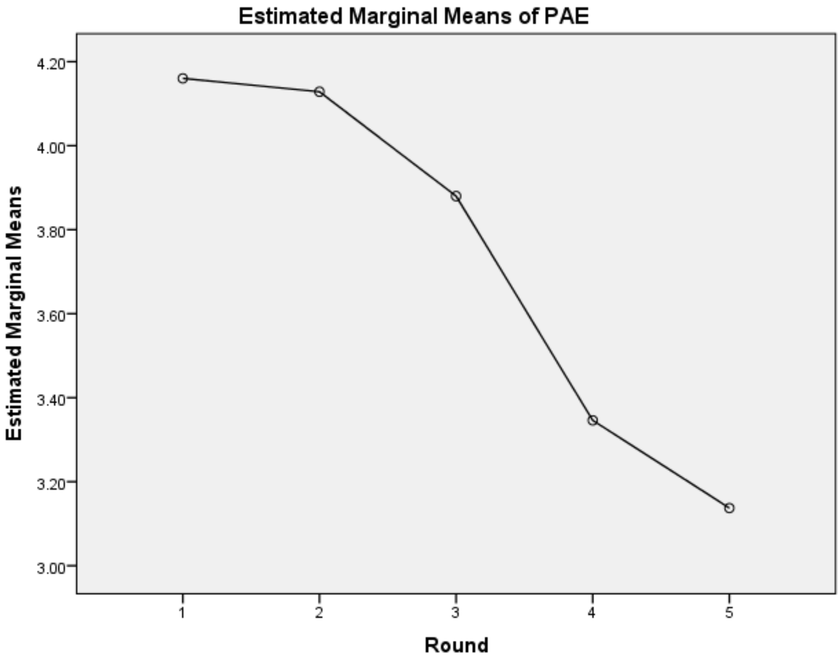


Figure 98 PAE of Marta's T2 estimates

Bonferroni-corrected post-hoc tests showed that the PAE of estimates in each of R1 and R2 differed significantly from the PAE of estimates in R4 ($p=.03$ and $p=.04$ respectively) and from PAE in R5 ($p<.005$ for both). Connecting this finding briefly to the chart of number structure representation by round (Figure 90), there are no striking differences at the level of categories of number structure representation to explain the significant improvement in linear accuracy. More detailed qualitative data, on the estimation strategies used, will be explored in RQ2b.

There was also a significant main effect of range on the linear accuracy of estimates, $F(3,33)=3.26$. Bonferroni-corrected post-hoc tests showed a significant overall difference between the PAE of estimates in the range 0-20 compared to the range 5-15 ($p<.05$), as shown in Figure 99 below.

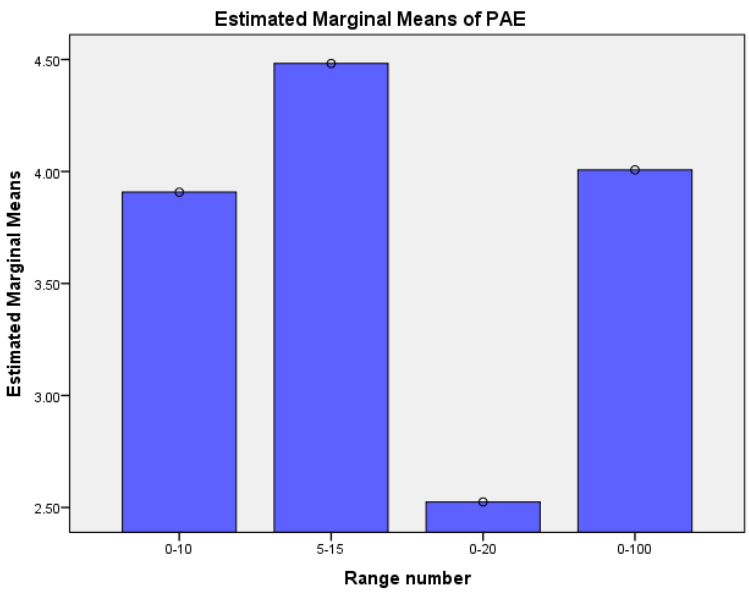


Figure 99 Estimated marginal means of PAE by range

It is interesting at this point to briefly compare the above chart with the chart of number structure representations by task condition (Figure 92). Figure 92 does reveal more representation of number structure in T2 0-20 than in T2 5-15, which the higher linear accuracy for range 0-20 suggests, but the difference is slight. Comparison between the above chart and differences in estimation strategy will be discussed in RQ2b.

There was a significant interaction effect between interview round and range, $F(12,132)=3.18$. This effect confirms the statistical significance of the pattern visible in the interaction graph below, which is that PAE changed by round differently for estimates in different ranges.

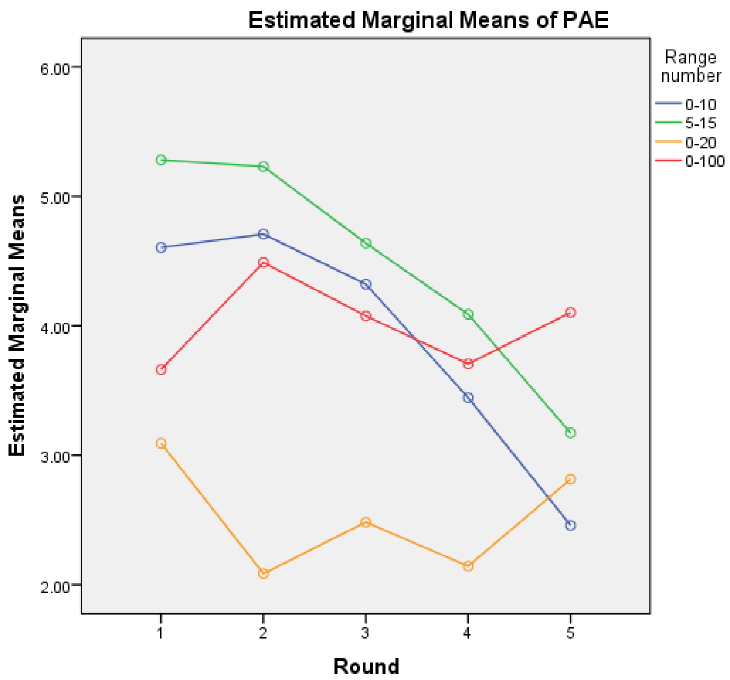


Figure 100 Estimated marginal means of PAE in each round

Most notably, estimates in rounds 0-10 and 5-15 showed negligible change between R1 and R2, before subsequently improving in linear accuracy with each round. For estimates in ranges 0-20 and 0-100 however, no such overall trend is apparent. Current hypotheses on children's number line estimation do not offer insight into why this effect should occur.

The series of charts indicating Marta's representation of number structure (Figure 93-Figure 97) show T2 estimates for range 0-10 and 0-20 featured some representation of relative numerosity in R3, R4 and R5 that was not present in R1 and R2, but the frequency of representation was still low. There was no change in number structure representation between R1 and R2 to give insight into why range 0-20 estimations became suddenly more linearly accurate, and why range 0-100 estimates became less accurate.

As for the main effect of round, investigation of the connection between estimation strategy and estimation result will be discussed in RQ2.

8.2.3.3.2 Task 4

To explore the within-task variation of T4, another two-way mixed ANOVA test was used. As noted before, Marta only attempted four trials of T4 in R2, after becoming tired, so R2 was excluded from the analysis.

Initial testing attempted to compare the PAE of T4 estimates across rounds 1, 3, 4 and 5, and the four ranges. However, Levene's test for homogeneity of variance between groups identified significant differences between ranges in R1. This result reflected a higher variance of PAE in range 0-10 estimations, correlated with a far higher mean PAE as well, as illustrated here:

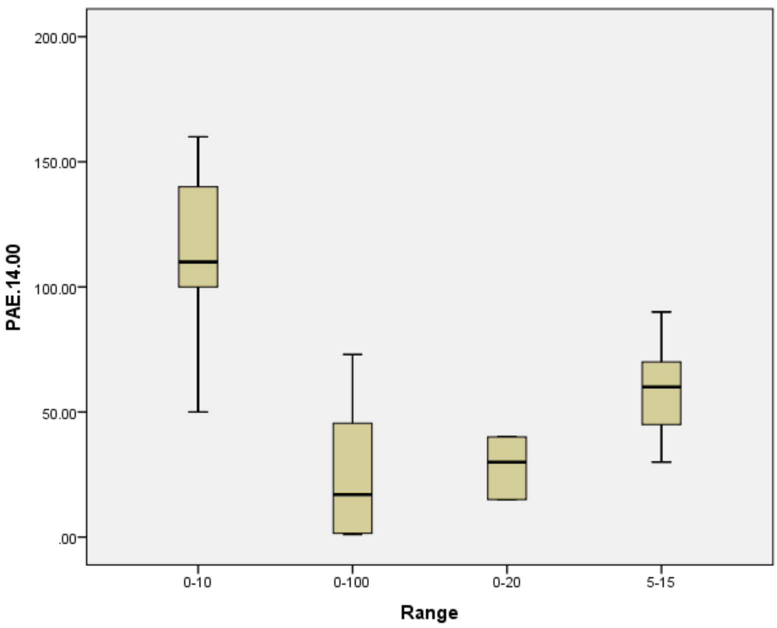


Figure 101 Boxplots showing PAE by range, T4 R1

Examining the graphs of Marta's T4 estimates for each target gives an important insight into this significant difference. In R1 (and only R1) Marta's 0-10 range estimates were such overestimations that they all but one fell outside of the given range, and for 6 out of 9 trials, by more than 50%:

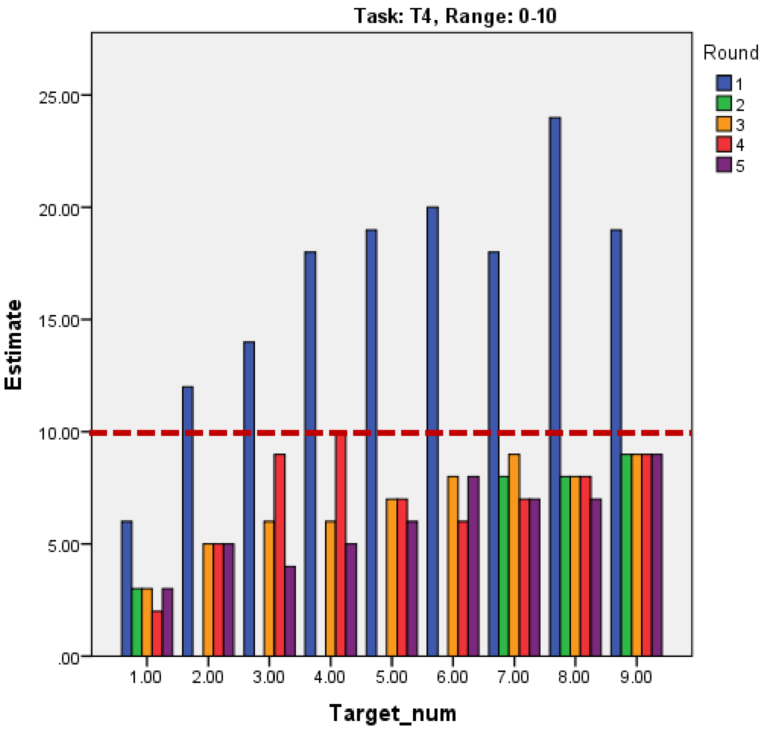


Figure 102 T4 0-10 estimations by round

In contrast, Marta's 0-10 range estimations for later rounds all fell within range.

In all other ranges in R1, Marta's T4 estimates were also within or close to the given range:

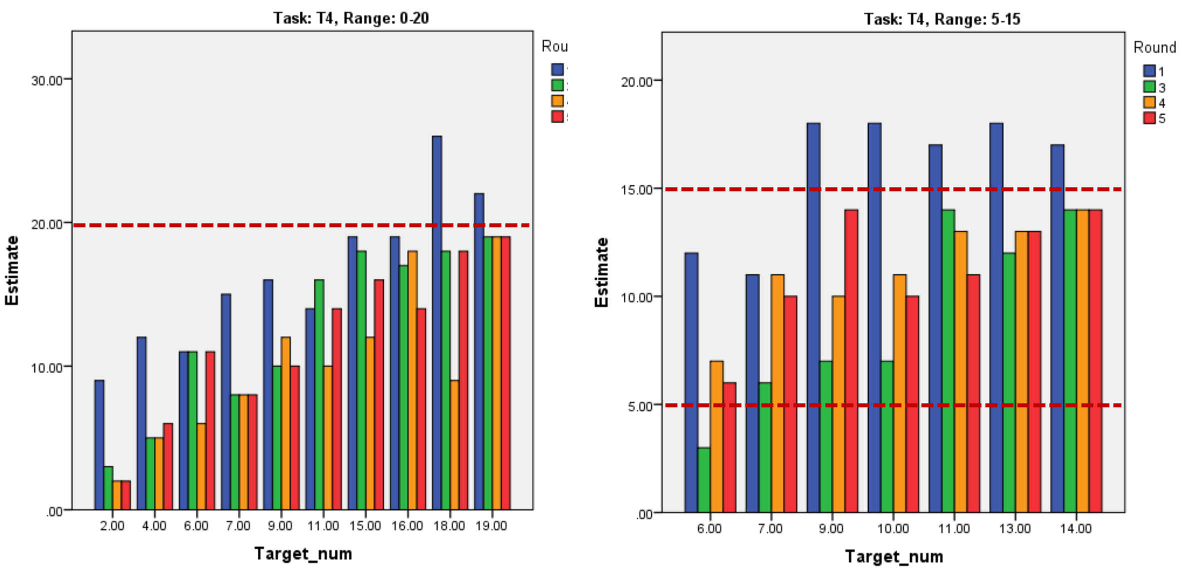


Figure 103 T4 0-20 estimations (above) and 5-15 estimations (above right)

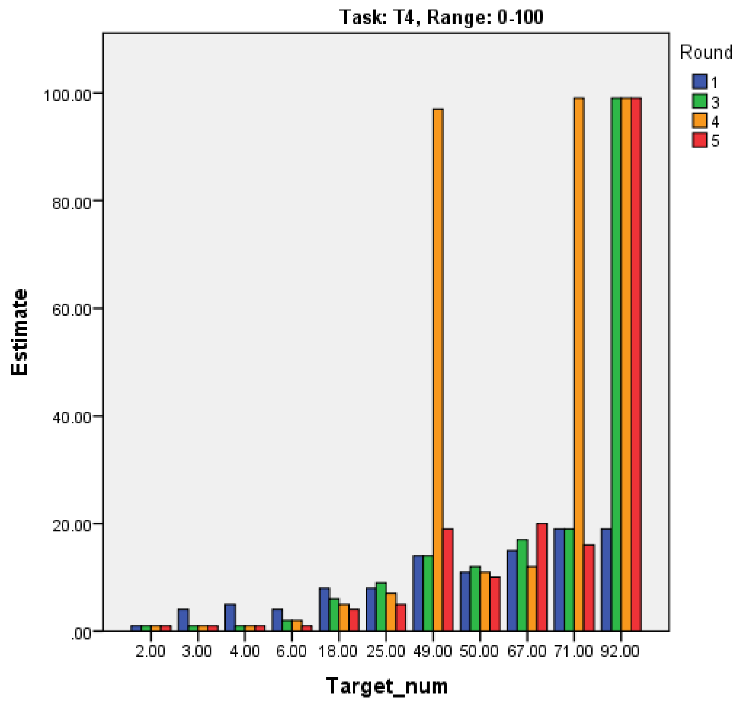


Figure 104 T4 0-100 estimations

This significant difference between the PAE on different ranges in R1 unsurprisingly remained under various transformations of the data, confirming that comparison between the PAE of T4 estimations including R1 was not appropriate. Marta's R3, R4 and R5 estimates in T4 were however suitable for comparison, and so a two-way mixed ANOVA was carried out to explore the between-task and within-task variation for these rounds.

As for Marta's T2 estimates, Levene's test for homogeneity of variance between groups was violated for the untransformed PAE scores (even once R1 was excluded) so the square-root transformation was used again. Mauchly's test was non-significant, indicating that sphericity could be assumed.

The ANOVA testing found no significant main effect of round for the PAE of T4 estimates in R3, R4 and R5. There was also no significant main effect of range, and no interaction effect between round and range.

The plot below shows the mean PAE for each range (T4 only), in each round. As the above testing confirms, the trend revealed is a significant change in estimation PAE for range 0-10 between R1 and later rounds, but little difference between ranges for R3, R4 and R5:

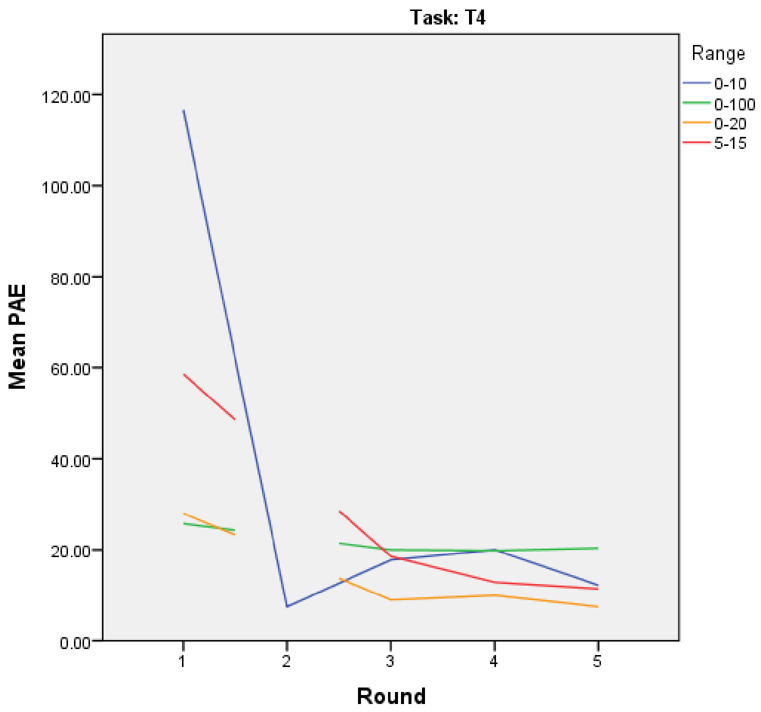


Figure 105 Mean PAE of T4 estimates by range and round

To summarise the between-task and within-task connections in terms of linear accuracy, for both T2 and T4 significant differences existed between ranges. In both tasks, estimations for the range 0-20 were the most linearly accurate, but the linear accuracy of other ranges varied between task and round. In all but one task condition (T2 0-100), estimations were more linearly accurate at the end of the year (R5) than at the beginning of the year (R1), but the trend of change was different for the two different tasks and for different ranges. In T4, linear accuracy did not change significantly between R3, R4 and R5 for any number range.

8.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks?

The strategies identified in Marta's T2 and T4 responses were the following:

- Reference to anchor point
    - Those included in the task environment (left and right endpoints)
    - Those created or visualised by child (midpoint only)
    - Previous trials of the task (immediately preceding trials only)
- Counting strategies
    - Count-on from LE to estimate
    - Count-back from RE
    - "Count-on queried" (see discussion below)
- Judgment using relative numerosities

- Estimate and/or further strategies apparently influenced by relative numerosity of target and task environment
- Mathematical contradiction spotted using knowledge of relative numerosities, leading to change of estimate or change of strategy

A response seen frequently in Marta's interaction with the number line tasks, and not in other children's interactions, was the *apparent* use of a counting strategy, combined with evidence strongly suggesting that the count was not the main influence on Marta's actual estimate. This task response forms the category above named "count-on queried". It was a frequent task response from Marta and combined representation of the count sequence with other behaviour undermining the interpretation of the response as a genuine count strategy.

Several factors were involved in the inference of "count-on queried" examples. In most cases, it was a combination of the factors listed below that led count-on representations to be coded as "count-on queried". The factors were:

1. Number of units counted-on having no mathematical relationship to target or estimated value
2. Long pause and further task response between count-on and giving the estimate
3. Representation of other number structure during count-on coinciding with change in the speed/unit size/direction/continuation of the count

An example combining factors (1) and (3) above is the following extract from Marta's R3 interview. The target to be estimated is number nine, in the range 0-10:

M: [*eyes to right endpoint, then left endpoint*]
J: Number nine [*holds out number nine rocket sticker towards M*]
M: [*takes sticker, looks at it; with sticker in hand, counts-on from LE x9 small hops, mouthing numbers 1-9 silently. Pauses. Eyes look to rocket in hand, then LE, then glance to RE. Counts-on x4 further hops to reach ~90% along line, sticks rocket*.]

The interpretation of this extract is that Marta began, as she did in many estimation trials, by counting-on from the left endpoint of the line with the rocket sticker in her hand. After counting-on the appropriate number of units (nine) Marta paused, itself a suggestion that she was not satisfied with the result. Marta then checked the target in hand, and the two endpoints of the line. After looking at the right endpoint (ten), Marta was reminded of the relative numerosity (or position in count sequence) of nine and ten, and thus extended the count-on by a further four units *in order that* the count's end would be in a position to reflect the proximity of nine and ten.

The above example is just one of many in which the apparent use of a count-on strategy, when examined closely, is revealed to be more complex. Note especially that Marta did not *abandon* the count-on strategy – for example by stopping and jumping in one go to the right endpoint – but made a change to the number sequence representation (here, extending the sequence by four units for no mathematical reason) which allowed the estimate to still *appear* the result of a count-on from left endpoint.

8.3.1 What patterns can be detected in the way children use or do not use these strategies?

Several patterns in estimation strategy are evident from Marta's case. The chart below shows the representations and strategies which Marta demonstrated in estimation trials, compared by target type: whether the target was close to a left endpoint, midpoint, right endpoint, or quarter/three-quarter point of the number line.

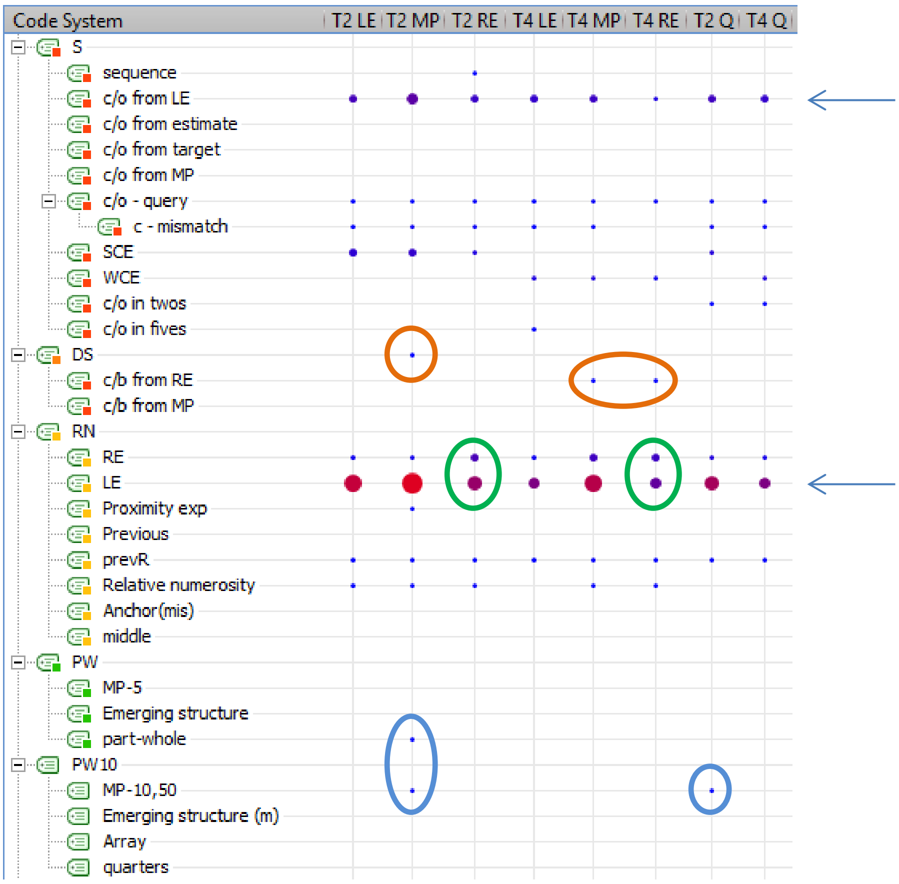


Figure 106 Marta's strategies and representations by target type

One important pattern in Marta's use of representations and strategies which is evident in the above chart and the series of charts showing representations by round (Figure 93-Figure 97) is the unusual dominance of number sequence representation in the form of count-on from left endpoint, and references to the left endpoint (blue arrows). These appear consistently and with high frequency in Marta's estimations, for both T2 and T4, for all ranges, and whether the target in question is close to the left endpoint, right endpoint, or middle of a given range.

A second important pattern in Marta's case is the consistency with which most of the representations and strategies she demonstrated appeared. The consistent frequency of left endpoint references and count-on from left endpoint strategies has already been noted, but references to the right endpoint, "count-on queried" strategies, and references to the previous trial also occurred very consistently between the target types shown in the chart.

The conclusion drawn from the above data is that Marta made relatively few adaptations to strategy based on the target and task conditions. One adaptation that is evident is the higher frequency of references to the right endpoint, accompanied by fewer references to the left endpoint, seen in targets which belong closer to the right endpoint (circled green). Another adaptation is the use of part-whole representation during T2 for targets close to the midpoint and quarter point only (circled blue). Similarly, count-back strategies (orange) occurred only for targets close to the midpoint or right endpoint. Whilst these examples do suggest adaptation informed by number structure, these categories represent only a small proportion of Marta's overall representations and strategies.

Another important pattern in the qualitative data, which is hinted at by the consistent appearance of count-on strategies but not directly shown, was Marta's propensity to use count-on and left endpoint strategies even in cases where she initially represented an alternative. Marta did achieve increasing linear accuracy over the five interview rounds, including in trials using these basic strategies of left endpoint and count-on, but in multiple cases these strategies led Marta to override what was a more linearly accurate initial response. The following extract illustrates such a situation. The extract is taken from T2, range 5-15, and Marta is asked to estimate the position of target 14:

J: Number ... 14.
M: [*taps line near RE, then jumps to LE. Count-on from LE x9 small jumps, to middle of line. Sticks rocket. Looks to right endpoint*.]

The interpretation of this episode is that Marta initially represented 14 in terms of proximity to 15 (the right endpoint). However, for some reason she did not estimate the target based on this first response, and instead went on to use the count-on from left endpoint strategy. This resulted in a final estimation position near the midpoint of the line, i.e. a less linearly accurate estimate than the original response. Marta's final glance to the right endpoint suggests that the relation of 14 and 15 had not been entirely forgotten, but she did not alter the final estimation position.

8.3.1.1 Differences between task conditions

Differences between number structure representations in each task condition were explored in RQ1c. However, within representations of a particular aspect of number structure can occur in multiple strategies. The following chart shows the estimation strategies used in T2 estimations in each range, for all five interview rounds.

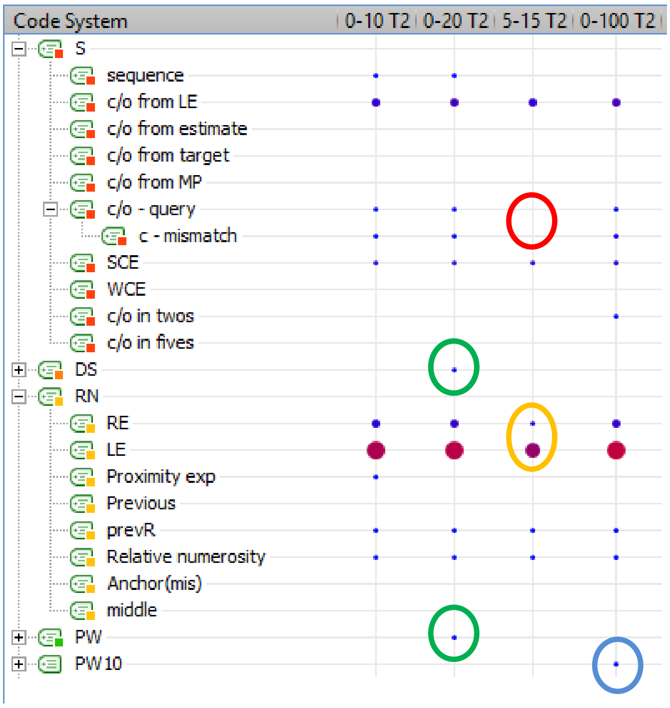


Figure 107 Structure represented during T2 trials

Differences are apparent between ranges. Firstly, for range 5-15, no count-on strategies were identified as "count-on query" or "mismatch" (circled red). In all other rounds, some of Marta's count-on strategies were identified as representations of increasing number sequence, yet not the primary means by which Marta actually determined the position of her estimate. For range 5-15 however, the count-on strategies were straightforward count-on strategies. Secondly, for range 5-15 again, the frequency of references to both right and left endpoints (circled gold) is lower than for all other rounds. Thirdly, estimates in range 0-20 featured the only representation of decreasing sequence (green), and also representation of part-whole structure (green). Estimates for range 0-100 also featured part-whole representation (blue).

8.3.2 How do the strategies used relate to children's estimation results, if at all?

One connection is between the chart above (Figure 107) and the significant difference in linear accuracy between T2 0-20 estimations and T2 5-15 estimations, as reported in RQ1c (see Figure 99). There is a correlation between strategies incorporating representation of more aspects of number structure (T2 0-20, compared to T2 5-15 trials) and more frequent references to the line's endpoints (T2 0-20 compared to T2 5-15), with higher linear accuracy of estimation. Most importantly of all, the absence of “count-on queried” strategies for range 5-15 was associated with lower linear accuracy. In other words, where Marta’s count-on strategies *were* “count-on queried” and not in fact the primary means of estimation (ranges 0-10, 0-20 and 0-100), the resulting estimations were more linearly accurate.

Another finding from the analysis of linear estimation accuracy carried out in RQ1c was the main effect of interview round: Marta’s T2 estimations in later interview rounds were significantly more linearly accurate than in earlier rounds.

The following chart plots Marta’s T2 representations and strategies by interview round, to see what connection between strategies and estimation accuracy are suggested.

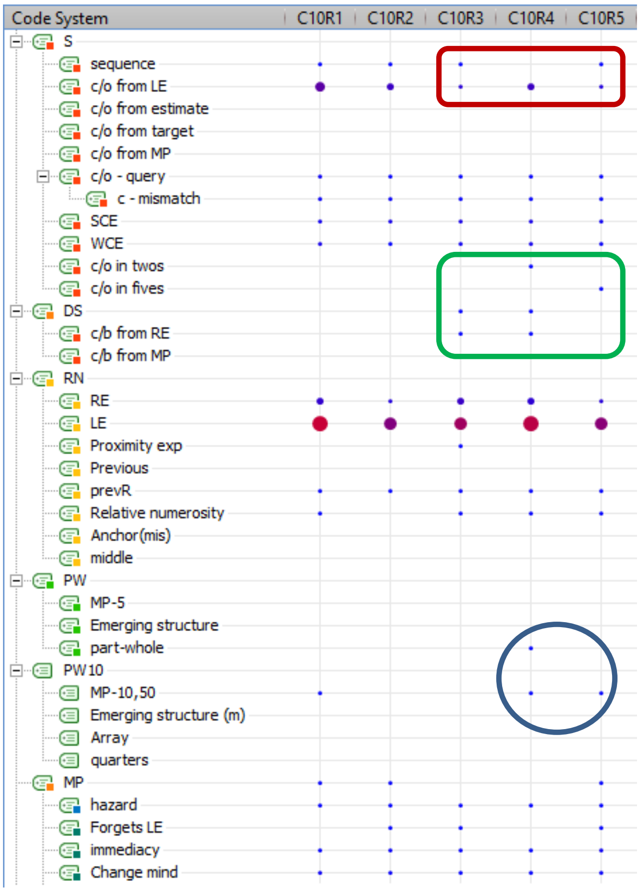


Figure 108 Marta's estimation strategies and representations by round

The chart highlights three changes between earlier and later rounds that might be involved in changing estimation accuracy. Firstly, the frequency of representation of increasing sequence and count-on from left endpoint strategies is slightly lower in later rounds than in the first two rounds (see red box), although this difference is small. The proportion of count-on strategies that are “count-on queried” remains even across rounds. A second more interesting difference is that more structurally complex aspects of number sequence – counting in twos and in fives, and counting-back – are only represented in later rounds (see green box). The third and final difference is that more part-whole structure is represented in later rounds (circled blue) than in earlier rounds. Overall, Marta's more linear estimations in later rounds are associated with representation of more complex number structures.

In the remainder of this section, visual plots of estimation accuracy by target are used to identify interesting patterns in estimation accuracy, especially large disparities between rounds and between targets in the same range. After identifying points of investigation, the plots are then compared to charts showing the strategies and representations that occurred during the relevant estimation, to see whether connections can be drawn that may explain some of the estimation results.

8.3.2.1 T2, range 0-10

The following chart shows Marta's estimates for T2 in range 0-10.

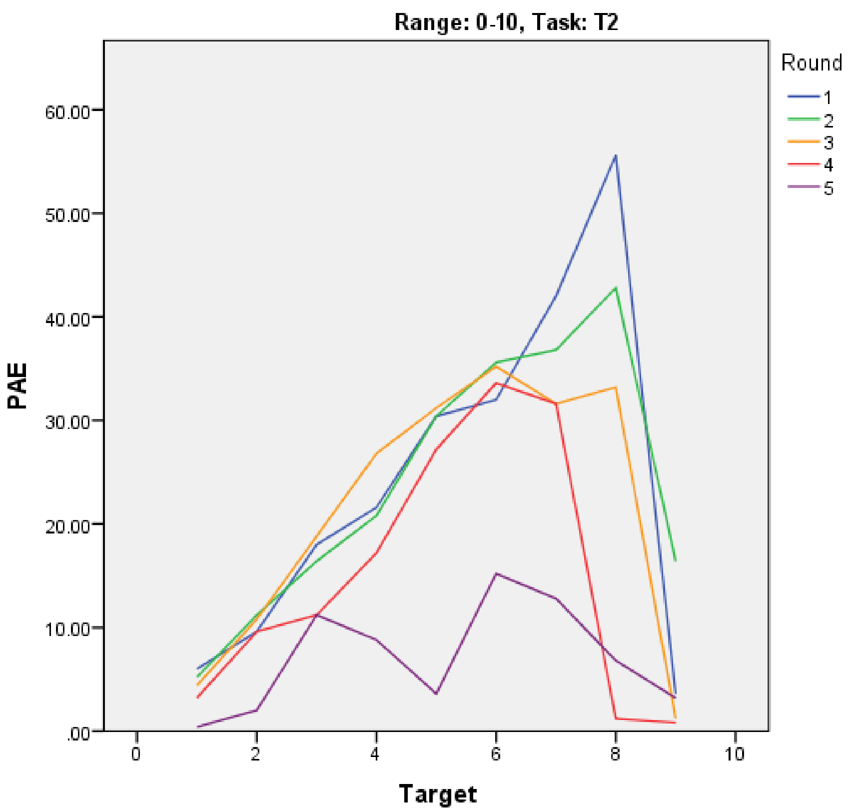


Figure 109 PAE of target estimates, T2 0-10

This graph reveals interesting between-rounds variation in the pattern of PAE by target. The graph shows that T2 estimates on 0-10 were accurate for target 1, and then in the first four interview rounds, that PAE increased linearly with target until target 8 or 9. In R1, R2 and R3, the PAE pattern suggests that the right endpoint was used as an anchor point for target 9, and in R4, PAE suggests an anchor point was used for both targets 8 and 9. In R5, PAE is low for targets near the left endpoint, midpoint and right endpoint, suggesting that all three of these were used as anchor points to aid estimation. The questions arising to be answered by the qualitative data are:

- Is there evidence for use of RE as an anchor for estimates of 9 (all rounds) and 8 (R4 and R5)?
- Is there qualitative evidence of a left endpoint and midpoint strategy being used in R5?

The charts below show the representations and strategies used in T2 0-10 estimations, firstly for R1-R4 inclusive, and then separately for R5. In answer to the first question, there is good evidence for use of the right endpoint as anchor point in all rounds. This can be seen from the references to right endpoint which are found for target 9 (circled green), but not at all, or infrequently, for all other targets on this range. In addition, although there are count-on from left endpoint strategies recorded for target 9 in R1-R4, they are identified as "count-on queried" in actual fact (circled orange). There is however no direct evidence for use of a right endpoint anchor for target 8 in R5: the only representations and strategies observed were count-on from the left endpoint, with the rocket stuck at the count's end.

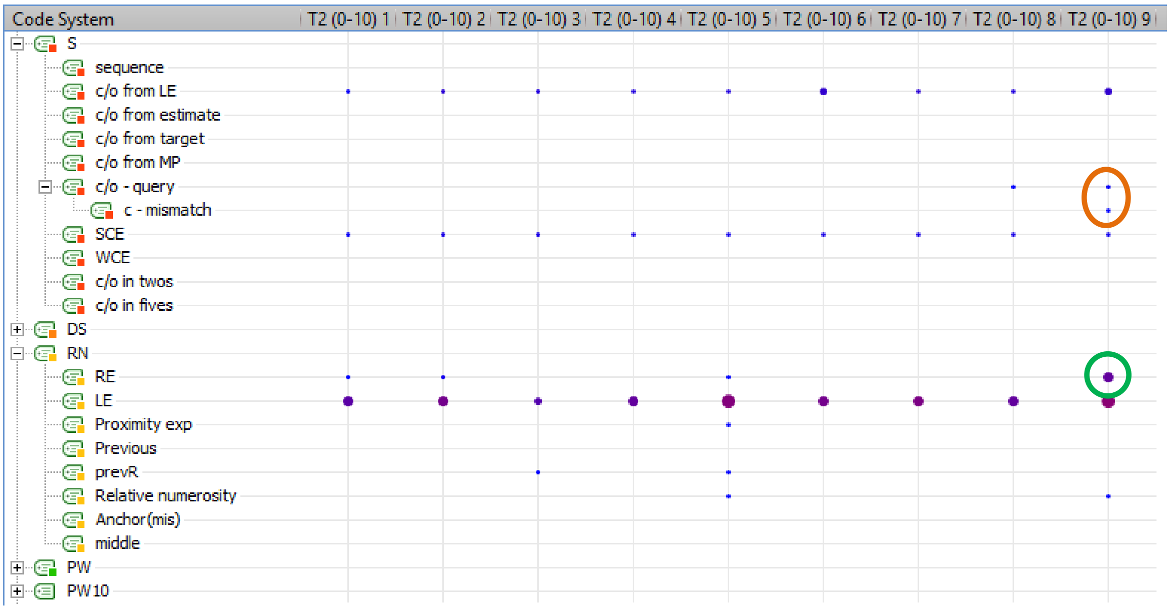


Figure 110 T2 0-10 estimates R1-R4

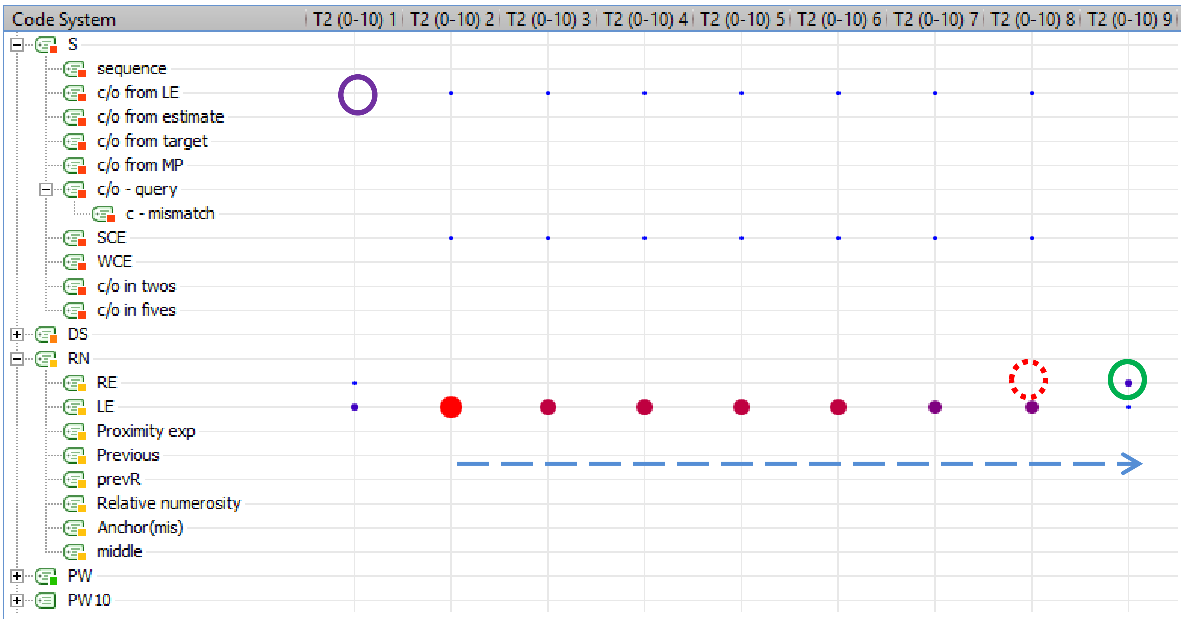


Figure 111 T2 0-10 estimates R5

In answer to the second question, the absence of count-on for target 1 in R5 (circled purple) does suggest the use of a left endpoint anchor for this target. However, there is no direct evidence for a midpoint anchor. As for the majority of targets in this task condition, the representations and strategies observed for target 5 are only reference to left endpoint, and count-on from left endpoint.

The PE graph of the same task condition reveals that for the first four rounds, Marta's T2 estimations for range 0-10 were all underestimations, which follows logically from her physical strategy of counting on from the left endpoint with small 'hops', smaller than one tenth of the line length. In R5, the PE of estimates resembles an S-shaped pattern of under- followed by over-estimation. A one-cycle proportion judgement model (e.g. Rouder and Geary, 2014; Slusser et al., 2013) also predicts an S-shaped curve of PE. However, for T2, a number-to-position task, a proportion judgement model would predict over- followed by under-estimation centred around an anchor point (here, the midpoint – the point after the endpoints with the lowest estimation error), rather than the under- followed by over-estimation seen here.

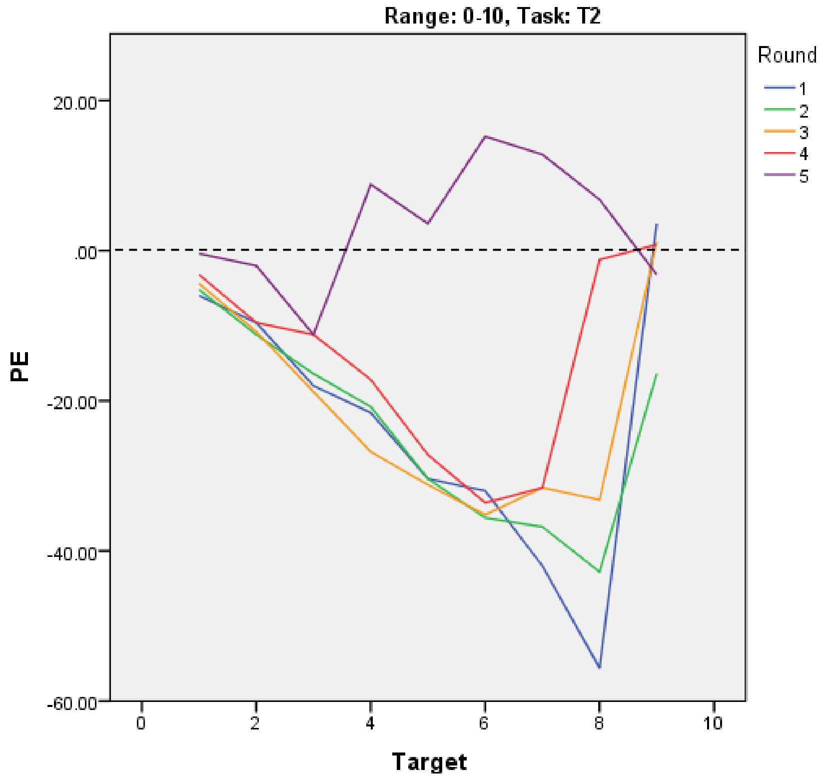


Figure 112 PE of T2 estimates, range 0-10

The conclusion from the above findings is that the S-shaped pattern of under- then over-estimation seen in Marta's T2 0-10 estimations in R5 is better explained by appeal to the observed representations and strategies than by a proportion judgment model.

8.3.2.2 T4, range 0-10

The graph below shows the PE of estimates for T4 range 0-10.

Figure 113 PE of T4 estimates, range 0-10

The most notable feature of this graph is the extremely high PE of estimates in R1 compared to later rounds. The lowest estimation error in this task condition occurred for targets 8 and 9. The two charts below display the representations and strategies for T4 range 0-10, firstly for R1, and secondly for R2-R5 inclusive. The most striking changes for individual targets between rounds are the changes in PAE for target 8 and 9, between R1 and all later rounds.

The following charts plot Marta's representations and strategies for R1 separately, and then R2-R5 inclusive, in order to investigate the above points.

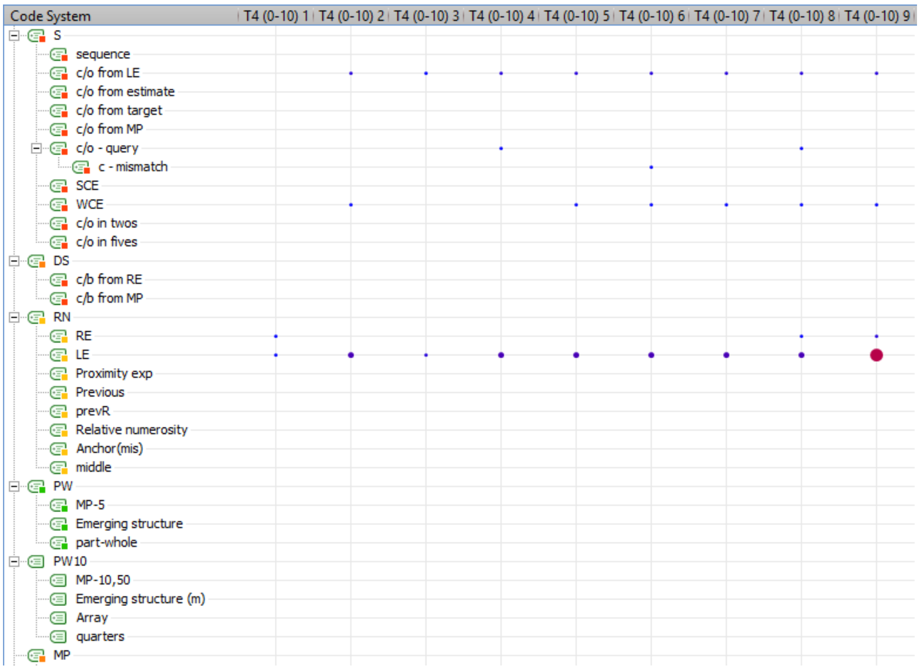


Figure 114 T4 0-10 R1 estimates

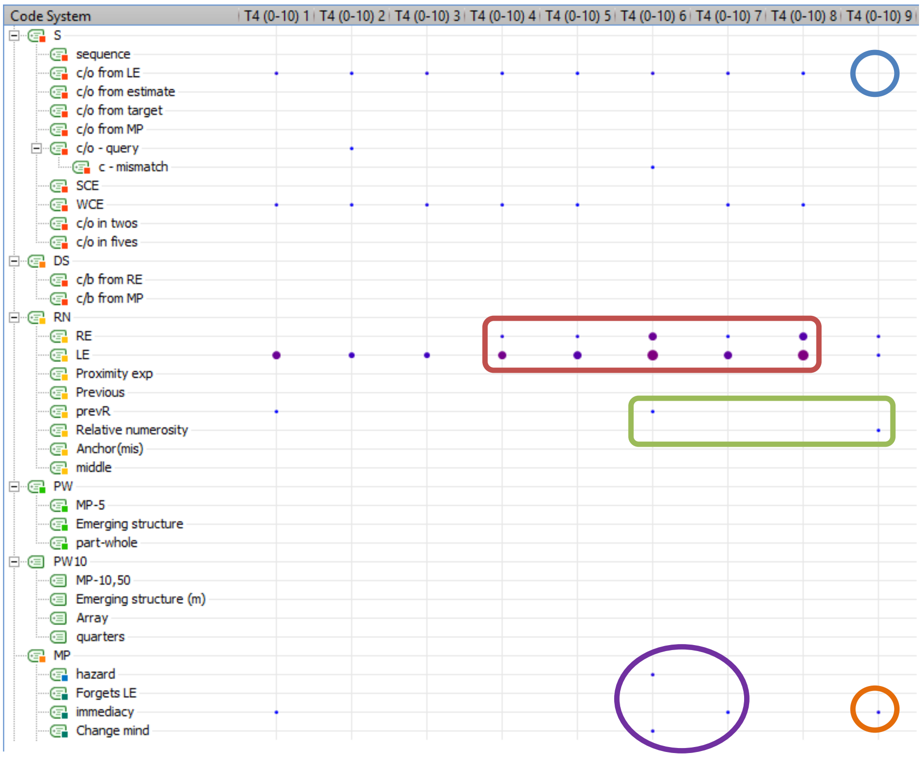


Figure 115 T4 0-10 R2-R5

Looking at what changed between R1 and later rounds for targets 8 and 9, the first key difference is that in later rounds there are no instances of count-on from left endpoint for target 9 (blue circle). There are also fewer references to each endpoint for target 9 in later rounds, and some signs of answering immediately (orange). Together, these findings suggest good evidence that a right endpoint anchor was being used to estimate target 9 for R2-R5. For targets 4, 5, 6, 7 and 8, later rounds show *more* references to the endpoints, and crucially, the right endpoint as well as the left endpoint (see red box). These later rounds also reveal some representations of relative numerosity (green box) and occasions on which Marta encountered mathematical contradictions ("hazard" code) and changed her mind from a count-on strategy (circled purple). The targets 4, 5, 6 7 and 8 where these differences apply were all markedly more accurately estimated in the later rounds than in R1.

8.3.2.3 T4, range 0-20

The graph below plots the PAE of Marta's T4 estimations for range 0-20. The first point to note is that there is no data for R2, since in this round Marta became tired and stopped T4 before reaching range 0-20. For most rounds, the estimations are most accurate close to the left endpoint and close to the right endpoint. The most interesting disparity between rounds is for target 18: in R3 and R5, the estimation of both 18 and 19 is very accurate (0% error), but in R1 and R4, there is a spike in PAE for target 18, before a comparatively accurate estimate for target 19.

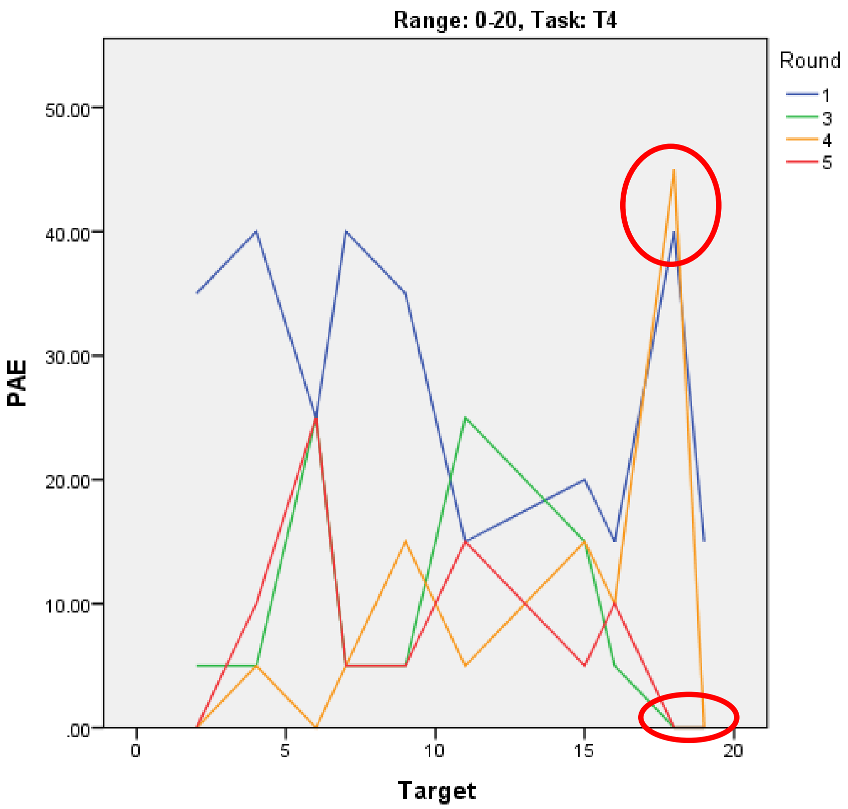


Figure 116 PAE of target estimates, T4 0-20

The charts below therefore plot Marta's representations and strategies for R1 and R4 together, and then for R3 and R5 together, to compare observations for targets 18 and 19.

The first difference to note is that in R1 and R4 (low estimation accuracy) but *not* in R3 and R5 (high accuracy), Marta used count-on from left endpoint strategies for targets 18 and 19 (circled red on both charts). Despite also changing her mind about some aspects, and at least one observed count being "count-on queried", for both targets Marta wrote her estimations at the count end (circled yellow). In the estimation of 18 and 19 in R3 and R5, there are fewer references to the line's endpoints (blue) although both right and left are referenced, and it was also observed that Marta made her estimations with immediacy (circled green). This evidence strongly suggests that a right endpoint anchor was being used in R3 and R5, whereas in R1 and R4, Marta relied heavily upon counting-on from the left endpoint.

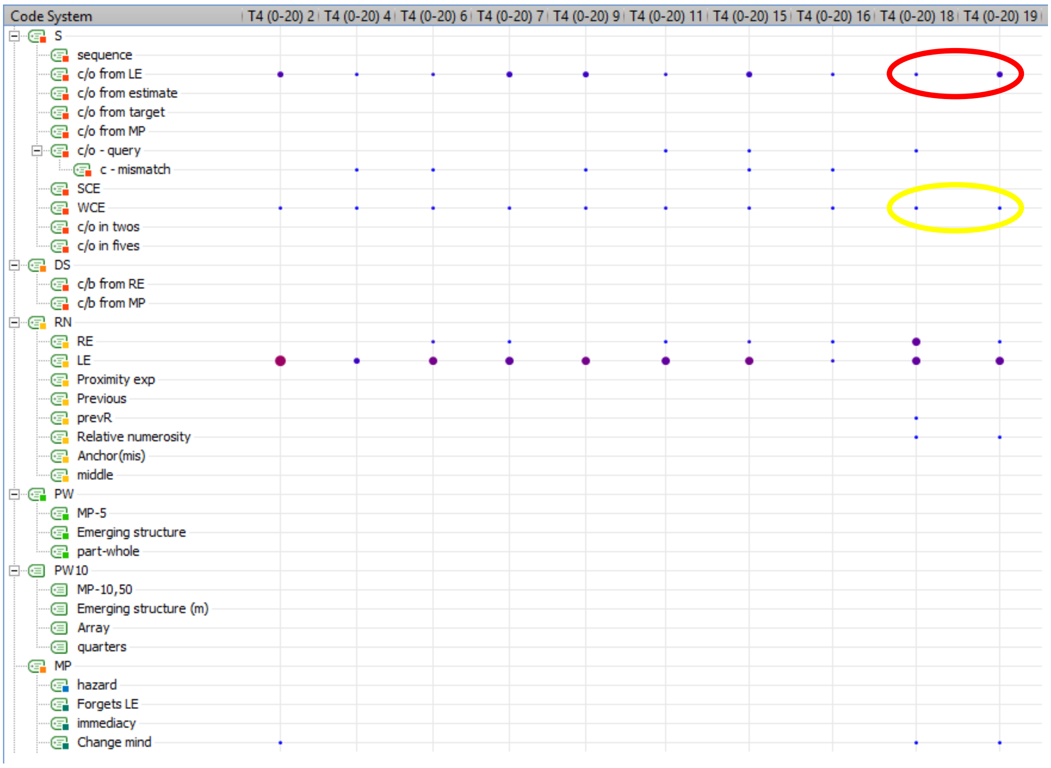


Figure 117 T4 estimations, range 0-20, R1 and R4

Figure 118 T4 estimations, range 0-20, R3 and R5

8.3.2.4 T2, range 5-15

This graph of linear estimation error for T2 estimates on 5-15 shows low PAE (high accuracy) for target 6, close to the left endpoint, followed by linearly increasing PAE for each subsequent target. In R3 (gold line) there is some evidence to suggest that the right endpoint was used as an anchor point for estimations of targets 13 and 14. In R2, there is also a drop in PAE for target 14. In R5, lower PAE at targets 10 and 14 suggests that a midpoint anchor *and* right endpoint anchor may have been used.

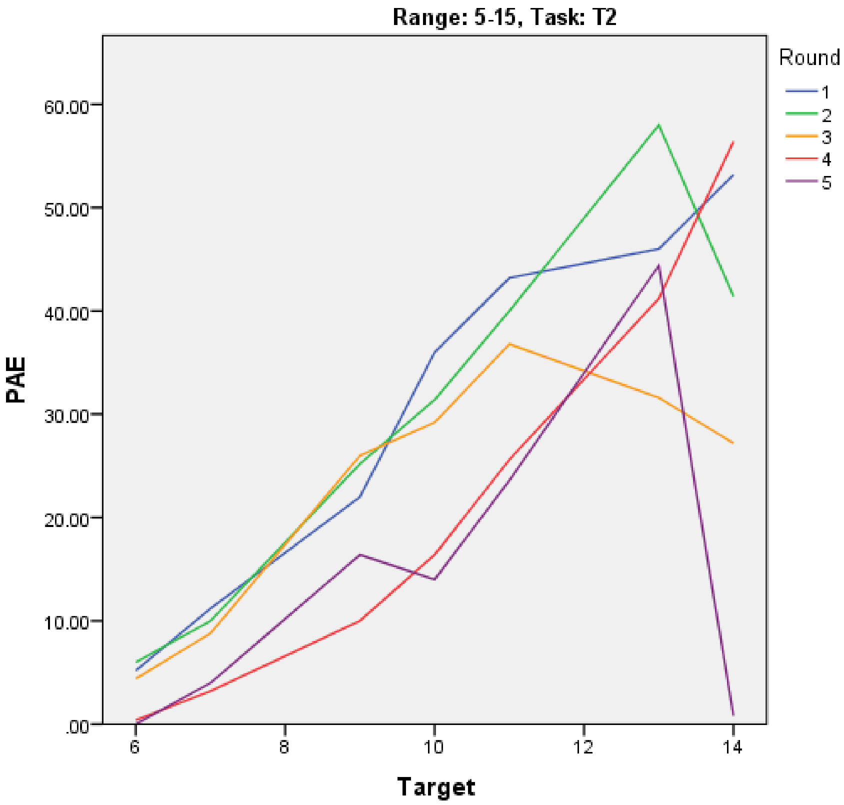


Figure 119 PAE of target estimates, T2 5-15

The questions to be answered by the qualitative data are the following:

- Is there any evidence for use of a right endpoint anchor for target 14 in R2, and targets 13 and 14 in R3?
- Is there evidence for the use of midpoint anchor in R5?
- Is there evidence for the use of a right endpoint anchor for target 14 in R5?

Since the R1 and R4 estimations for this range show similar patterns of PAE (both linearly increasing with target number), Marta's representations and strategies for these two rounds together are plotted first.

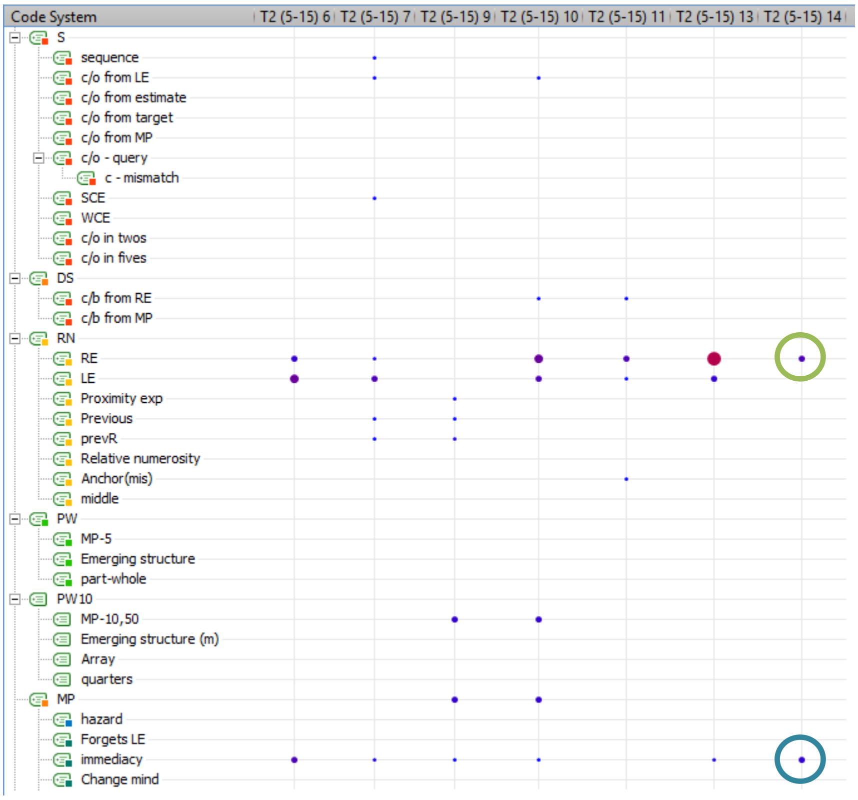


Figure 120 T2 estimation, range 5-15, R1 and R4

For estimations of target 14 in R1 and R4, Marta referenced only the right endpoint (green), and with immediacy (blue). This combination suggests a right endpoint anchor. The observations for target 13 are also suggestive of a right endpoint anchor. The target 13 and 14 estimations in these two rounds were however highly linearly inaccurate. If Marta did indeed use a right endpoint anchor for targets 13 and 14 in these rounds, it did not help her generate linearly accurate estimations.

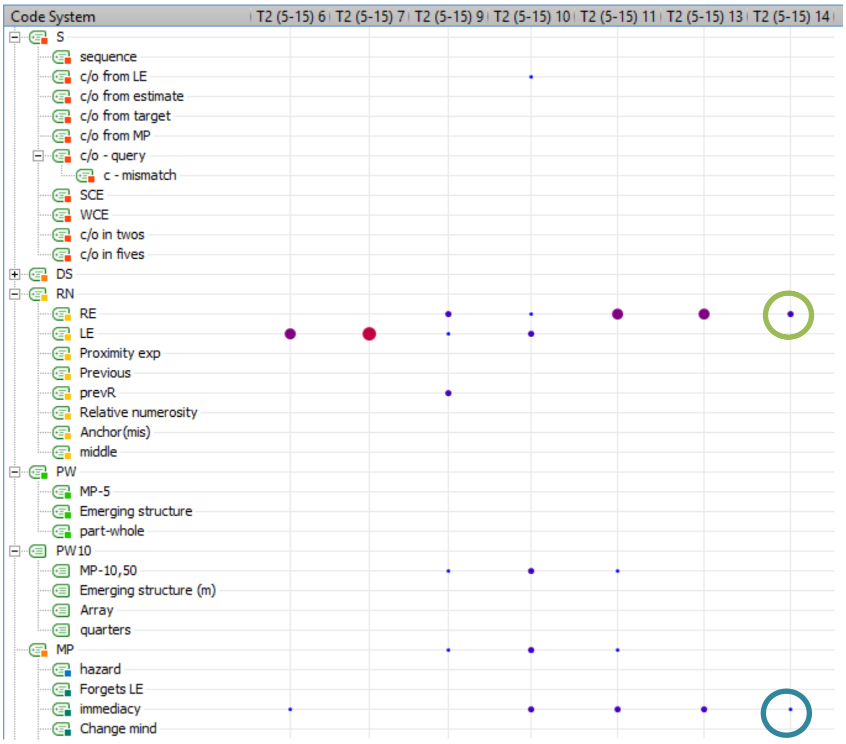


Figure 121 T2 estimation, range 5-15, R2

This chart plots Marta's representations and strategies for the same task condition in R2. There is **no change** for target 14 compared to R1 and R4: once again, only reference to the right endpoint and immediacy of estimation are observed. As in R1 and R4, above, the observations for both targets 13 and 14 suggest a right endpoint anchor. However, whilst Marta's target 14 estimate in this round is more accurate than in R1 and R4, her target 13 estimate is less accurate than in R1 and R4.

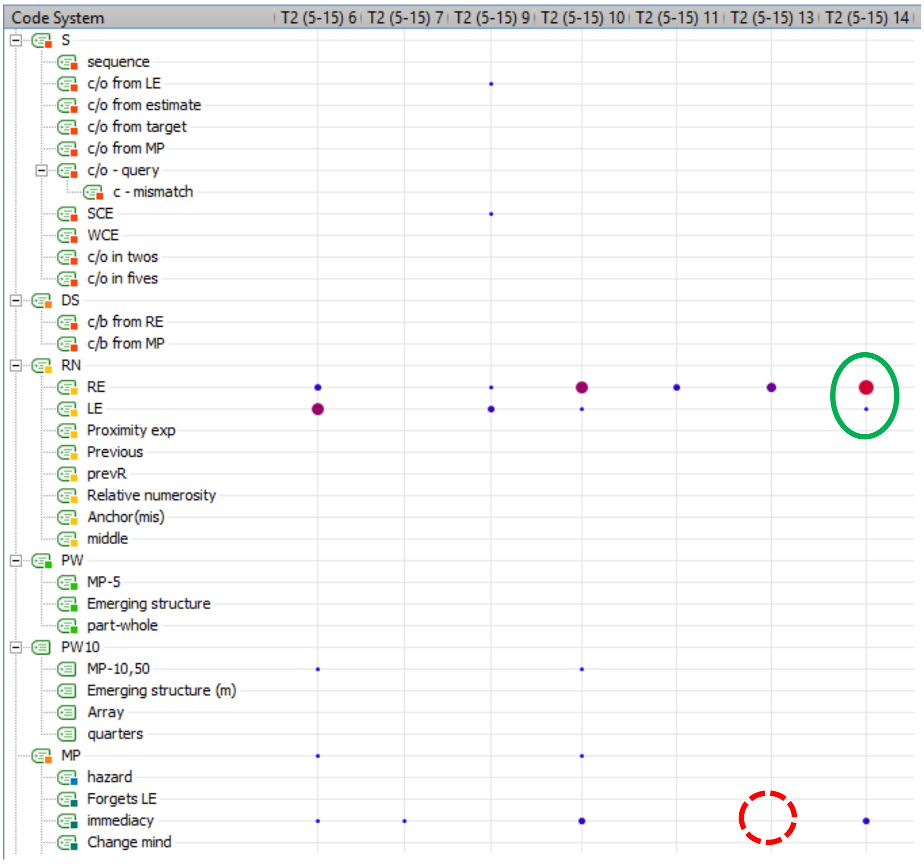


Figure 122 T2 5-15 estimates, R3

The next chart plots Marta's representations and strategies for R3, revealing differences from R1 and R4, and R2. For target 14 this time, there are references to both left and right endpoints, and a much higher frequency of references to the right endpoint than previously. These changes are associated with a much lower PAE for the estimation of 14.

Marta's estimation of 13 in this round is the most accurate of all rounds, and it differs from all other rounds by *not* being estimated with immediacy.

Finally, the following chart shows Marta's representations and strategies for R5. For target 14 in this round, Marta refers only to the right endpoint (as in R1 and R4, and R2) but with higher frequency than in these rounds. Target 14 is also again estimated with immediacy, suggesting a right endpoint anchor.

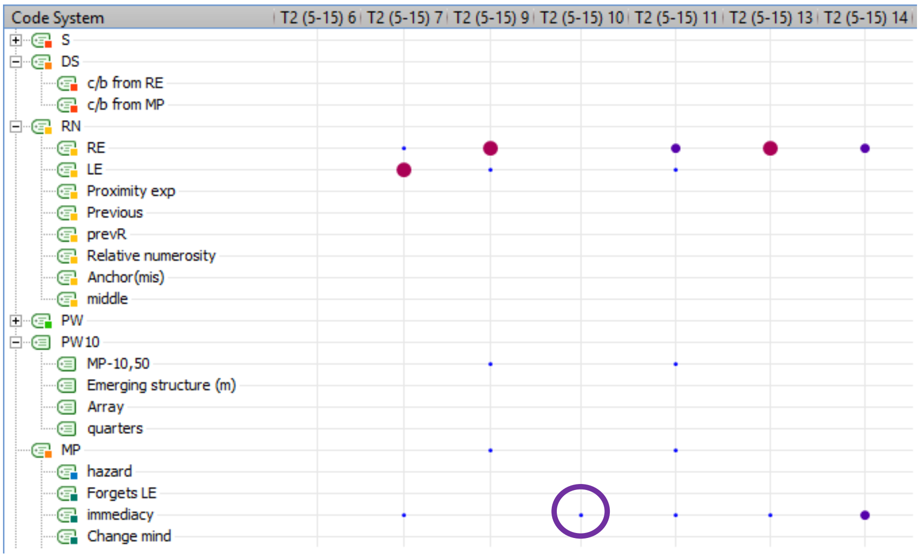


Figure 123 T2 5-15 estimates, R5

For target 10 in R5, which was estimated with lower PAE than in other rounds and neighbouring targets, the only observed feature of Marta's estimation was the immediacy with which she made it. This suggests the use of some inferred anchor point, to make such a rapid decision, but no firm conclusion can be drawn. It contrasts strongly with the observations for target 10 in all other rounds, where Marta represented more number structure but estimated with lower linear accuracy.

This task condition presented some interesting estimation differences to analyse. In conclusion, Marta's representations suggest that target 13 was estimated relative to the right endpoint in every round. It was estimated most accurately when *not* immediately placed (R3). Target 14 was estimated most accurately in R4 and R5, which are the rounds with strongest evidence for use of a right endpoint anchor. However, the same representations were observed (to lesser extent) during the less accurate estimations of target 14 in other rounds too. The evidence for a midpoint anchor in R5 is inconclusive; the immediacy of Marta's target 10 estimation suggests some anchor point was used, but there is no further evidence.

Looking at the following graph of PE (percent error, as opposed to percent absolute error), it becomes clear that *all* of Marta's estimation error on this range for T2 was underestimation. Unlike in other task conditions, this cannot be attributed to small unit 'hops' in Marta's count-on strategies, since as the preceding series of charts demonstrates, Marta employed count-on strategies very seldom in this task condition.

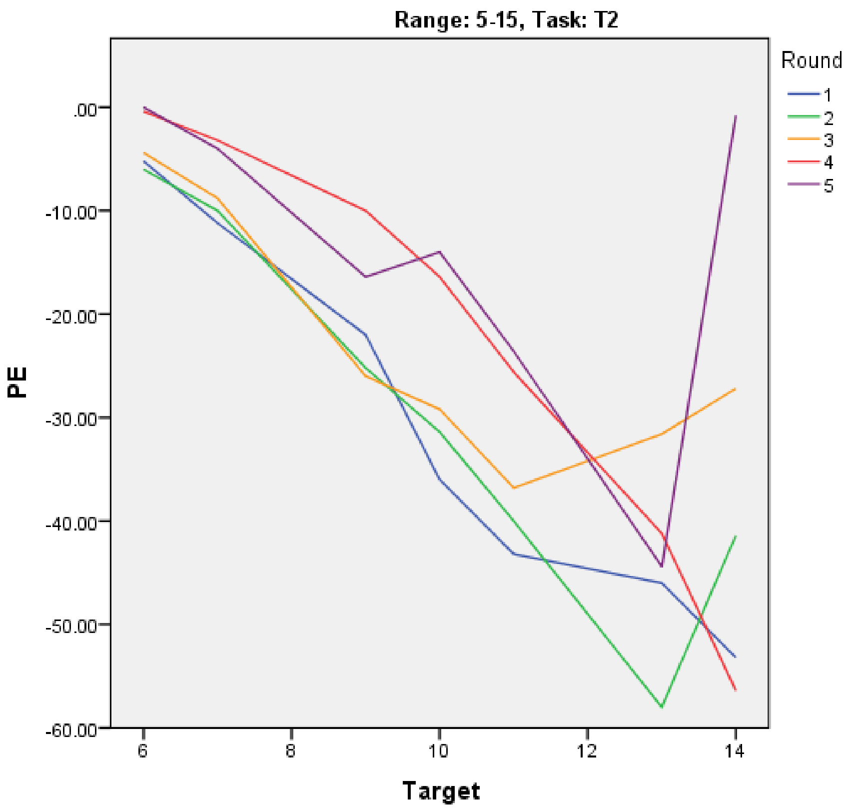


Figure 124 PE of target estimates, T2 5-15

8.3.2.5 T2, range 0-100

The graph below plots Marta's estimates for T2 range 0-100.

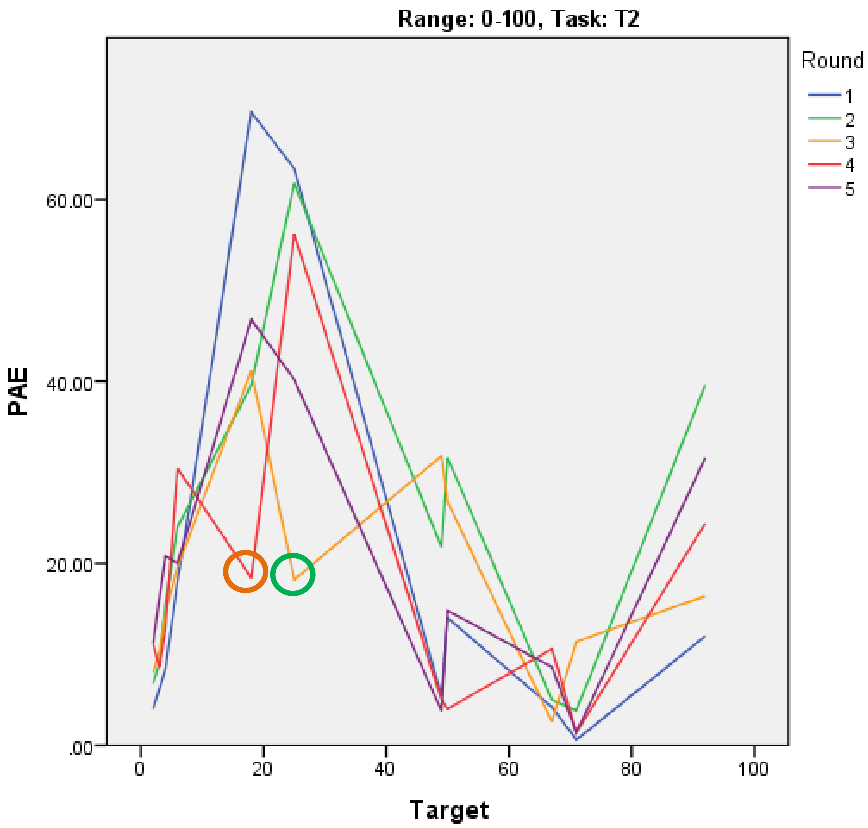


Figure 125 PAE of target estimates, T2 0-100

For this task condition, there are few striking disparities in PAE between rounds. Instead, PAE is largely correlated between rounds. The questions that arise are the following:

- Is there evidence of midpoint anchor being used in R3, R4 and R5?
- Is there evidence that could explain the estimation accuracy for 67 and 71 in all rounds?
- Is there evidence about Marta's estimation of target 92 that could help explain the poor linear accuracy in almost all rounds?
- What was different about the estimation of target 25 in R3 (circled green)?
- What was different about the estimation of target 18 in R4 (circled orange)?

Marta's representations and strategies for T2 estimates in range 0-100, for each round in turn, are shown in the sequence of charts following. With regard to the first question, about the use of midpoint anchors, evidence suggesting use of midpoint anchors is circled in dark blue, and absence of midpoint evidence where it could reasonably be expected (targets 49 and 50) is circled in red. Evidence suggesting use of a midpoint anchor includes reference to a midpoint, immediacy of estimation, lack of count-on sequence, or count-on but "count-on queried". Multiple references to the left endpoint and straightforward count-on strategies are evidence against a midpoint anchor.

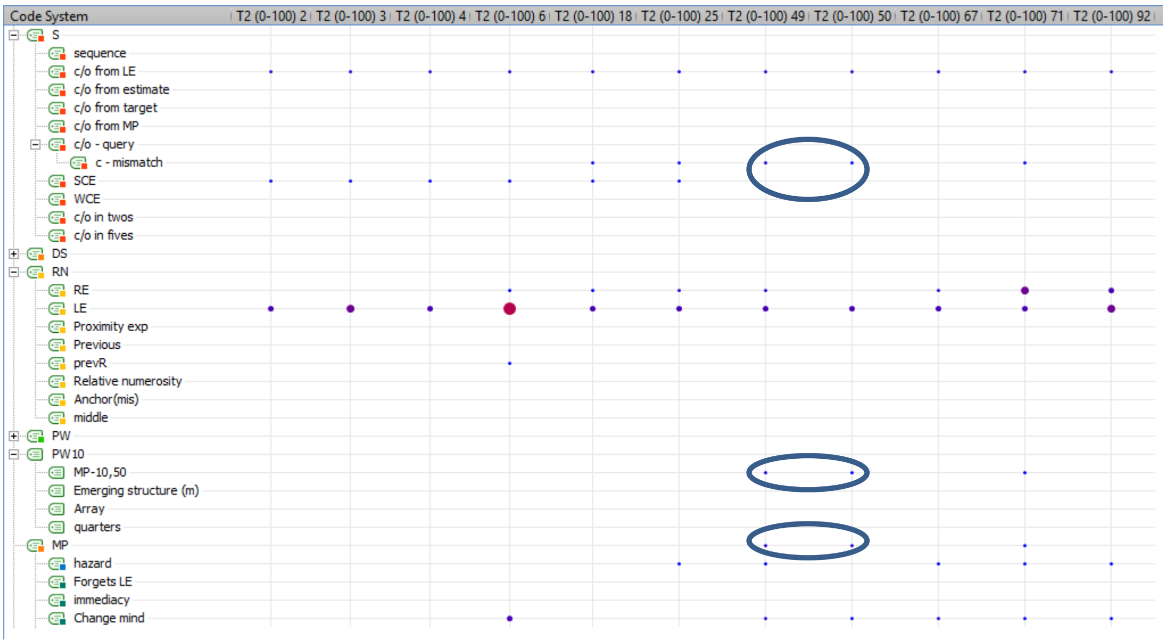


Figure 126 T2 estimates, range 0-100, R1

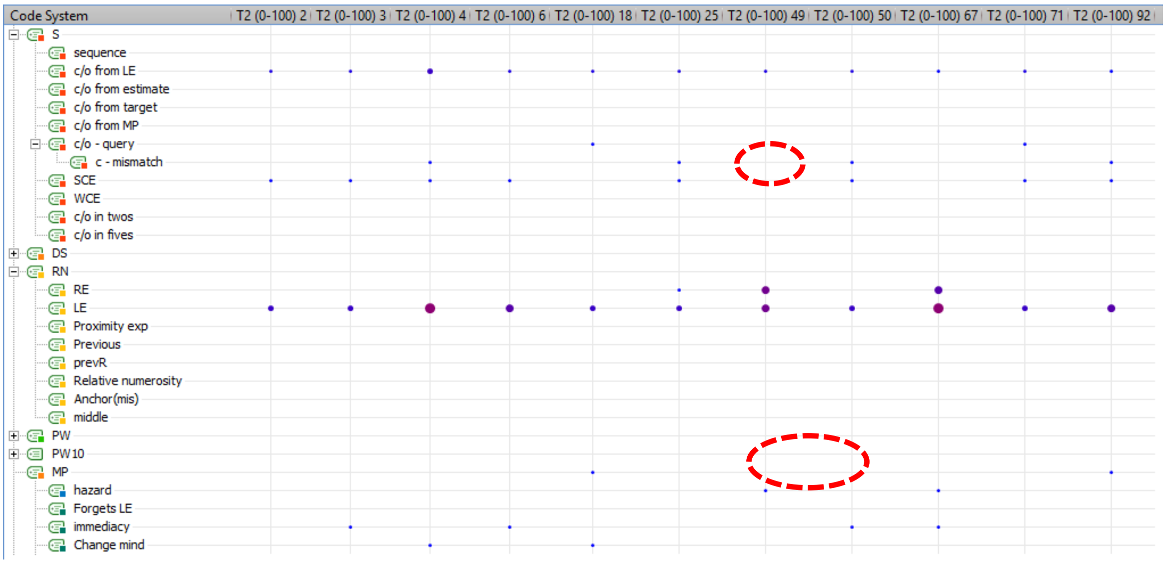


Figure 127 T2 estimates, range 0-100, R2

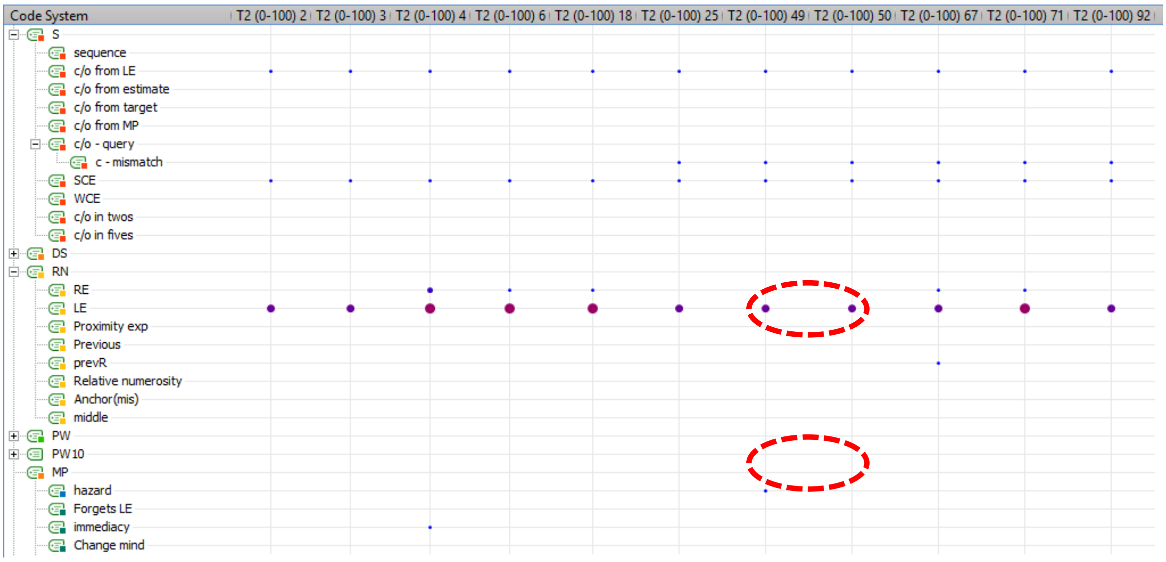


Figure 128 T2 estimates, range 0-100, R3

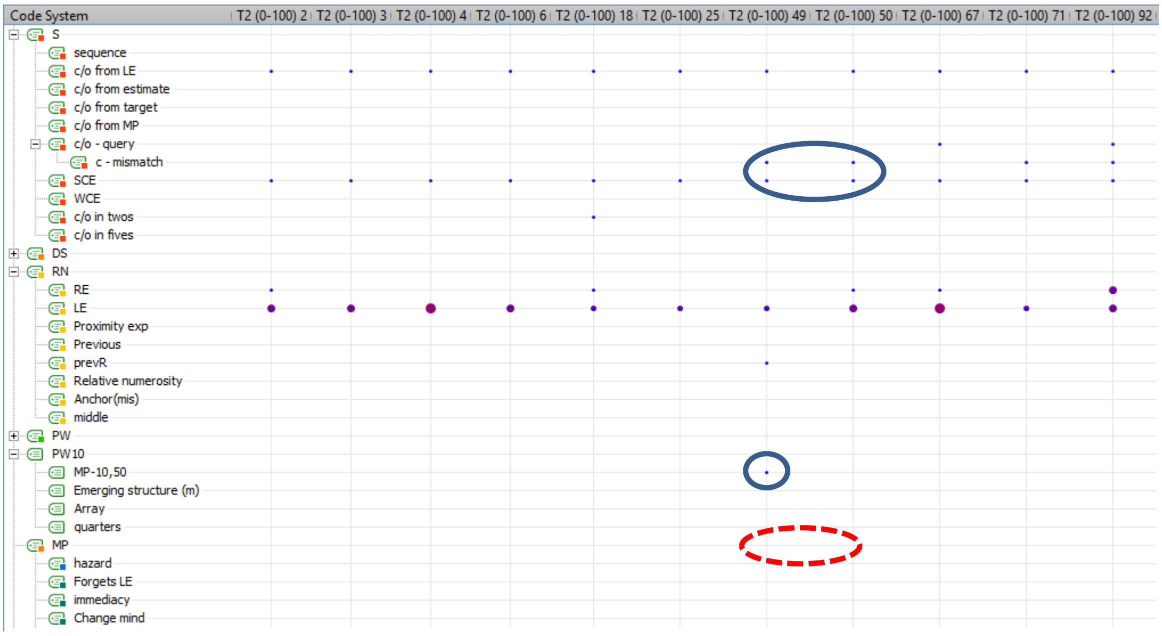


Figure 129 T2 estimates, range 0-100, R4

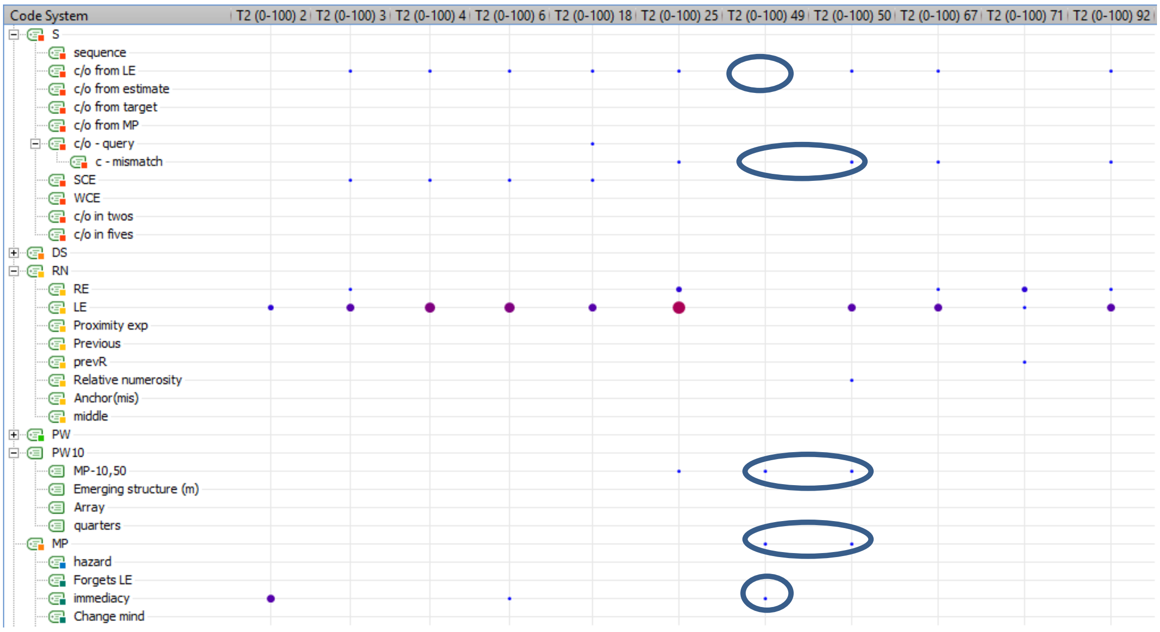


Figure 130 T2 estimates, range 0-100, R5

The charts above show that there was good evidence for midpoint strategies for targets 49 and 50 in R1, R4 and particularly R5, but not in R2 or R3. This coincides with the estimation accuracy for targets 49 and 50 in the PAE graph.

With regard to the second question, on the accuracy of estimations for 67 and 71, the above charts do not give much insight. For the most part, the strategies and representations observed during estimation of 67 and 71 are the same as those observed for other targets in the same round, except with far fewer midpoint references than targets 49 and 50. In particular, there is no striking difference between the strategies and representations observed for 67 and 71 (estimated very accurately) compared with target 92 (estimated with high PAE). An hypothesis about Marta's estimation of target 92 is that it was not close enough to 100 for Marta to consider proximity to 100 an important feature of its representation. Whereas Marta made changes to her estimation strategies to represent the proximity of 14 to 15 (during range 5-15 estimations) and the proximity of 8 and 9 to 10 (during range 0-10 estimations), target 92 was not, for Marta, represented as 'close to one hundred'.

The final two questions posed by the PAE graph relate to two occasions in which Marta estimated a target with notably lower PAE than in all other rounds. The first of these was target 25 in R3, and the charts of representations and strategies suggest no plausible reason for Marta's suddenly more accurate estimation; the strategies and representations observed are similar or identical to those observed for target 25 in all other rounds. The second example was target 18 in R4. Here, the representations and strategies are the same as observed in other rounds except for the addition of count-on in twos. There is of course no direct link between this representation and higher estimation accuracy, but it is an example of more sophisticated number structure than represented in most of Marta's estimations. It is possible, therefore, that it reflects that Marta had in mind a higher level of number structure thinking than her usual during this estimation.

The following graph of PE (percentage error) demonstrates initial over-estimation, followed by high accuracy around targets 67 and 71 (and in some rounds also targets 49 and 50), and then under-estimation of target 92 in all rounds. This is to a reasonable extent in accordance with a proportional judgement explanation of the task, with an inferred midpoint of fifty. However, as noted elsewhere, the current task design provides insufficient evidence with regard to estimation of targets close to the right endpoint to properly assess the fit of the proportional judgment account.

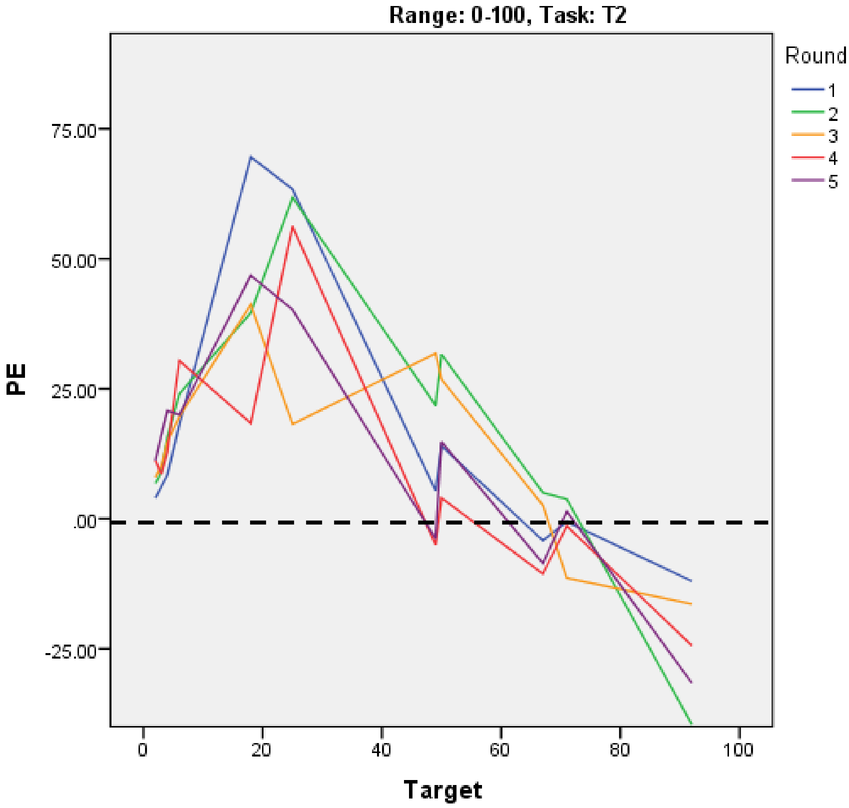


Figure 131 PE of target estimates, T2 0-100

8.3.2.6 T4, range 0-100

The following graph plots the PAE of Marta's T4 estimates for range 0-100. There is very high correlation between rounds for all targets except 92, which features a very dramatic difference between the PAE of the R1 estimate and all PAE of all later estimates. For targets except 92 (and even including 92 in R1), PAE increases linearly with the target's magnitude.

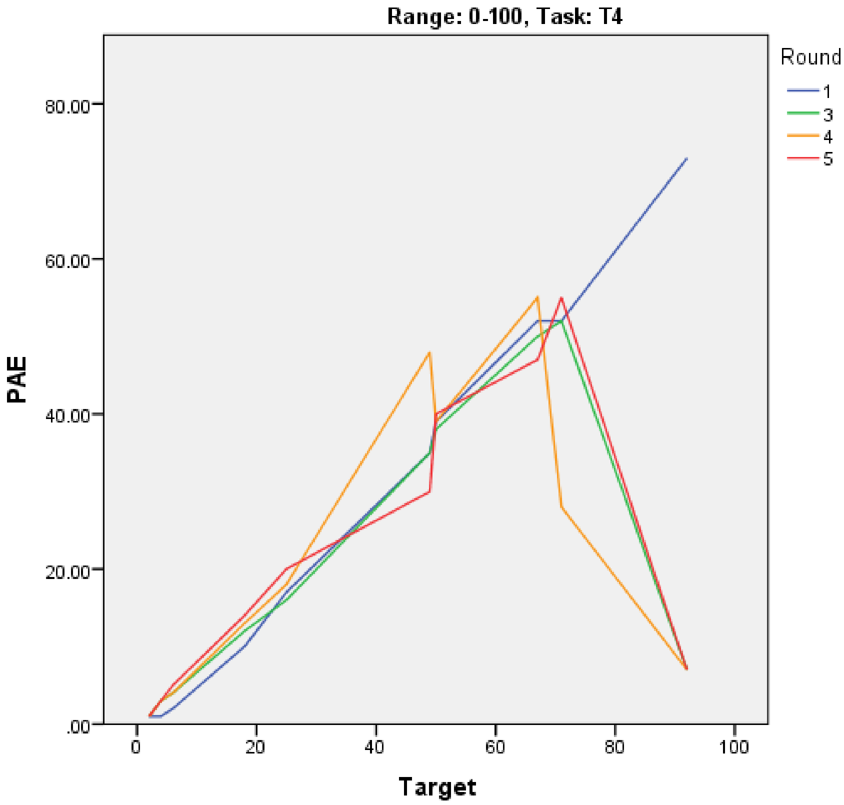
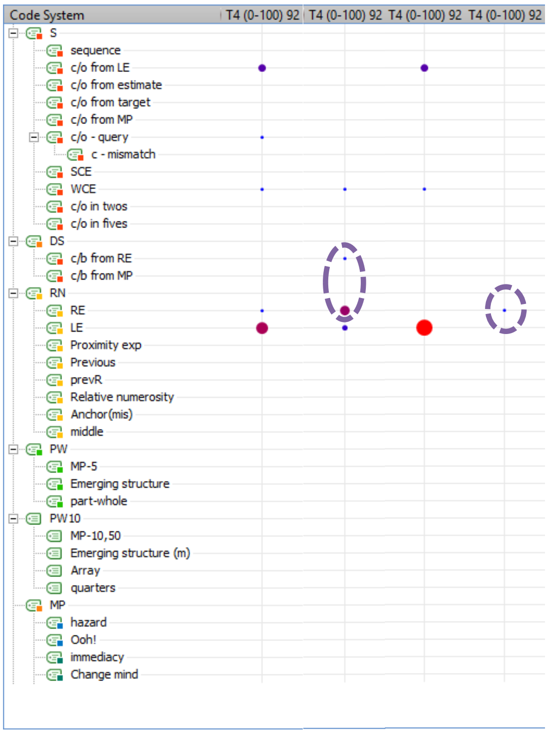


Figure 132 PAE of target estimates, T4 0-100

The question arising from this graph of Marta's T4 0-100 estimates is: what changed between R1, and R3, R4, and R5, in the estimation of target 92? Figure 133 plots Marta's representations and strategies for the T4 estimation of 92 in each round.



This chart reveals nothing in common between R3, R4 and R5 to explain why target 92 was estimated in these rounds with such strikingly lower PAE than in R1. In both R3 and R5, there is good evidence for a right endpoint anchor being used, but there is no such evidence for R4.

Figure 133 T4 estimates of target 92, range 0-100. Left to right: R1, R3, R4, R5

The next chart plots Marta’s representations and strategies for all targets in range 0-100, for T4 estimations in R3, R4 and R5. Although the graph of PAE for T4 0-100 estimations shows much lower PAE for target 92 compared with other targets (for which PAE increases linearly with distance from the left endpoint), the Figure 134 shows no striking differences between target 92 (green) and other target numbers. In fact, the representations and strategies observed for target 92 are identical to those observed for target 49 (orange), which was estimated with much higher PAE in all three rounds considered here. A possible reason for this finding returns to the argument that error is highest farthest away from anchor points, and in R3, R4 and R5 a right endpoint anchor was used. Although Marta did not significantly alter her strategies or representations for target 92, the mere fact of its proximity to the numbered right endpoint meant that the same strategies could lead to lower PAE. The proximity to right endpoint was a key factor in T4 in a way that findings suggest it was *not* in T2 (see previous section) because in T4 the spatial representation of target and endpoint is given. In T2, a notion of the “proximity” of 92 and 100 was something that participants could either represent or not, whereas in T4 it was represented spatially in the task itself.

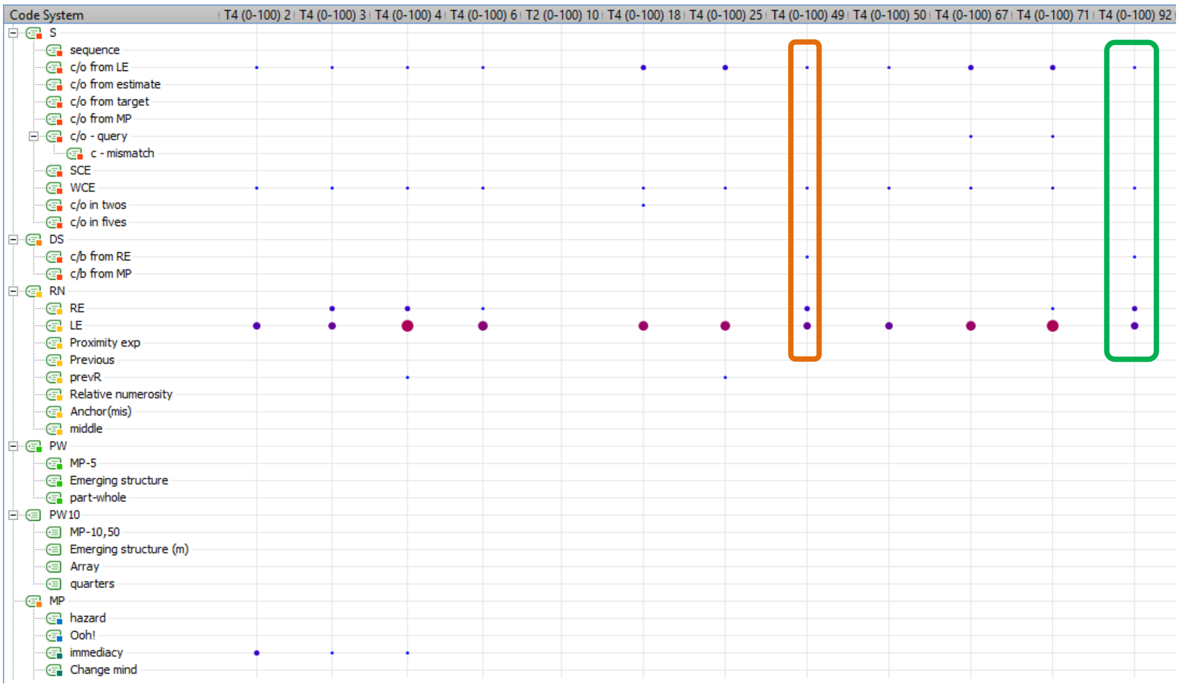


Figure 134 T4 0-100 estimates, R3 R4 and R5

To summarise this section on the connection between strategy and estimation accuracy, increases in Marta’s estimation accuracy were in some cases concurrent with observed changes in the representation of relevant number structure, for example the right endpoint of number ranges. In these cases, a fairly firm conclusion can be drawn that the number structure represented was an important factor in the increased linear accuracy of estimation. In other cases however, relevant number structure was represented in trials which had very different levels of linear accuracy. It is certainly not the case that strategy and estimation accuracy can be linked in all trials. At a more general level, tasks in later rounds in which estimation accuracy was higher were associated with fewer instances of count-on from left endpoint and reference to left endpoint only.

8.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling?

This question involves examining the changes evidenced by both the qualitative and quantitative estimation data collected from Marta's interviews. Before considering changes in estimation results, the diagram below summarises Marta's representations and strategies in each round of interviews. The chart used at this point, as in the previous case study, compares each frequency only to the frequencies of the same category of representation in other rounds, in order to better highlight trends within individual categories of representation. In Marta's case, these trends were not visible when frequencies were compared across the whole chart, due to the dominance of left endpoint and count-on representations. The following chart is not suitable for making any comparisons *between* categories.

Figure 135 Marta's representations and strategies by round

A number of trends are apparent from the above data. Firstly, the frequency of count-on from left endpoint strategies decreases over the course of the five interviews. The frequency of "count-on queried" observations also decreases between the first and final interviews. Next, there is an increase in representation of more mathematically advanced types of sequence, namely decreasing sequences and counting by twos or fives, which do not appear at all in the first two interview rounds. There is considerably more representation of relative numerosity and reference to previous targets in later interviews compared with the first two interviews. There is also more representation of part-whole structure in R4 and R5 compared with the first interviews. In the final interview, there is a substantial increase in the frequency of representations and estimations made with immediacy.

Overall, these changes could be summarised as small but noticeable increases in the representation of increasingly sophisticated number structure, and a steady decrease of the most basic estimation strategies. For several categories of number structure shown above, the change is most striking between R1 and R2 (grouped) compared with R3, R4 and R5 (grouped).

In terms of the linear accuracy of number line representations, Marta's representations became more linearly accurate over the course of the year (see RQ1c). However, Marta's number-to-position estimations (T2) showed a different trend to her position-to-number (T4) estimations. For T2 estimations, ANOVA testing revealed a significant main effect of interview round on the linear accuracy of representations. For T4 estimations, there was a striking reduction in the PAE of range 0-10 estimates between R1 and later rounds, but for other ranges much less change. Between R3 and R5, there was no statistically significant increase in linear accuracy for any range.

8.4.1 In what ways does evidence support or not support the log-linear hypothesis?

Marta's case provides only limited support for the logarithmic-to-linear shift hypothesis. The main piece of evidence it provides in support is an increase in the linearity of number line estimations. However, the increases in linearity of estimations tended to occur to specific target numbers rather than to ranges (which would be the case if a mental representation of a range made a logarithmic-to-linear 'shift').

Linear, logarithmic and exponential models were fitted to each range of Marta's T2 and T4 estimates for each round, to compare model fit and assess how well the logarithmic-to-linear shift hypothesis seemed to apply in Marta's case. The following table lists the R^2 model fits for each model in each task condition. Where the best fitting model is the one predicted by the log-linear hypothesis (i.e. logarithmic for T2, and exponential for T4), the R^2 figure is highlighted yellow.

Range	Task	Round	R Square Linear	Logarithmic	Exponential
0-10	T2	1	.581		.842
		2	.796	.611	.949
		3	.767	.545	.979
		4	.803	.595	.945
		5	.939	.885	.871
	T4	1	.721	.877	.653
		2	.987	.989	.979
		3	.888	.949	.808
		4		.469	
		5	.845	.793	.830
0-20	T2	1	.943	.830	.918
		2	.959	.867	.910
		3	.946	.809	.953
		4	.928	.855	.843
		5	.840	.879	.769
	T4	1	.869	.774	.885
		2			
		3	.904	.910	.808
		4	.692	.706	.691
		5	.902	.918	.727
5-15	T2	1	.897	.844	.914
		2	.769	.715	.831
		3	.881	.798	.927
		4	.774	.862	.760
		5	.715	.659	.805
	T4	1		.647	.570
		2			
		3	.817	.819	.820
		4	.788	.807	.735
		5		.614	
0-100	T2	1	.435	.705	.500
		2	.430	.768	.471
		3	.707	.897	.651
		4	.551	.726	.570
		5	.530	.793	.536
	T4	1	.939	.879	.713
		2			
		3	.577		.902
		4	.585	.448	.814
		5	.587		.942

Figure 136 Model fits for log-linear hypothesis

The above table reveals that it was only for range 0-100 that the models predicted by the log-linear hypothesis were consistently best-fitting. Even for this range, however, the R^2 figures show that model fits were not particularly good (for example, not as good as for the exponential models in T2 range 0-10, even though the log-linear hypothesis gives no reason why these estimates should follow an exponential distribution). There is no task condition which revealed a shift from logarithmic (or exponential) to linear distribution over the course of the five interviews, even though increases in linearity (as measured by PAE) occurred. Range 0-10 in T2 also shows this increase in linearity in the increasing R^2 figures for the linear models, but even in the earliest interviews, the logarithmic model described Marta's estimates less well than the linear model anyway.

8.4.1.1 Alternative accounts

Marta's case provides qualitative data on the estimation of individual target numbers. In addition, plotting PAE by target number, as in RQ2b, allows assessment of the hypothesis of proportional judgment accounts that estimation accuracy should be highest near to anchor points and lowest at the farthest point from anchor points. For ranges 0-10, 5-15 and 0-100, the graphs of PAE by target number (see RQ2b) strongly support this hypothesis, though range 0-20 does not.

The observation that changes in the linearity of estimation tended to occur to individual targets numbers rather than to ranges is also more in accordance with a proportional reasoning account than a log-linear shift account. The target numbers where striking changes between rounds were observed were most often those pre-identified as likely candidates for the use of anchor points: targets close to each endpoint, and to the midpoint of ranges. These changes (as explored in RQ2b) are in accordance with the idea that children's increasing linearity of estimation stems from the gradual incorporation of more anchor points as well as a general reduction in estimation bias.

8.4.2 **What is the intra-child variability of children's numerical magnitude representations in estimation tasks at different times?**

There are two aspects to consider in answering this question. First is the variability in observed representations and strategies, which can be seen most clearly in the analysis of RQ2. In Marta's case, one overall finding was high consistency between trials within a given task condition – the representations and strategies did not typically vary very much at all between different target numbers. Where variations of representation and strategy (from Marta's 'default' representations of left endpoint and count-on) did occur, it was often for targets at the midpoint and close to the right endpoint, and these variations occurred more frequently in later interview rounds than in early interview rounds. The charts showing representations and strategies in T4 range 0-10 illustrate this, with R1 (Figure 114) showing very similar observations for all targets, in contrast to all later rounds (Figure 115).

The second important aspect to this question is the variation of estimation error (PAE) in different task conditions at different times. The graph below shows the standard deviation for T2 and T4, in each round.

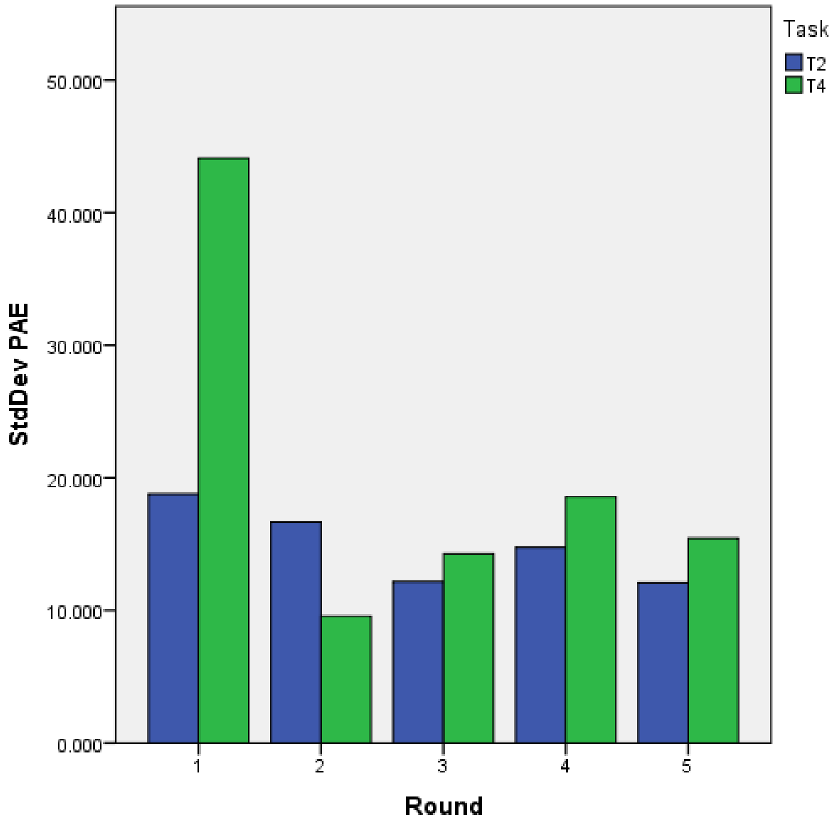


Figure 137 Standard deviation of PAE for T2 and T4 by round

The graph demonstrates a slightly lower variation in PAE in R5 compared to R1, but no strong downward trend. Variability of PAE was in most rounds higher in T4 than in T2. The most noticeable feature of this graph is of course the far higher standard deviation in PAE for T4 in R1. However, previous analysis (see RQ1c) identified significant differences between PAE variance in different ranges in R1, indicating that the above graph might not accurately represent a trend in variability for all ranges. The series of graphs on the following split the data according to range, to allow examination of variability over time for each range separately.

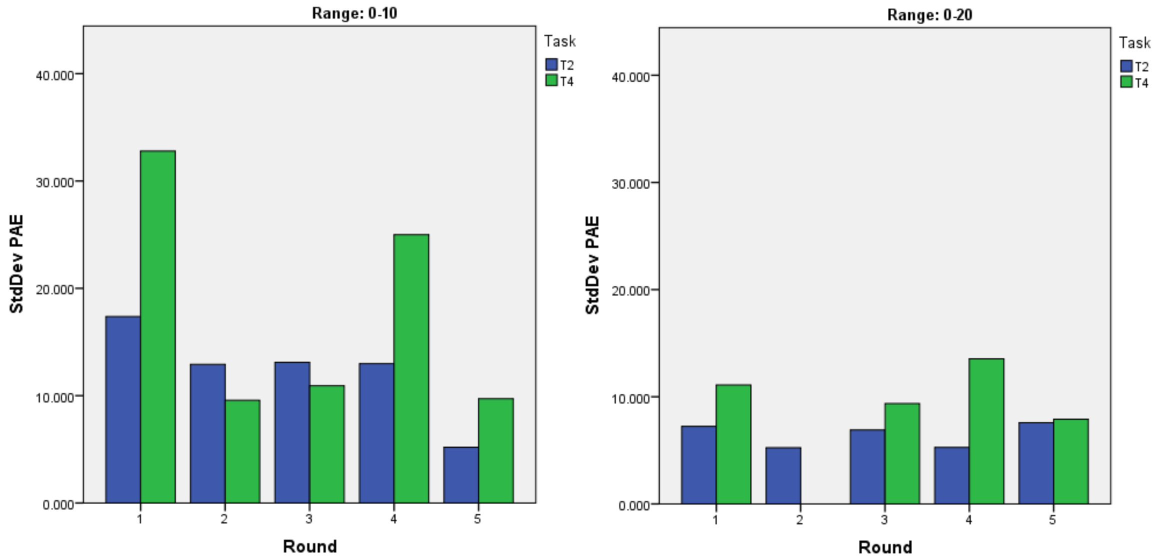


Figure 138 Ranges 0-10 (above) and 0-20 (right)

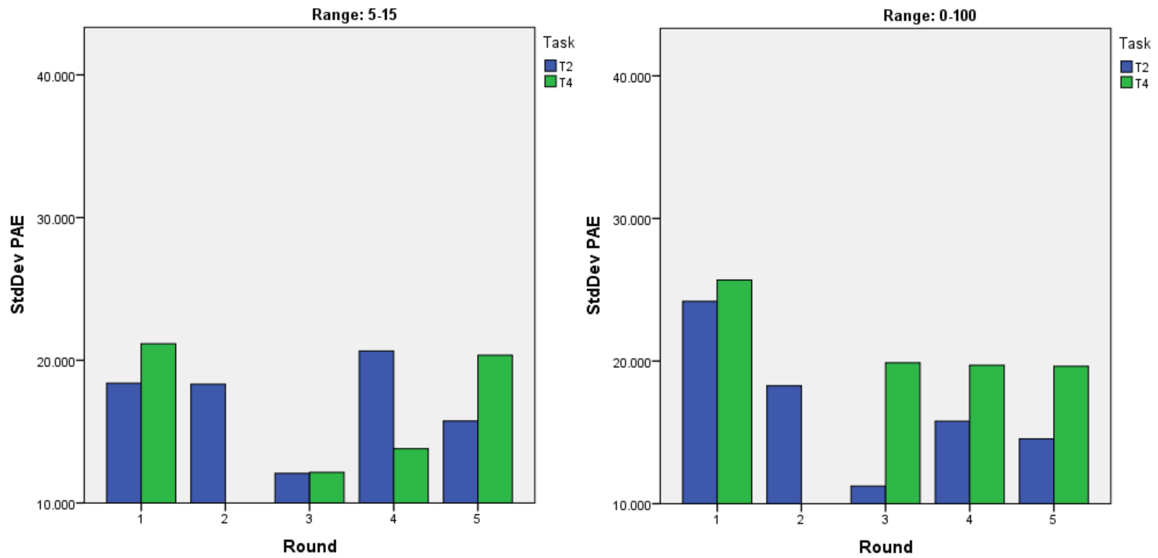


Figure 139 Ranges 5-15 (above) and 0-100 (right)

For ranges 0-10 and 0-20, variability of estimation PAE was lowest in R5. For ranges 5-15 and 0-100 on the other hand, variability of PAE was lowest in R3, although for T4 range 0-100, it did not change between R3, R4 and R5. The biggest differences in variability of PAE were between R1 and later rounds.

8.4.3 Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?

Marta's case illustrates both increased linear accuracy of representation over time, and representation of increasingly advanced mathematical structure. Importantly however, the representation of more advanced structure was not always accompanied by increased estimation accuracy on a target-by-target basis, and vice versa. For some targets, it proved possible to identify Marta's increased accuracy with changes in representation, but this was not the case for all targets.

Marta's case also illustrates an interesting trend of reluctance to change strategy. The basic strategy of count-on from left endpoint continued to be used and imitated even when Marta demonstrated more mathematically advanced representations of the same target numbers, for example representations involving the relative numerosity of targets and endpoints, and part-whole structure. This finding was somewhat unexpected, and the present methodology was not designed to investigate it, but two possible explanations are proposed.

One suggestion is that since the task-based interviews involved no feedback other than encouragement, the initial unfamiliarity of the tasks followed by repetition throughout the year encouraged Marta to continue the strategy with which she had felt comfortable and successful in the first experience – R1. The pattern of holding on to initial strategies was not seen to the same extent in other children, but this does not rule out the possibility that the research design encouraged this response in Marta.

The second suggestion focuses on the familiarity and reliability of the increasing-sequence representation of numbers. For any given target number in this study, the representation of that target as "*n* units on from zero" (or the adaptation for range 5-15) is a correct mathematical representation, and one that children in Year One are, and have been, exposed to very regularly indeed. It is plausible that even when Marta saw an alternative way to represent the target number for the purposes of the estimation task, she remained keen to use, or appear to use, the increasing-sequence representation of the target due to its high familiarity and perceived reliability.

Chapter 9 Case study: Catharina

Catharina was selected as the final in-depth case study because her responses to the task-based interviews had neither particularly high nor particularly low levels of mathematical sophistication. When the children were approximately ordered according to the amount and level of number structure they represented in their task responses, Catharina fell in the middle of the group. Her number line estimations were close to the median levels of linear accuracy.

9.1 What does this case say in response to important ideas in the literature?

This case provides further strong qualitative evidence for the use of lower and upper endpoints and midpoints as 'anchor' points during number line estimations. Such a hypothesis has been advanced in various forms by Siegler and Opfer (2003), Barth and Paladino (2011), White and Szucs (2012), Slusser et al. (2013) and Rouder and Geary (2014), and Catharina's case provides qualitative support for it. The order in which Catharina included anchor points is in line with the specific predictions of Ashcraft and Moore (2012): the left endpoint is always included as an anchor point, then the right endpoint (for Catharina, the only occasions on which it was not included were in R1 and R2), and then a midpoint (for Catharina, the clearest evidence occurs in R5).

To a greater extent than in Patrick's case, the qualitative data in Catharina's case generally coincided with the suggestions from the quantitative data alone. In other words, lower estimation error around potential anchor points, the observation that led many of the above researchers to hypothesise use of anchor points, did coincide with qualitative evidence for their use, and vice versa. The examination of estimations on a trial-by-trial basis in RQ2b demonstrates this most fully. To a greater extent than other cases therefore, Catharina's case provides some support for deducing estimation strategies from quantitative estimation results.

Catharina's case does not provide good support for the log-linear shift account of number line estimation. As explained in RQ3a, Catharina's data only convincingly agree with the log-linear account in showing an overall increase in linearity of number line estimation, and this prediction is not unique to the log-linear shift account. In terms of the other specific predictions of the account, namely initial estimations being well fit by logarithmic models, and changes in representations for a given range occurring rapidly, Catharina's case refutes rather than supports the theory.

Catharina's case does not provide good support for a proportional reasoning account of number line estimation either. Graphing linear estimation error by target did not reveal the patterns of over- and under-estimation predicted by proportional reasoning models, and in fact the observed patterns were much better explained by appeal to the observed representations and strategies during estimation trials. Catharina's case provides particularly clear evidence for how the estimation strategies a child deploys can impact on the specific estimation errors recorded, for example the unit size of a consistently applied count strategy leading to consistent under-estimation.

The evidence of this case provides good support for the prediction of White and Szucs that "specific numbers could exhibit unique behaviors as a function of the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy" (2012, p. 9). Charts displaying Catharina's representations and strategies according to target type (proximity to various potential anchor points) shows very clear variation in line with expectations, for example the dominance of right-endpoint references, coupled with reduced use of count-on from left endpoint strategies, for targets close to the right endpoint.

In this remainder of this chapter, I present the findings from Catharina's case study in relation to each of the research questions.

9.2 RQ1: In what ways do children appear to cognitively represent number during the different tasks of the interviews used in this research?

9.2.1 What are the modes and component signs used in the representations?

Catharina produced inscriptions on paper during T1 of each interview, and the component signs of these inscriptions were notational (only).

In R1, T1 was the only task in which Catharina made inscriptions on paper. In R2, Catharina quite frequently drew on or annotated the page in T2 and T4, as well as T1. She marked points or dots for the units in counting, as in this example from T2, R2:

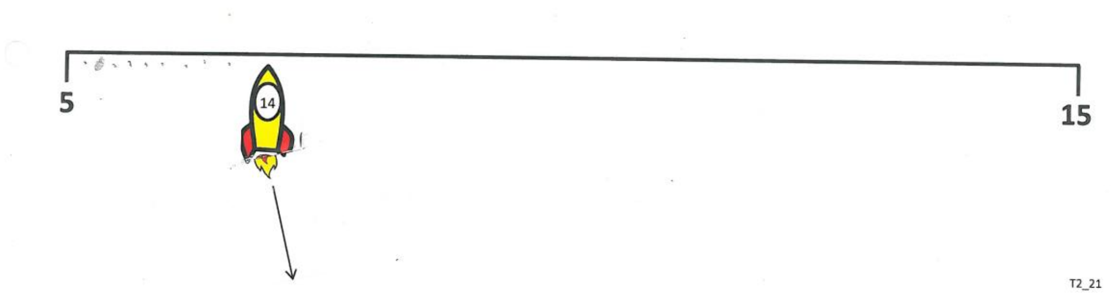


Figure 140 T2 R2, estimation of target 14, range 5-15

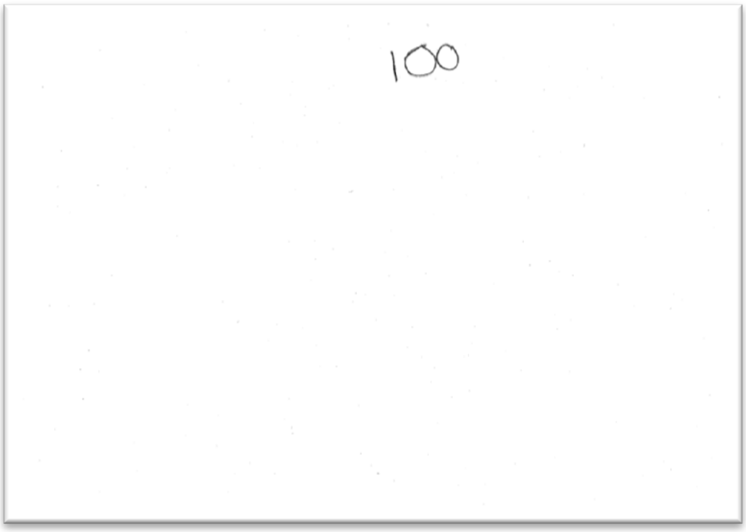
This form of annotation also appeared in R3, though infrequently, but not at all in R4 or R5. Catharina produced verbal, gaze and gesture representations throughout all five interviews. The discussions of Catharina's representations in the following sections therefore include verbal, gesture and gaze representations throughout, as in the previous two case studies.

9.2.2 What aspects of number structure are represented?

9.2.2.1 T1 representations

In terms of Thomas and Mulligan's (1995) framework of analysis, Catharina's T1 representations represent no number structure. The inscriptions, together with any accompanying representations in speech and gesture, are nevertheless included here and very briefly discussed in order to show the aspects of number that Catharina *did* consistently represent in her T1 responses.

Other children interviewed also responded to T1 with inscriptions made up of notational components, but the majority of these inscriptions were numbers up to and including one hundred. Catharina, in contrast, made notational representations of numbers no *smaller* than one hundred.

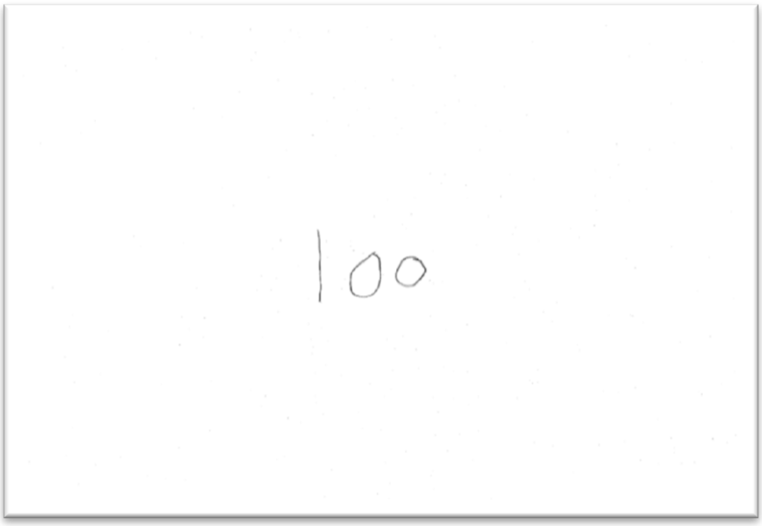


Catharina's drawing in T1 R1, left, was not accompanied by any representation in gesture. Verbally, Catharina confirmed that the drawing showed "one hundred".

Figure 141 Catharina T1 R1

Catharina announced "One … a ten!" after producing this drawing in R2, pointing to the final "10". The researcher checked "Is that what comes into your head?" and Catharina nodded.

Figure 142 Catharina T1 R2



Catharina's R3 response was identical to that of R1. There was no representation in gesture, and verbally Catharina confirmed to the researcher that the drawing showed "one hundred".

Figure 143 Catharina T1 R3

Figure 144 Catharina T1 R4

After producing this drawing in R4, Catharina made a spontaneous comment:

C: Twelve ... hundred! [*pointing to numerals*]

J: What does that say?

C: Twelve hundred! [*points again to numerals*]

The researcher's question double-checked that Catharina was equating "120" and "twelve hundred".

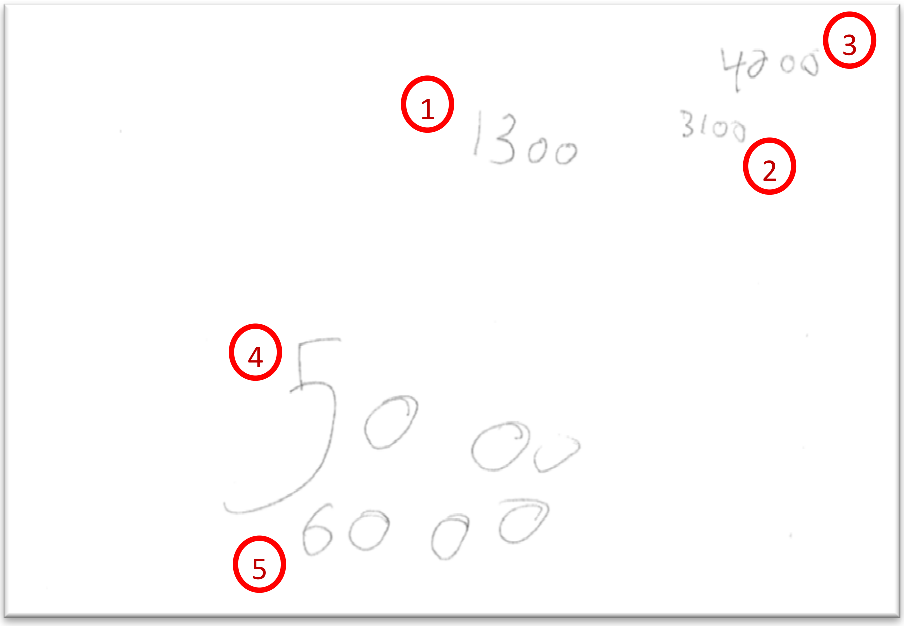


Figure 145 Catharina T1 R5

The above drawing, Catharina's response to T1 in R5, was drawn in stages, as indicated by the red circled numbers. Follow-up questions or probes, which had not been used in earlier rounds on the basis that they would influence T1 responses in future interviews, were included here since it was the final interview. For this reason, only Catharina's initial (stage 1) T1 response in R5 should be considered directly comparable to that of R1-R4. The further representations invited by the follow-up questions do however provide interesting supplementary evidence for the T1 responses in all rounds.

J: ... One three zero zero? [*reading Catharina's drawing – only "1300" at this stage*]

C: [*nods vigorously*]

J: So, can you explain that to me a bit?

C: [*smiles*] Thirty [NB: not thirteen] ... hundred! [*reading the numerals, then smiling at J*]

J: OK... what about other numbers?

C: What other numbers? [*shaking head*]

J: Well ... I asked you to think about all the numbers from one all the way to one hundred...

C: I've got another numbers ...

J: How do they look in your head?

C: [*leans over and writes "3100"*] … Thirty one! … [*eyes wide open, grinning at J*]

J: That's a high number, isn't it? [*sounding impressed*]

C: [*nods vigorously*] And I've got another one ... I know … [*slowly writes "4000"*]

J: Yes, that's high as well. [*sounding impressed*]

C: Forty hundred!

J: So are those the numbers that come into your head?

C: [*nods vigorously*] Mmhmm.

J: Yes? OK.

C: Forty hundred and another … [*writes "6000"*] and sixty hundred.

This episode shows Catharina representing increasingly large numbers. There is no obvious mathematical connection between the numbers represented and the task itself other than the fact that all are 'large' by the standards of a Year One classroom.

The follow-up questions to Catharina's initial response invited further representation of some sort, and may have influenced more specifically what Catharina decided to represent. However, there is evidence that the response 'represent large numbers' occurred before any influence of follow-up questions. Firstly, the inscriptions made in R1-R4 and stage one of R5 all show numbers greater than or equal to one hundred. Secondly, and no less importantly, Catharina presented each of these numbers to the researcher with a flourish: her body language was triumphant, her tone of voice was exclamatory, and she looked immediately to the researcher apparently for acknowledgment. Catharina was highly engaged with the researcher throughout all parts of the interviews, but this precise set of behaviours was only observed when Catharina presented the researcher with large numbers, for example when estimating "ninety nine" for target 92 in T4, range 0-100. For these reasons, it seems plausible to say that Catharina was responding to T1 by representing numbers that she considered large and impressive.

Although in T1 Catharina represented no aspects of number structure according to the frameworks of Thomas and Mulligan (1995) and Resnick (1983), her responses consistently represent an association between the task request ("think about the counting numbers from one up to one hundred") and numbers that are 'large' in the context of a Year One classroom.

9.2.2.2 All tasks

Catharina’s representations during all tasks were analysed according to Resnick’s microstage theory. The following table contains the aspects of number structure in Resnick’s theory, and Catharina represented all categories of number structure in this table. Grey text shows the few specific procedures and strategies that Catharina did *not* represent.

Number structure	Detailed stages	Procedures or strategies in which this number structure may be represented
S Sequence	Count sequence from 0 or 1	Count-on from LE.
	Count sequence from other start point	• Count-on from non-zero LE. • Count-on from estimate to RE for confirmation.
DS Decreasing sequence	Count sequence reversed	Count-back from RE.
RN Relative numerosity (quantity comparison)	Ability to represent (esp visualise) number’s magnitude without counting up to it	• Applying efficient count strategy, using RN to choose shorter available count. • Checking judgements, querying whether an estimate is likely. • Use of endpoints as “anchor” points.
PW Part-whole	Partition and recombine numbers (<20)	Use of “anchor” points based on partitioning the range.
	Use of number bonds to ten	• Checking estimations using number bonds to ten. • Using own midpoint anchor on 0-10 range.
	Multiple partitions of multi-digit numbers	Using own midpoint anchors on 0-20 and 5-15 ranges.
PW10 Part-whole understanding of base ten structure	Numbers as compositions of tens and units	• Using own midpoint anchor on 0-100 range. • Count-on and count-back by tens. • 10x10 array structure of 0-100
	Multiple partitions of larger multi-digit numbers	• Count-on or count-back by any given number. • Partitioning into quarters, thirds, other.

Figure 146 Aspects of number structure in Resnick’s theory represented by Catharina

The following chart shows the relative frequencies with which Catharina represented each category of number structure across the five interviews. The circles are coloured and sized to represent the varying frequencies with which representations occurred, relative to all the other frequencies in the chart. As for other children interviewed, the most common representations were of the left and right endpoints of the number lines in estimation tasks. Catharina represented most categories of number structure in each interview.

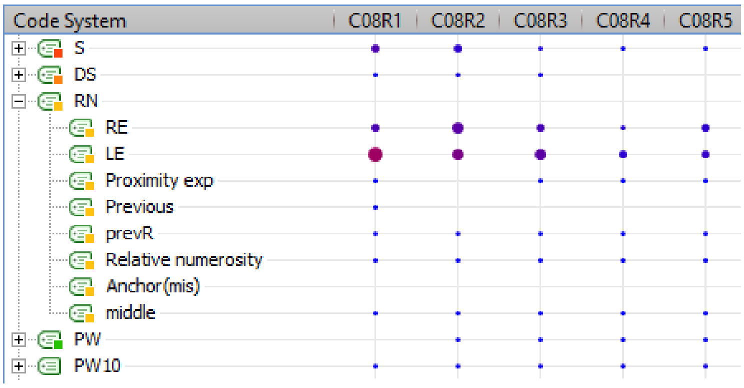


Figure 147 Frequency of number structure representations by Catharina

Three examples are used here to show the kind of episodes from Catharina's interviews that were inferred to show representation of different aspects of number structure. The first extract is taken from Catharina's R4 interview, and shows representation of relative numerosity, reference to the right endpoint, and increasing number sequence. The task is T2, and Catharina is asked to estimate the position of target 18 on a 0-20 number line.

J: Where does 18 go?

C: [*takes rocket to midpoint mid-air, adjusts leftwards a little, then suddenly jumps to RE. Sticks rocket next to RE.*] There. And then 19 would go there [*points to gap between rocket and RE*].

The next episode is taken from Catharina's R2 interview, and demonstrates representation of decreasing number sequence, and again reference to the right endpoint. The task is T4, the range is 0-20, and Catharina is estimating the number of a blank rocket that represents target 18:

C: [*eyes trace line from left endpoint to rocket and right endpoint, slowly*] Twenty, nineteen [*counts back from right endpoint with small unit 'hops', writes "18"*] … I put eighteen!

The final example, taken from Catharina's R1 interview, shows some representation of part-whole number structure. The task is T2, the target is 50 and the range is 0-100:

J: And … the last one is fifty.

C: Fifty … [*pauses, points to RE, then looks to J*] …. There?

J: Wherever you think best.

C: [*looks to midpoint, moves suddenly to midpoint*] I'm going to stick it IN the MIDDLE [*sticking rocket at midpoint*]

J: Oh - why?

C: Because … it belongs in the middle!

9.2.3 What are the notable between-task and within-task connections between representations?

In this section, Catharina's representations are compared between tasks and within tasks, asking what varied and what was consistent.

9.2.3.1 Task 1

Catharina's T1 representations were presented in RQ1a above, and the following table summarises them according to the framework of Thomas and Mulligan (1995). In terms of the component signs used and absence of number structure representation, Catharina's T1 representations were highly consistent.

		Interview round				
		R1	R2	R3	R4	R5
Structure, classification from Thomas and Mulligan (1995)	No structure	X	X	X	X	X
	Linear structure only					
	Emerging structure					
	Emerging structure (multiplicative)					
	Partial array					
	Array: 10x10					
	Component signs	Notational	Notational	Notational	Notational	Notational
	Static/dynamic	Static	Static	Static	Static	Static

Figure 148 Structure of Catharina's representations of number during Task 1 (Thomas & Mulligan, 1995)

The connections between Catharina's T1 representations and T2 and T4 representations are limited. A particular disparity is the representation of number structure, with none represented at all in T1, in contrast to quite frequent representation of multiple aspects of number structure in T2 and T4. The clearest connection between T1 representations and other tasks' representations is Catharina's enthusiasm for large numbers, as discussed in RQ1b.

9.2.3.2 Task 2 and Task 4

In this section, representations in different task conditions are compared to look at the between-task and within-task connections for representations during number line estimations. As in the previous case studies, this analysis includes representations in gaze, in gesture and in speech.

The following chart reveals highly consistent representation of number structure in different task conditions. Sequence structure, decreasing sequence structure, and relative numerosity are represented in all task conditions, and part-whole structure is represented in all but one condition (T4, range 0-100 is the only one).

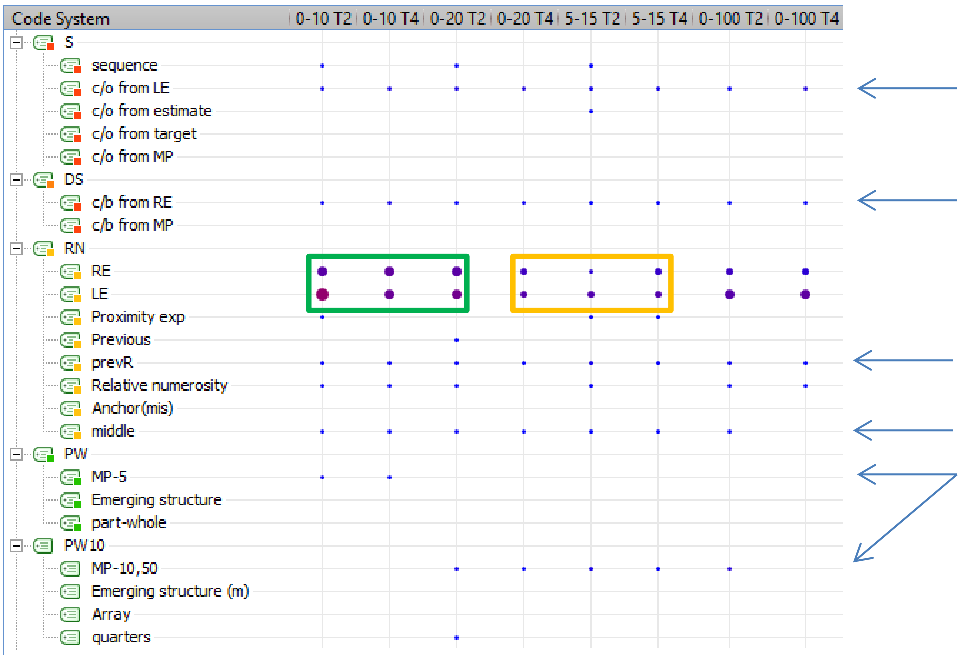


Figure 149 Structure represented by range and task

Furthermore, for many of the sub-categories of number structure, representations were observed more or less equally across task conditions, as indicated by the blue arrows. References to right and left endpoints were not equally prevalent in all conditions; more were observed some conditions (green box) compared with others (orange box).

9.2.3.3 Linear accuracy of representations

Between-task and within-task connections between representations also occur in the linearity of number line estimations. The extent to which Catharina's representations of targets coincided with the linear representation of number for each given range was measured using the percent absolute error (PAE). A three way mixed ANOVA was carried out to investigate the variation of PAE with round, task and range. As in the previous case studies, PAE scores were transformed using the square-root transformation, and all results are reported as significant (or non-significant) at p<.05. The assumption of sphericity was met for interview round, but not for the interaction of interview round and task, $\chi^2(9)=27.15$, p=.001. For the interaction of interview round and task, degrees of freedom were therefore corrected using the Greenhouse-Geisser estimate of sphericity (ϵ=.738).

There was a significant main effect of interview round on PAE scores, *F*(4,132)=7.25. Planned contrasts revealed a significant linear trend for PAE scores overall, *F*(1,33)=20.57, reflecting that the linear accuracy of estimation overall increased (PAE decreased) proportionately with interview round.

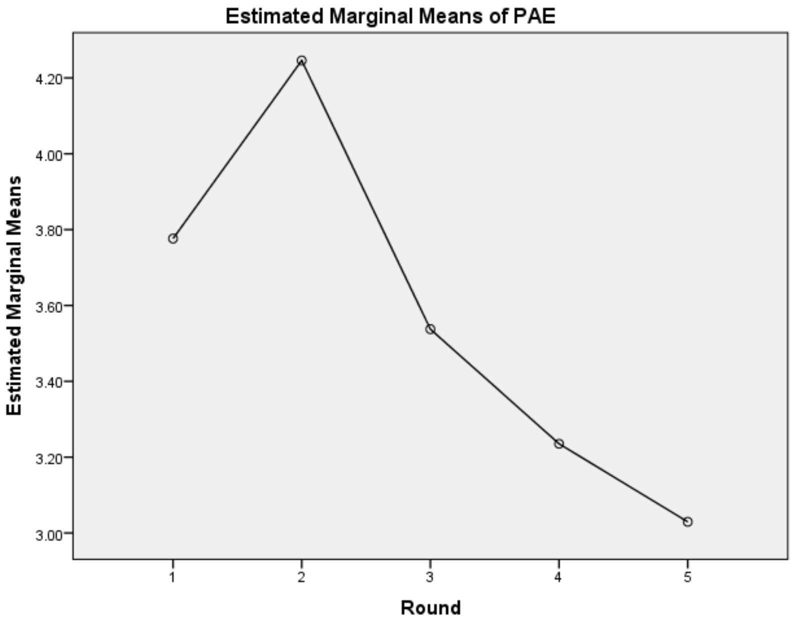


Figure 150 Catharina, mean PAE by round

Bonferroni-corrected post-hoc tests confirmed that estimation PAE in R4 and R5 was significantly lower than in R2.

There was no significant main effect of task, but there was a significant interaction effect of interview round and task, *F*(2.95,97.39)=3.54. The graph below of estimated mean PAE by round and task reveals that PAE increased for both tasks between R1 and R2. However, after this point PAE decreased with interview round at a much faster rate for T4 than for T2, so that by R4 and R5, the mean PAE of T4 estimates was actually lower than for T2.

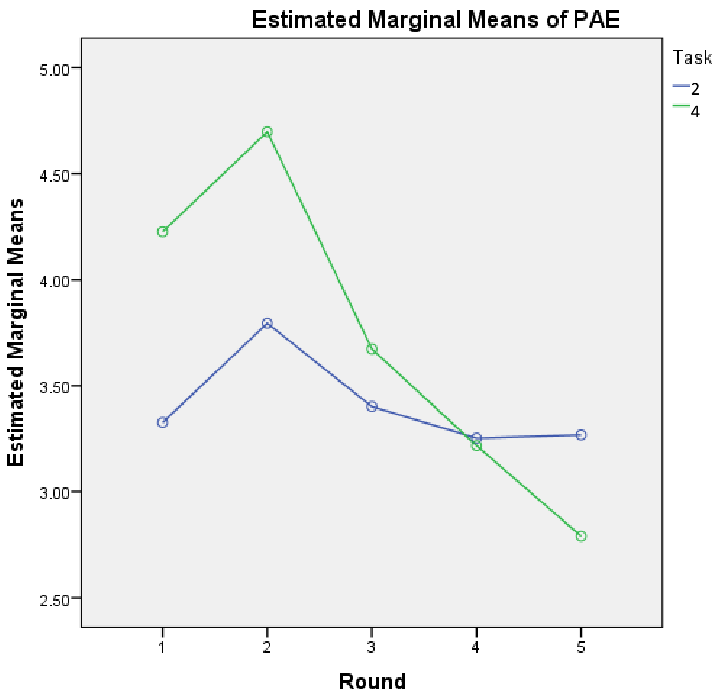


Figure 151 Mean PAE by round and task

No significant main effect of range was found. An interaction of interview round and range was found, $F(12,132)=4.38$; however this finding is of limited reliability since for several (three of ten) round-by-task conditions, Levene's test for homogeneity of variance between groups was significant at 5%. Violating the assumption of between-groups homogeneity of variance in these instances makes the significance of the interaction effect involving range groups less reliable. The finding is nevertheless reported, with the graphs of estimated mean PAE displayed, since Catharina's estimates in the different ranges display different trends over the five interview rounds (see Figure 152) and it is interesting to investigate this. Whether the variation adds up to a statistically significant interaction effect is not proven.

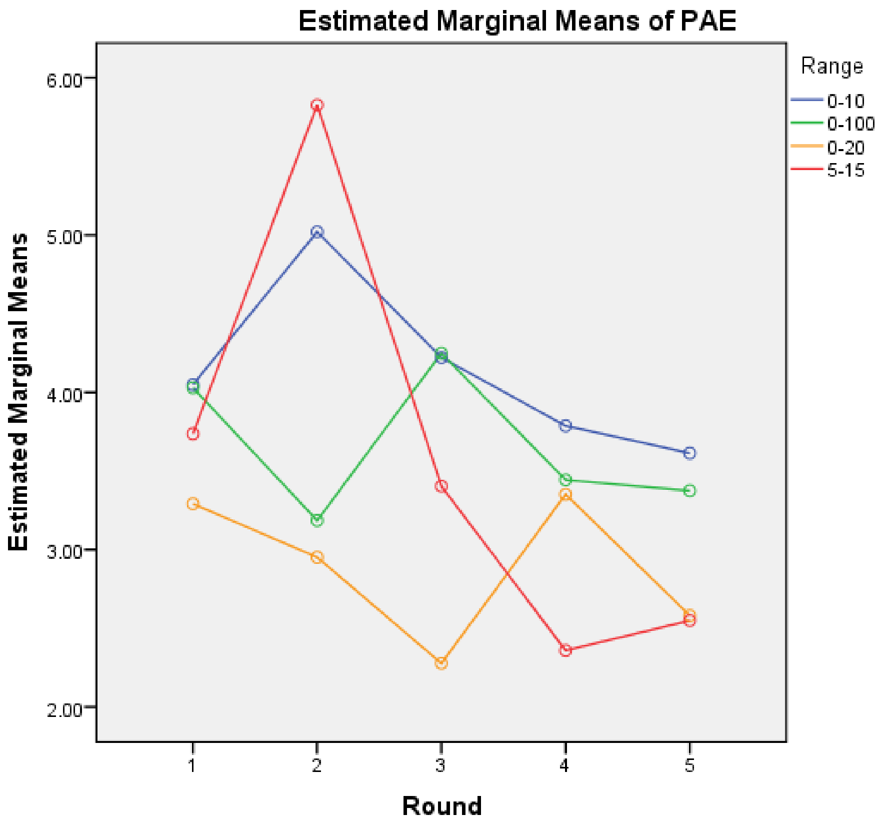


Figure 152 Mean PAE for each range, by round

This graph shows little correlation in between-rounds change for the different ranges. In particular, compared to R1, R2 estimations in ranges 0-10 and 5-15 became much *less* linearly accurate, whilst estimations in ranges 0-20 and 0-100 became much *more* linearly accurate.

The fact that there was no task and range interaction strengthens the argument that different number ranges themselves had a strong impact on estimations. In fact, for each range, there is striking similarity between estimation accuracy in T2 and T4. For example, whilst it is not clear from the quantitative data alone why Catharina's estimation in range 5-15 should have become suddenly *less* accurate in R2, the change occurred equally in T2 and T4 estimates, as shown in the graph below.

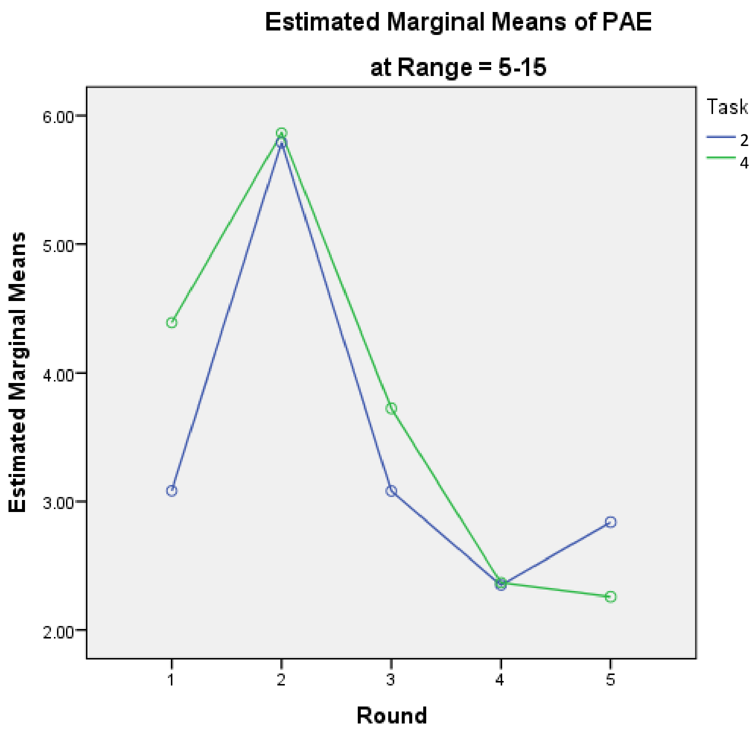


Figure 153 Mean PAE by round for T2 and T4 separately, range 5-15

9.3 RQ2: What strategies can be identified in children's interactions with number line estimation tasks?

The strategies identified in Catharina's T2 and T4 responses were the following:

- Reference to anchor point
    - Those included in the task environment (endpoints)
    - Those created or visualised by child
        - Midpoint
        - Quarter-point
    - Previous trials of the task
        - Within the same range
        - Within a different range
- Counting strategies
    - Count-on from LE to estimate
    - Count-on from estimate to anchor point (e.g. RE)
    - Count-back from RE
- Judgment using relative numerosities: particularly clear when estimate, or further strategies, appeared to be influenced by initial use of relative numerosity of target and task environment.

9.3.1 What patterns can be detected in the way children use or do not use these strategies?

Several patterns are evident from Catharina's case. Figure 154 shows the representations and strategies which Catharina demonstrated in estimation trials, compared by target type: whether the target was close to a left endpoint, midpoint, right endpoint, or quarter/three-quarter point of the number line. As in previous charts, representations and strategies are grouped according to Resnick's microstage account, and circles are coloured and sized to represent frequencies relative to all the other frequencies in the chart.

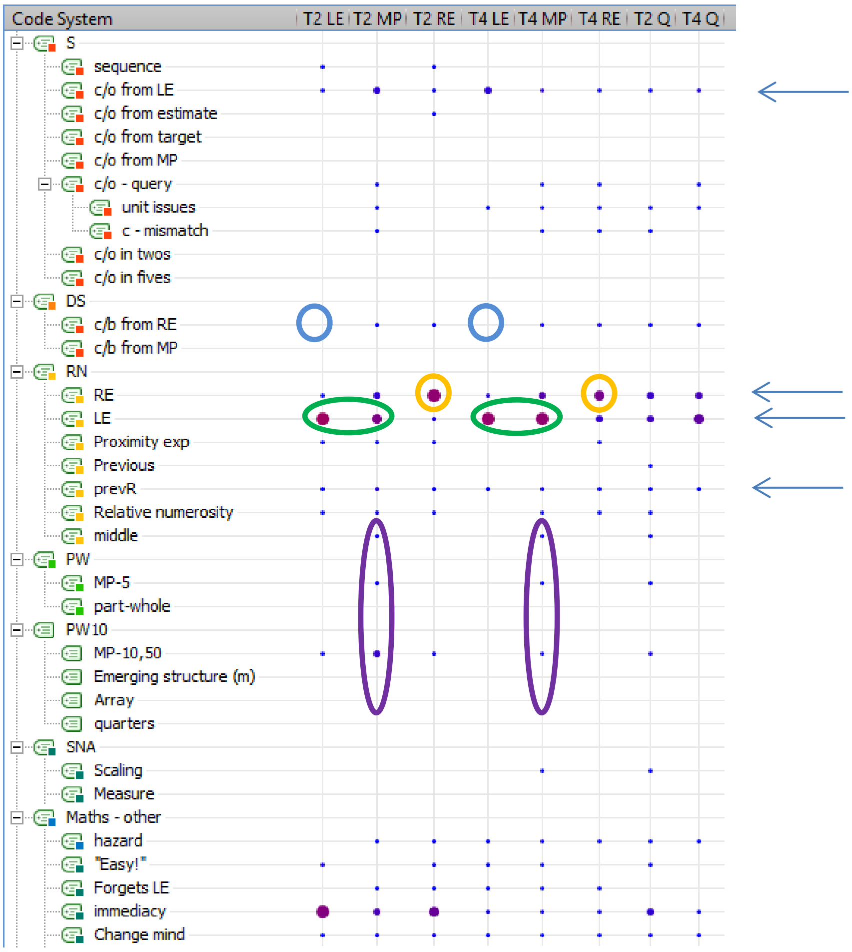


Figure 154 Catharina's strategies and representations by target type

This diagram shows four key representations that were observed for every category of target number (blue arrows), but also highlights some important variations between target type that appear to reflect flexibility in response to the mathematical features of trials.

The four types of representation that appeared for every target type are count-on from the left endpoint, references to left and right endpoints, and references to previous trials. There is no notable variation in frequency for references to previous trials, and there is no mathematical reason for this representation to be more or less prevalent for any particular target type. The frequency of references to the left and right endpoints, however, varied considerably. References to the left endpoint were more frequent for targets close to the left endpoint or midpoint (green) than for targets close to the right endpoint. References to the left endpoint were also more numerous than references to the right endpoint for targets close to the left endpoint or midpoint. Conversely, references to the right endpoint outnumbered left endpoint references for targets close to the right endpoint. Right endpoint references were also more frequent for targets near the right endpoint than for targets near the left endpoint or midpoint. These variations are in line with expectations based on the conditions of individual trials: it would be expected that references to the nearest endpoint would be dominant.

Other variations in representations and strategies are also in line with expectations based on the target type. Counting back from the right endpoint occurred for all target types except for targets close to the left endpoint (circled blue), for which count-back from right endpoint would have been a counter-intuitive choice. References to the midpoint were more common for targets close to the midpoint (purple) than for targets close to the left or right endpoints.

9.3.2 How do the strategies used relate to children’s estimation results, if at all?

In the remainder of this section, graphs of linear estimation accuracy are used to identify trials or rounds in which changes in children’s estimation results are evident. After identifying points of investigation, the graphs are then compared to charts showing the strategies and representations that occurred during the relevant estimation, to see whether connections can be drawn that may explain changes in estimation results.

9.3.2.1 T2, range 0-10

The following chart shows Catharina’s estimates for T2 in range 0-10.

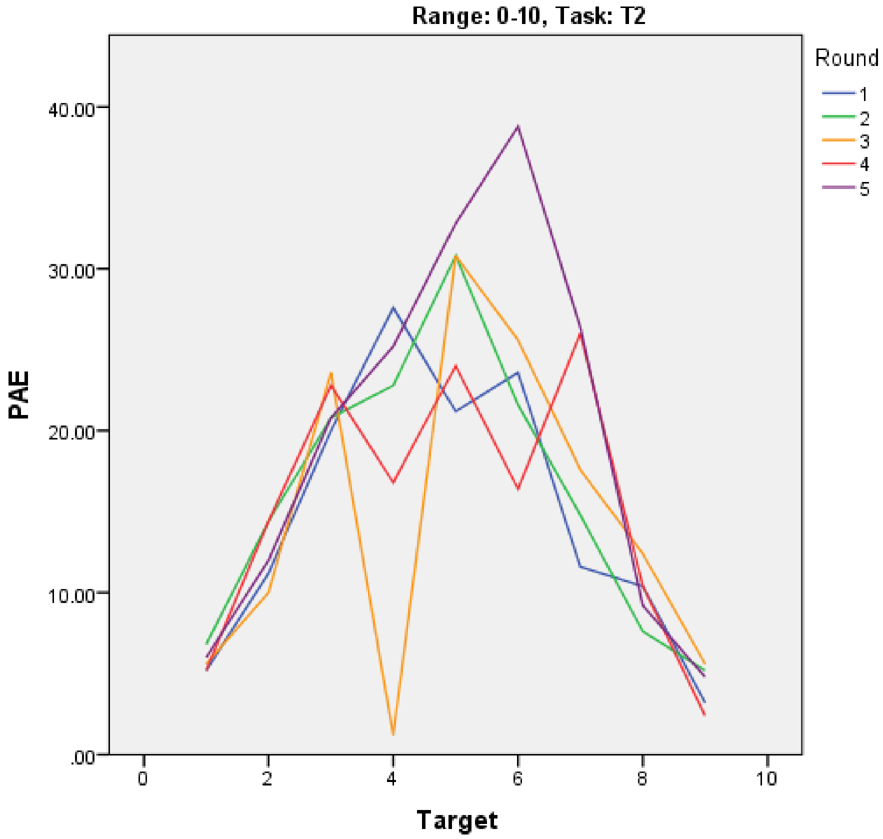


Figure 155 PAE of target estimates, T2 0-10

The graph demonstrates that T2 estimates on 0-10 were highly correlated between rounds, particularly for targets near the left endpoint (1, 2 and 3) and right endpoint (8 and 9). The questions arising for the qualitative data are:

- What was different about the estimation of target 4 in R3?
- Were similar representations and strategies observed for 1, 2, 3, 8 and 9 in the different interviews?
- What was different about the estimation of target 6 in R5?

Figure 156 T2 0-10, R1, R2, R4 and R5

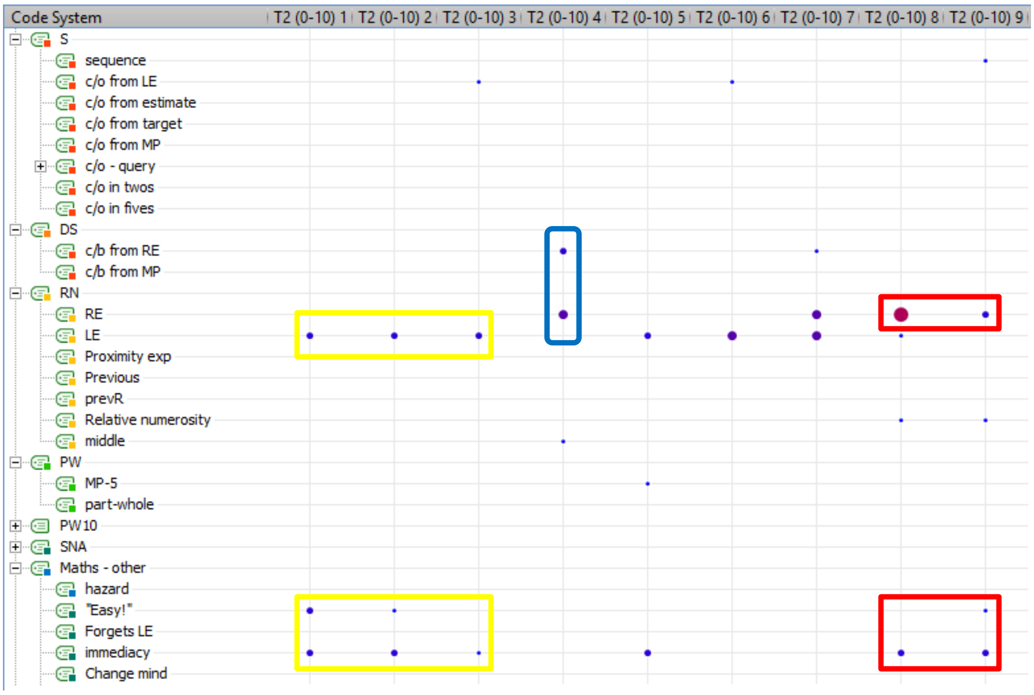


Figure 157 T2 0-10, R3

The above two charts reveal an answer to the first question, which is that the estimation of target 4 in R3 differed in its use of strategies focused on the right endpoint. In R1, R2, R4 and R5 (Figure 156), target 4 is estimated using count-on from left endpoint and reference to left endpoint (green box). In R3 however (Figure 157), target 4 is estimated using count-back from right endpoint (blue box).

The following series of charts plots Catharina's representations and strategies for T2 range 0-10 by individual round:

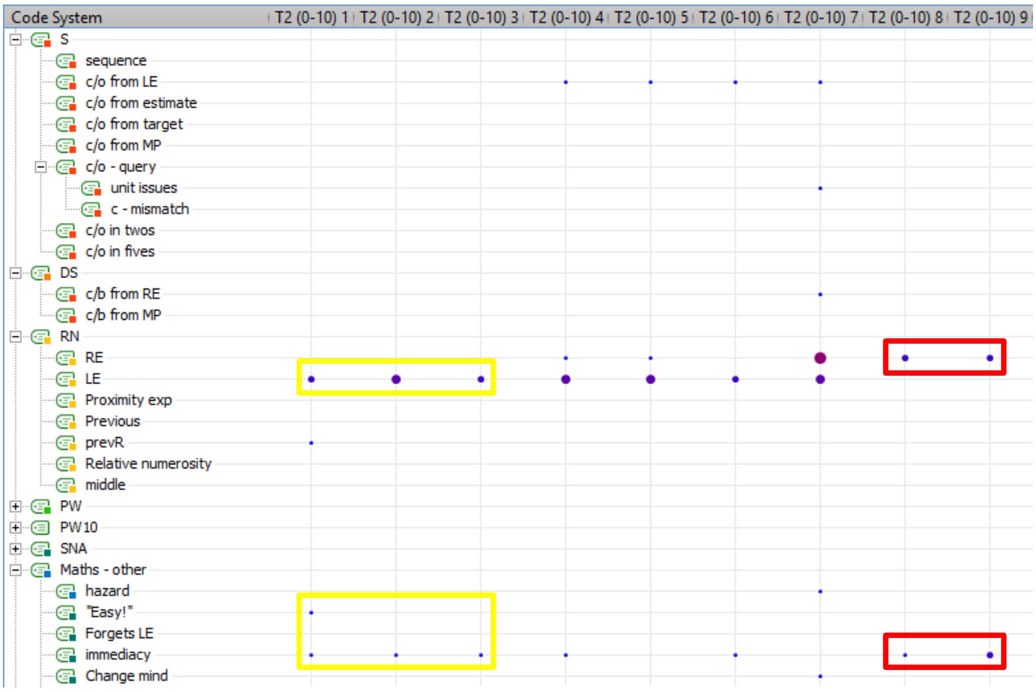


Figure 158 T2 0-10, R1

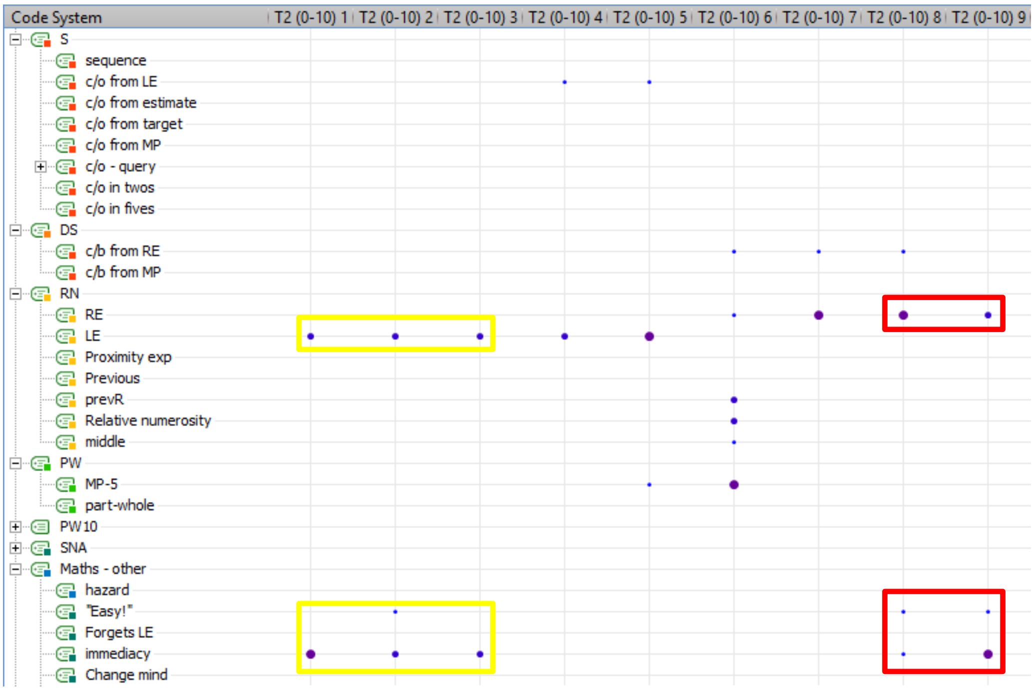


Figure 159 T2 0-10, R2

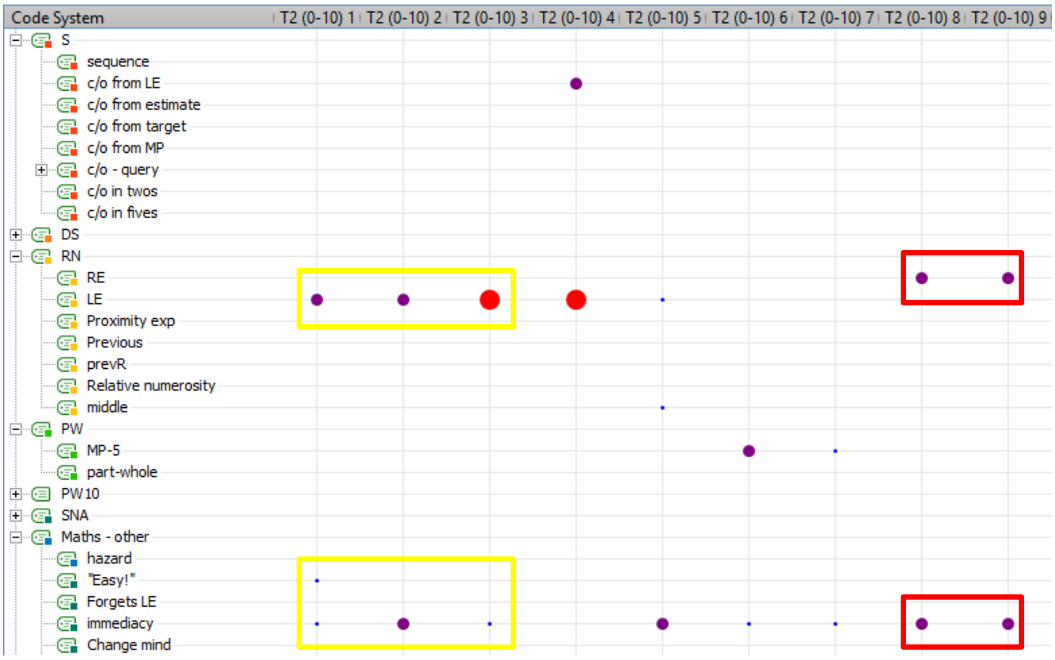


Figure 160 T2 0-10, R4

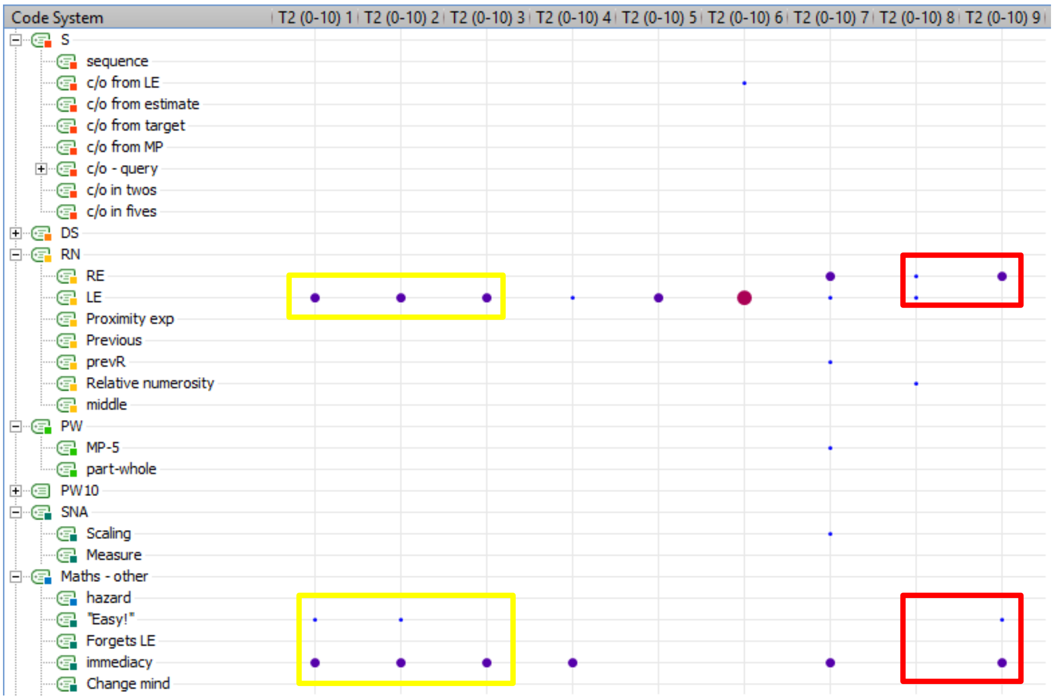


Figure 161 T2 0-10, R5

In answer to the second question raised by the estimation accuracy data, similar representations were indeed used for the estimations of targets 1, 2 and 3, and 8 and 9, in the different interviews. In every interview, targets 1, 2 and 3 were estimated using reference to the left endpoint and estimations were made with immediacy, strongly suggesting the use of a left endpoint anchor (see yellow boxing). In one instance (target 3 in R3), count-on from left endpoint was used. Targets 8 and 9, on the other hand, were estimated using reference to the right endpoint in every interview, and again with immediacy, suggesting the use of a right endpoint anchor (see red boxing). Again, count strategies were used only once (target 8 in R2).

The data presented here does not offer a clear answer to the final question, asking what was different about the estimation of target 6 in R5. The target was estimated using count-on from left endpoint, a strategy which was also seen in R1 and R3.

The next graph shows the percentage error (as opposed to absolute percentage error) of the same estimates:

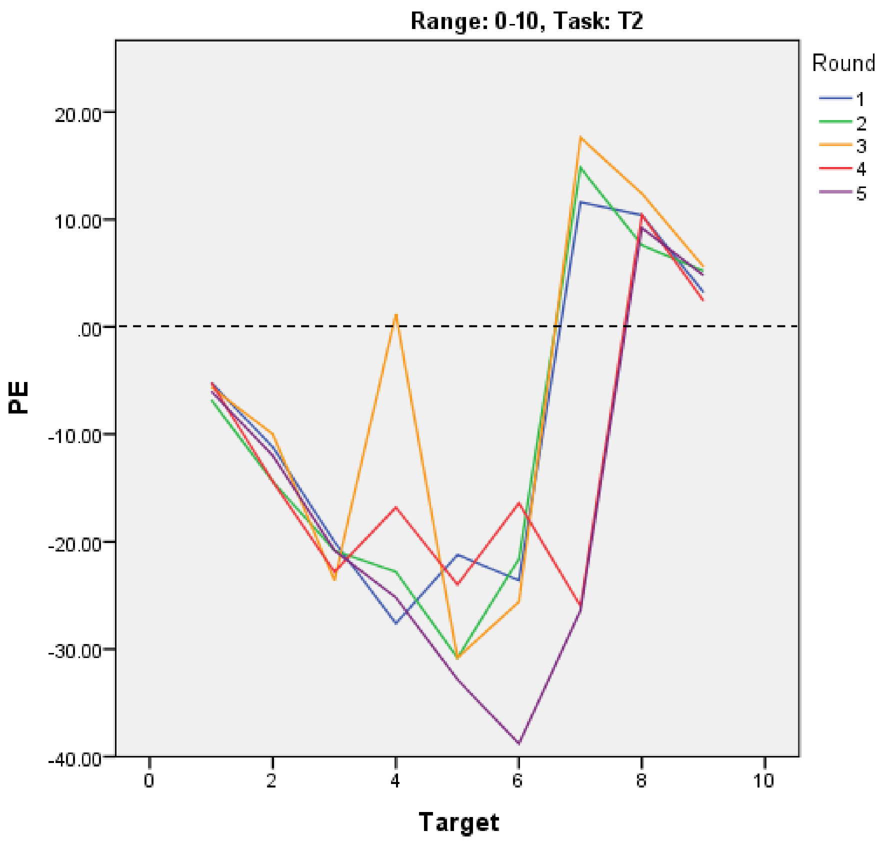


Figure 162 PE of T2 estimates, range 0-10

The graph of PE shows a clear pattern of underestimation for lower targets, followed by over-estimation for targets closer to the right endpoint. This is the reverse of the pattern predicted by proportional judgment models for a number-to-position estimation task. Furthermore, there are no targets in the middle of the range for which estimations became more accurate.

Instead, the estimation strategies that Catharina employed for each target, and their physical properties, provide a convincing explanation for the pattern shown above. The previous series of charts demonstrated that lower targets were estimated relative to the left endpoint, and higher targets relative to the right endpoint. This is compatible with error that increases with distance from either endpoint (which the above graph shows). The correlation between estimates and linear trends in the error indicate that Catharina's estimates implicitly represented a regular unit size, and the underestimation and overestimation indicate that this unit size was smaller than that required for a linear representation for this range. Examples of count strategies – e.g. target 3 in R3 – provide direct representation of unit size, and units were indeed smaller than required for linear representation.

9.3.2.2 T4, range 0-10

The graph below shows the PAE of T4 estimates for range 0-10.

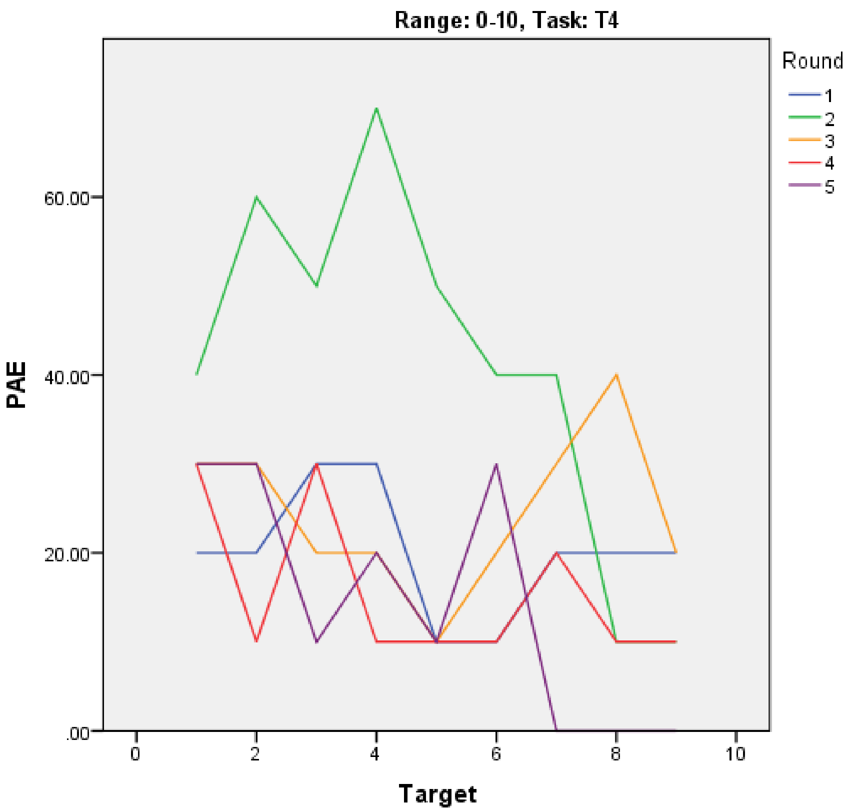


Figure 163 PAE of T4 estimates, range 0-10

The questions raised are the following:

- What was different in R2 compared to all other rounds?
- What could account for the highly accurate estimates of 7, 8 and 9 in R5?

Figure 164 shows the representations and strategies observed for T4 range 0-10 in R2.

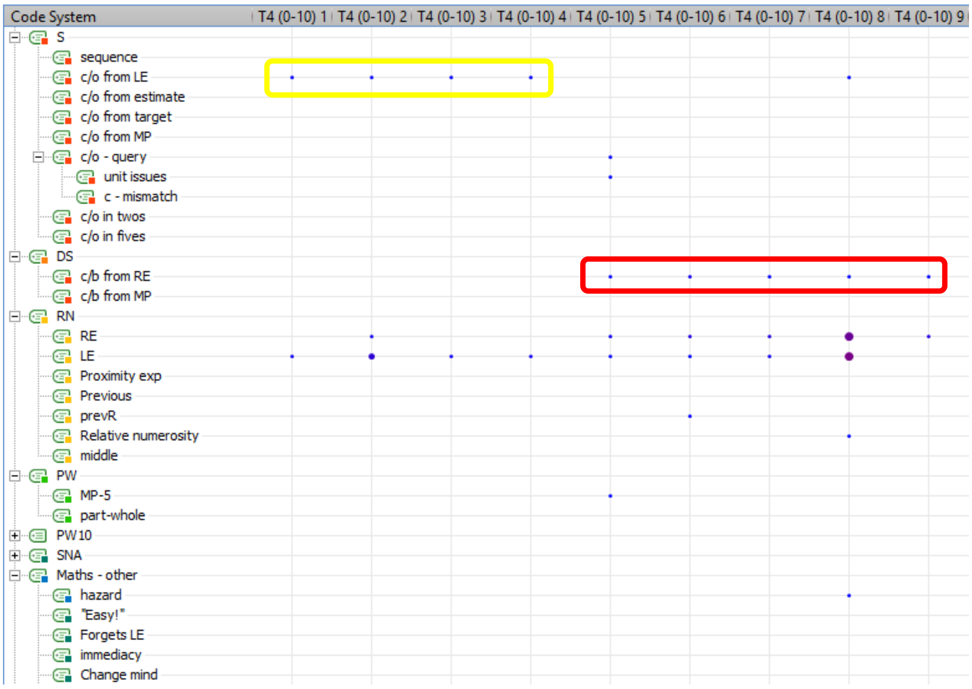


Figure 164 T4 0-10, R2

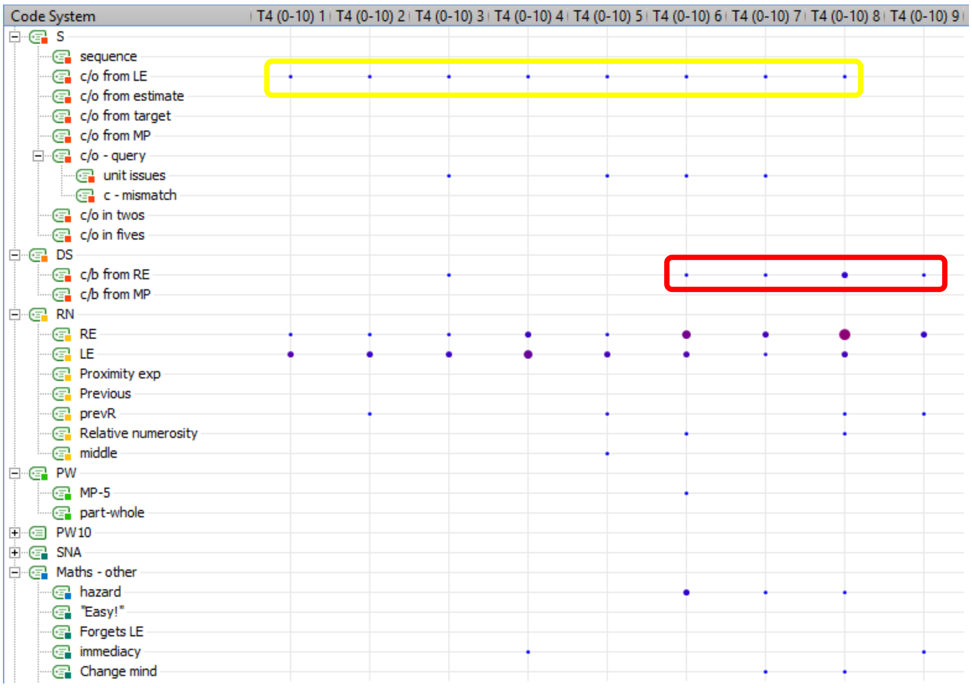


Figure 165 T4 0-10, R1, R3, R4 and R5

The above two charts do not provide any strong evidence for why estimates in R2 should have differed so markedly from estimates in other ranges. Estimates of lower target numbers were made using count-on from left endpoint (yellow) and left endpoint references, and estimates of higher target numbers were made using count-back from right endpoint (red) and right endpoint references, but this same pattern was observed in other rounds too. For this question, the most helpful data are the inscriptions Catharina made during T4 estimates in R2. In six out of the nine trials for T4 range 0-10 in R2, Catharina drew a sequence of dots as part of her count-on or count-back strategy, and these reveal a very small unit size. An example is shown below:

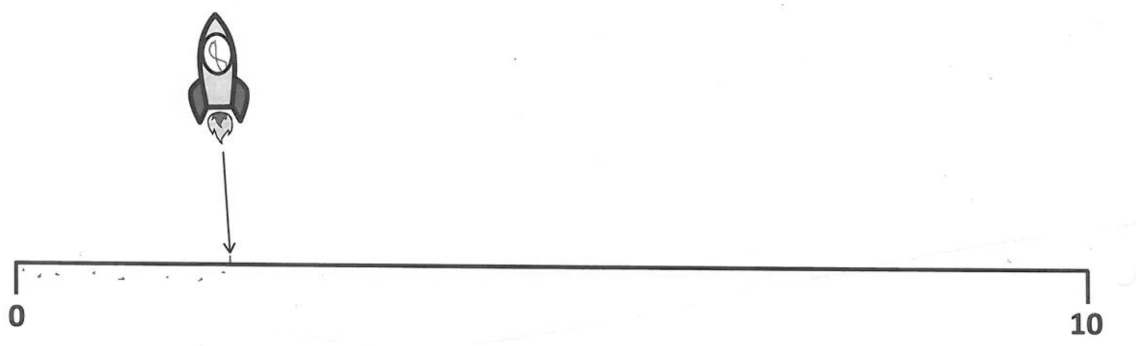


Figure 166 Catharina's T4 estimate of target 2, range 0-10, R2

The conclusion from this data is that whilst Catharina used count strategies in all rounds, in R2 she for some reason chose a very small unit size and used it consistently. In this way, the same type of strategy led to less linearly accurate estimations than in other rounds, where higher linear accuracy arising from count strategies (and this includes only trials in which the estimate was actually made according to the result of the count) indicates that unit size was necessarily more appropriately scaled.

The following chart, which shows representations and strategies for R5 only, was produced in order to investigate the second question – on the accuracy of estimation for targets 7-9 in R5:

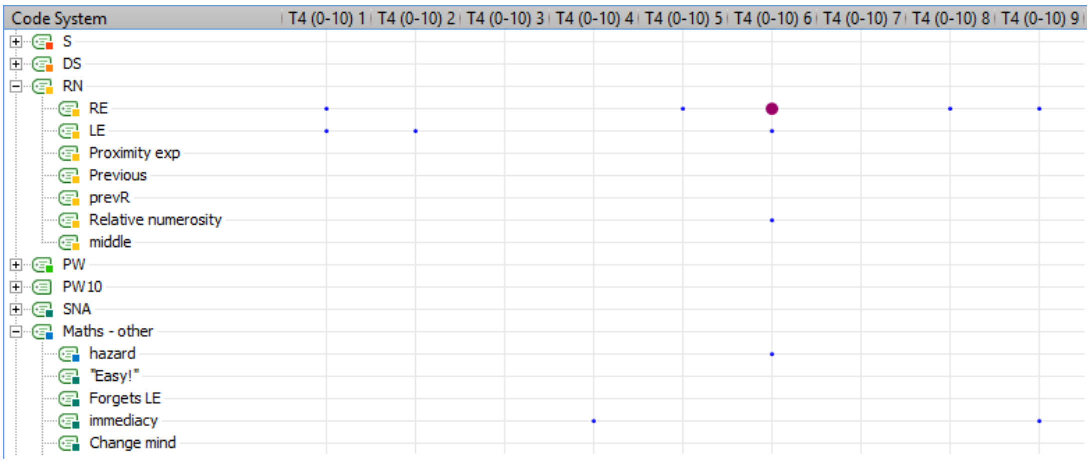


Figure 167 T4 0-10, R5

This data suggests that a right endpoint anchor was used to estimate target 9, and probably target 8, but it does not itself explain why the estimates should have been more linearly accurate than those in other rounds.

The following graph shows the estimation PE (as opposed to PAE) for this task:

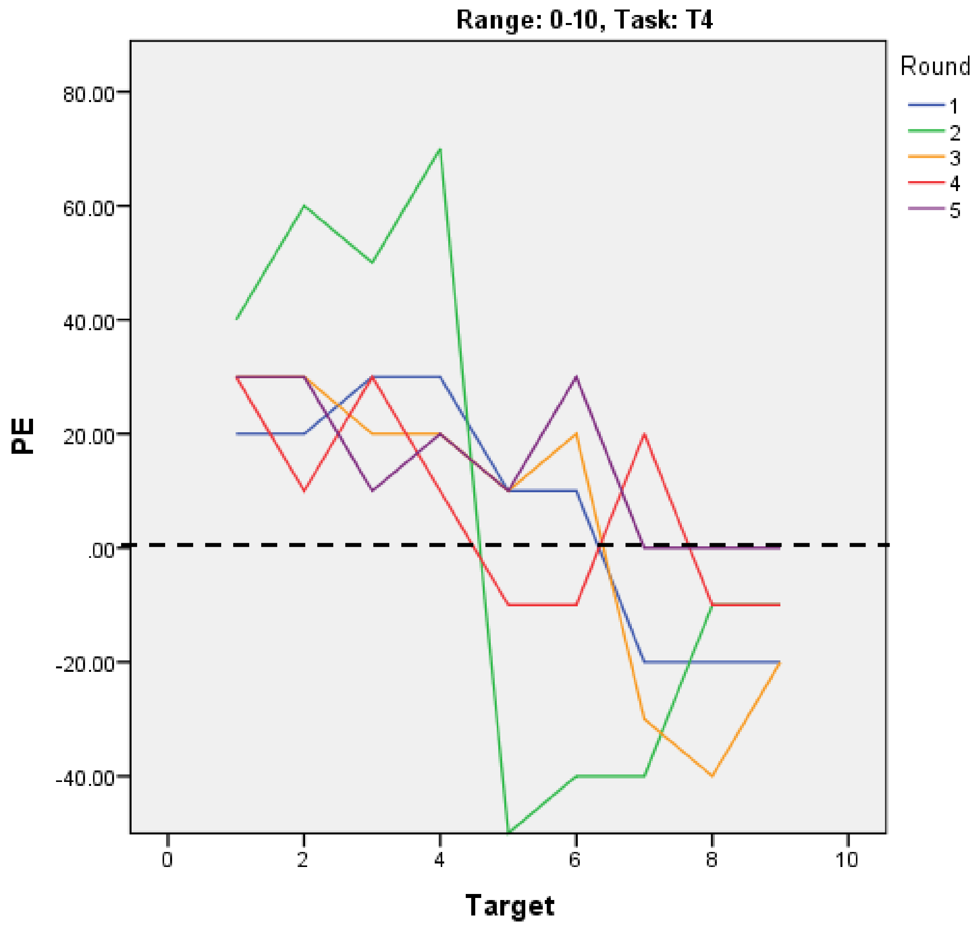


Figure 168 PE of T4 estimates, range 0-10

This graph indicates over-estimation for lower target numbers, followed by underestimation for higher target numbers. Once again, this is the reverse of the pattern predicted by a proportional judgment model. The pattern of over- followed by under-estimation occurs in R2 estimates just as in other rounds, except with larger magnitudes of estimation error.

This finding is consistent with the explanation suggested in the previous section for T2 range 0-10, namely that the over- and under-estimation pattern witnessed results from Catharina's choice of strategies and their physical properties. As in T2, the strategies depend upon the left endpoint (lower target numbers) or right endpoint (higher target numbers), and explicitly or implicitly represent a small unit size – in R2, particularly tiny.

9.3.2.3 T2, range 0-20

The following graph shows Catharina's T2 estimates for the range 0-20.

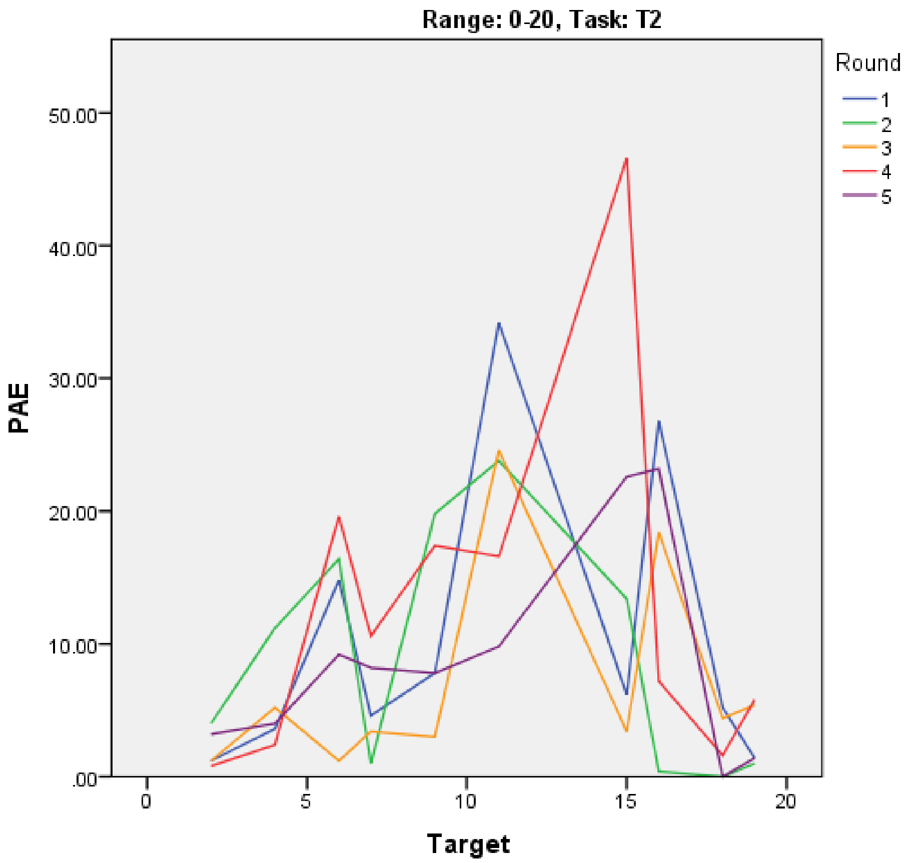


Figure 169 PAE of target estimates, T2 0-20

The questions that this poses for the qualitative data are:

- Is there evidence for use of a left endpoint anchor for targets 2 and 4?
- Is there evidence for use of a right endpoint anchor for targets 18 and 19?
- What was different about the estimation of target 15 in R4?

To explore these questions, the following charts show Catharina's representations and strategies for T2 0-20 estimates for R1-R3 and R5 together, and then for R4 separately.

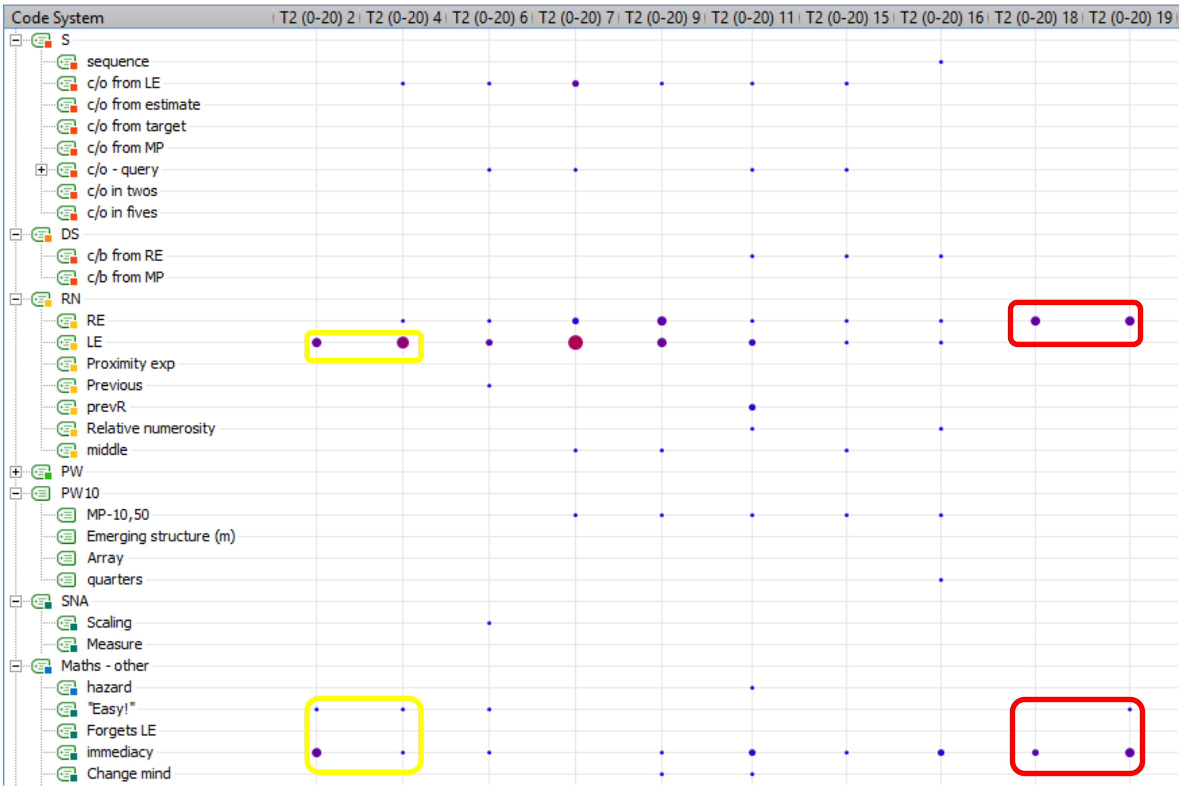


Figure 170 T2 0-20, R1-R3 and R5

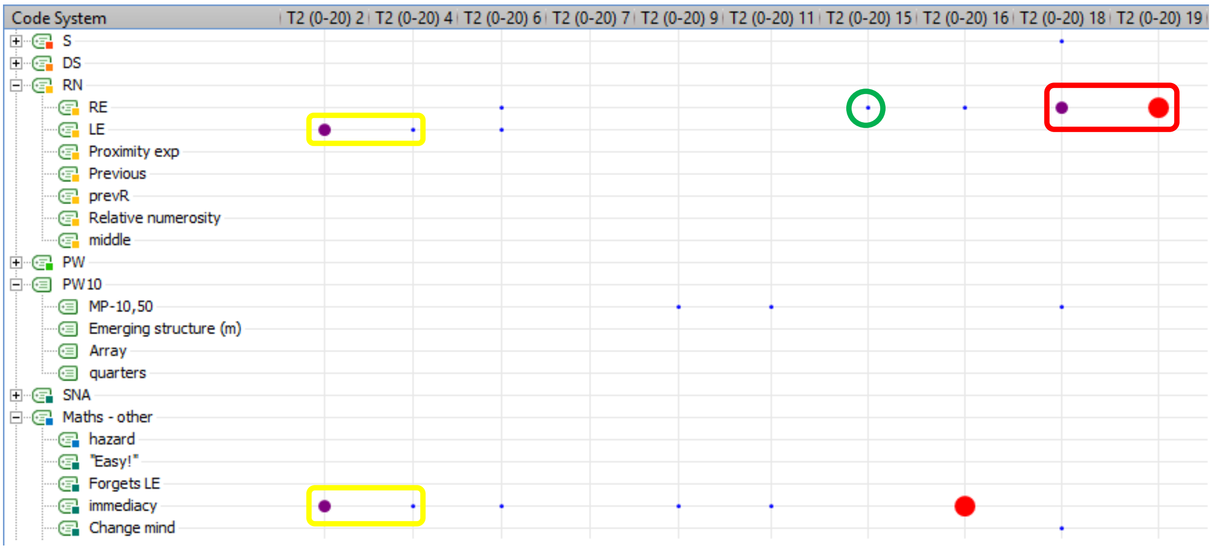


Figure 171 T2 0-20, R4

The evidence for use of endpoint anchors is strong: for targets 2 and 4, Catharina referred to the left endpoint almost exclusively, and with immediacy and comments that it was “easy”. For targets 18 and 19, Catharina only made reference to the right endpoint, and there were again observations of immediacy and comments that the trials were “easy”.

For target 15 in R4, Catharina referred only to the right endpoint (circled green). This representation was also observed in other rounds, but other rounds also included a wide variety of other number structure. This data does not provide a clear answer for why the estimation in R4 should have been markedly less accurate than in other rounds.

Looking at the PE graph of the same task condition, it becomes clear that all but one of Catharina's estimations in the range 0-20 were underestimations. This finding differs from both T2 and T4 for range 0-10, in which there was both under- and over-estimation.

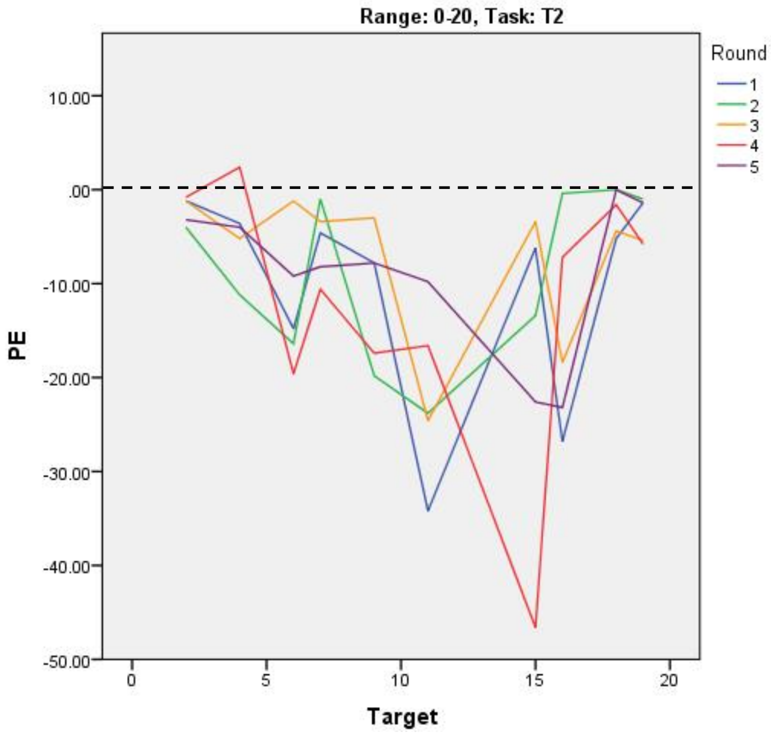


Figure 172 PE of T2 estimates, range 0-20

The conclusion is that this under-representation followed from Catharina's use of left endpoint oriented strategies, with an implicit unit size too small to result in linear representation of the range 0-20. The charts of representations and strategies indicate that left endpoint oriented strategies were predominant except for at targets 18 and 19, and this corresponds to the underestimation followed by accuracy at targets 18 and 19 demonstrated by the graph above.

9.3.2.4 T2, range 5-15

The following graph shows Catharina's T2 estimates for range 5-15:

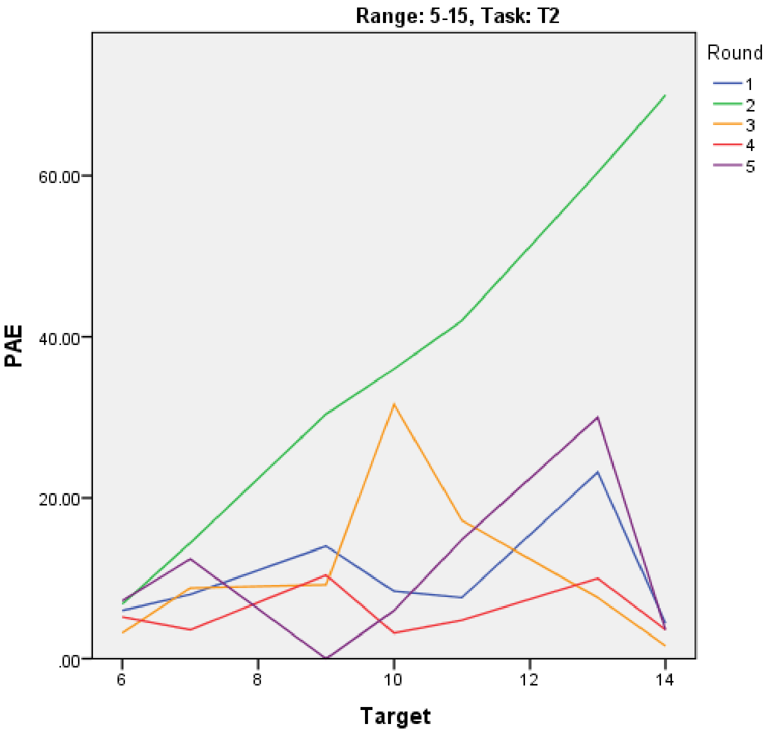


Figure 173 PAE of target estimates, T2 5-15

Two main questions arise from this graph:

- What evidence is there for the use of left and right endpoint anchors, for targets 6, 7 and 14?
- What was different about the estimations in R2?

To investigate these questions, the following charts display the observed strategies and representations for T2 range 5-15 for R1 and R3-R5 together, and then R2 separately.

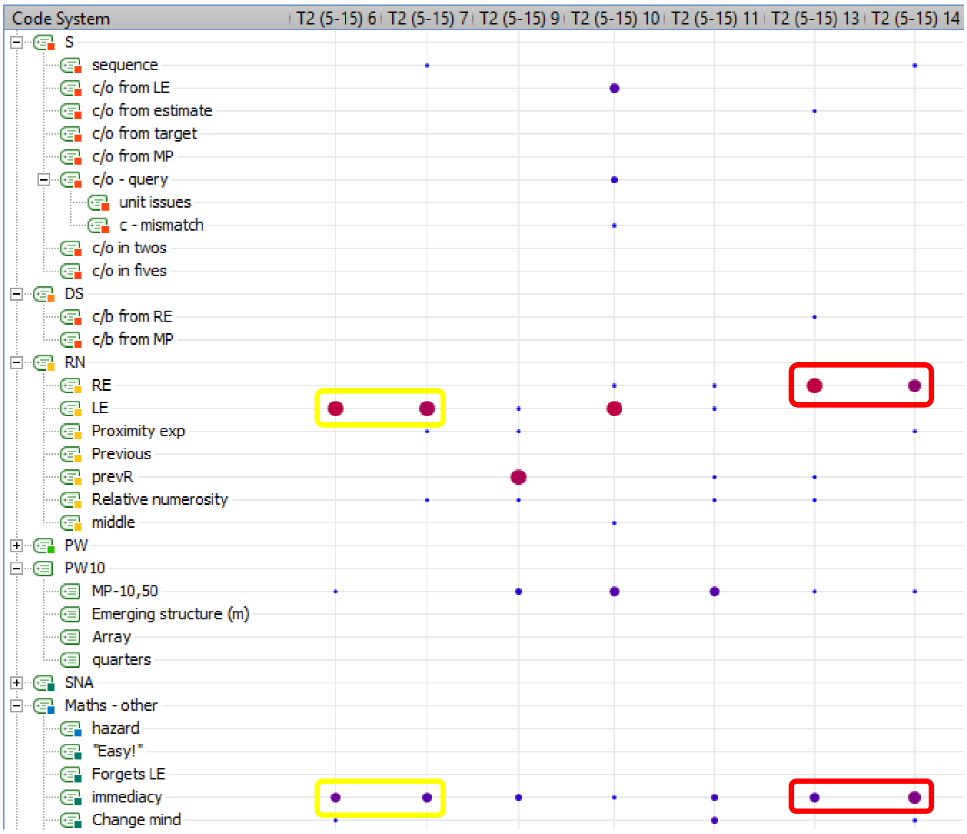


Figure 174 T2 5-15, R1 and R3-R5

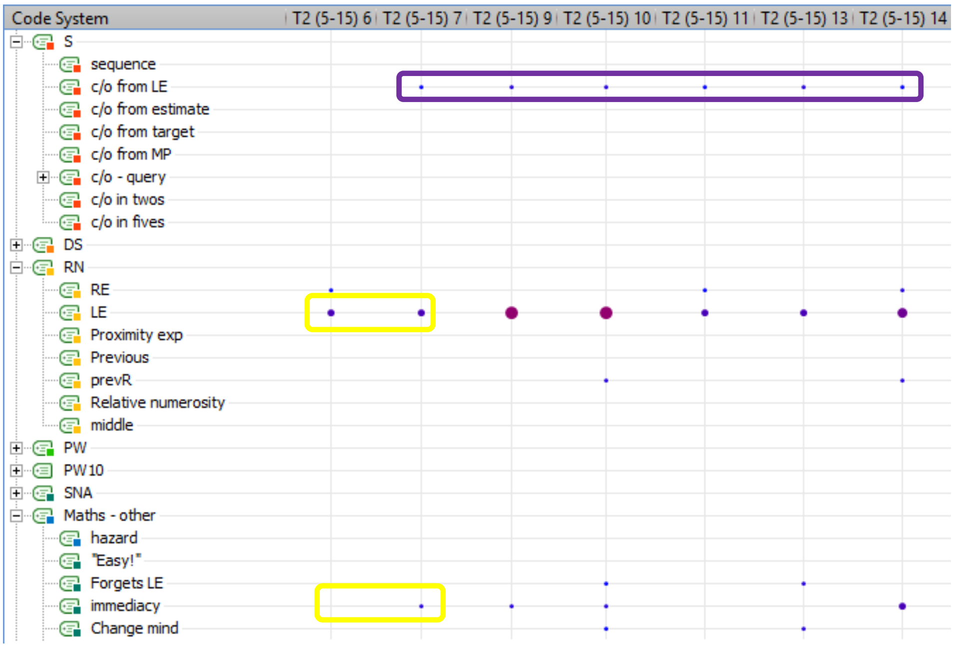


Figure 175 T2 5-15, R2

The evidence (yellow) for the use of a left endpoint anchor for targets 6 and 7 is good: references to the left endpoint predominate in all rounds, and there are observations of immediacy in the estimations. For target 14, there is good evidence (red) for use of a right endpoint anchor in all rounds except R2: there are frequent references to the right endpoint and observations of immediacy, although other representations are also observed to a lesser extent. For R1, R3, R4 and R5 (but not R2) there is also good evidence for use of a right endpoint for target 13, even though the graph reveals that this did not lead to particularly accurate estimates.

The chart showing R2 estimations reveals that count-on from left endpoint was used for all targets except target 6 (see purple box). The graph below of percentage error shows that this led to underestimation for all targets in R2, and that this underestimation increased proportionally with distance from the left endpoint.

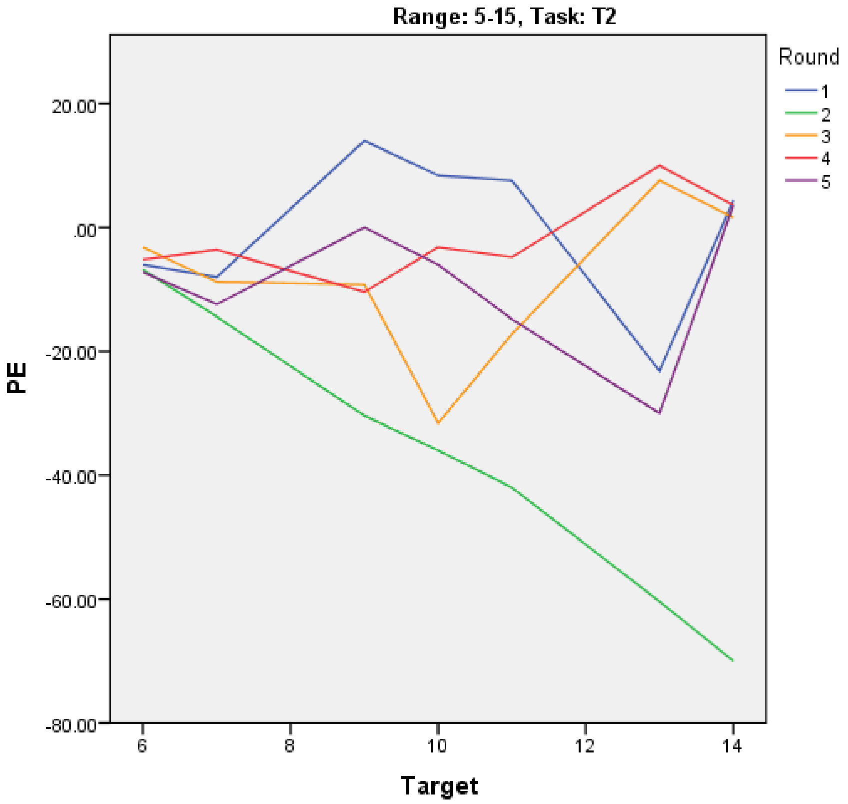
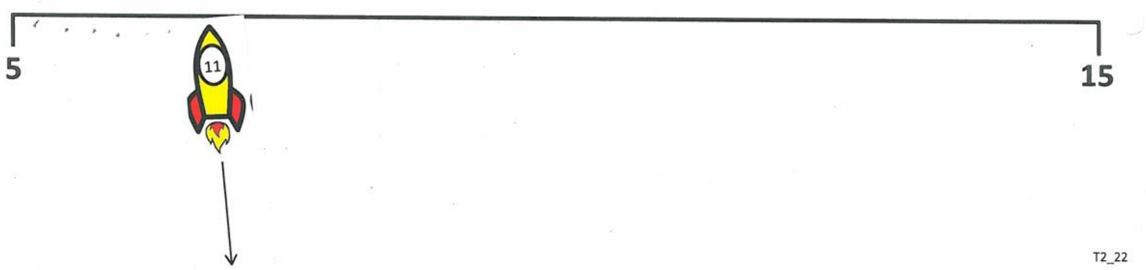


Figure 176 PE of target estimates, T2 5-15

In this instance, the qualitative data provides a convincing explanation for the pattern of estimation error recorded. As in other task conditions in R2, Catharina marked dots showing the unit size for the count-on strategies used in T2 5-15, and these confirm what the resulting underestimation already indicates, that the unit size chosen was too small to result in a linear representation for the range. As an example, the following shows Catharina's estimation of target 13 for T2 range 5-15 in R2:



9.3.2.5 T2, range 0-100

This graph shows the PAE of Catharina's T2 estimates in the range 0-100:

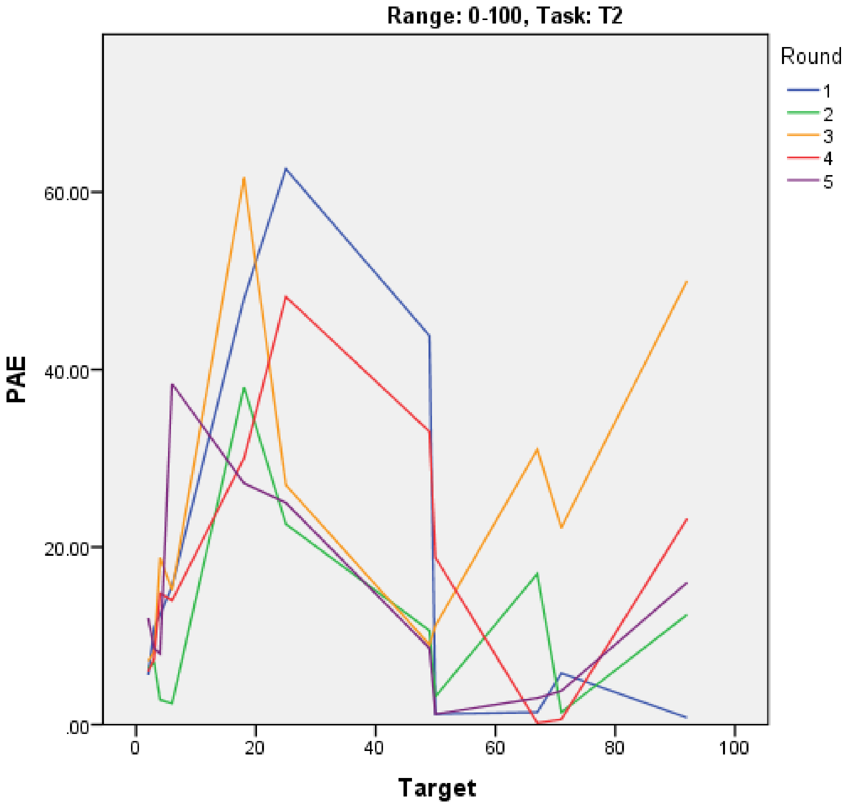


Figure 177 PAE of target estimates, T2 0-100

Once again, points of disparity between rounds are interesting. The questions that arise are:

- Is there evidence of midpoint anchor use for target 50 (and to a lesser extent 49)?
- Can qualitative data account for the less accurate estimates that occurred for targets 18, 67 and 92 in R3?

To investigate these questions, the following charts show Catharina's representations and strategies for T2 range 0-100, for R1, R2, R4 and R5, and then R3 separately.

With respect to the first question, there is evidence for the use of a midpoint anchor for targets 49 and 50, although this evidence is not as conclusive as for endpoint anchors in other rounds. During estimation of targets 49 and 50, there were references to the midpoint accompanied by observations of immediacy (green), and an absence of count-on or count-back strategies from either endpoint (orange boxes).

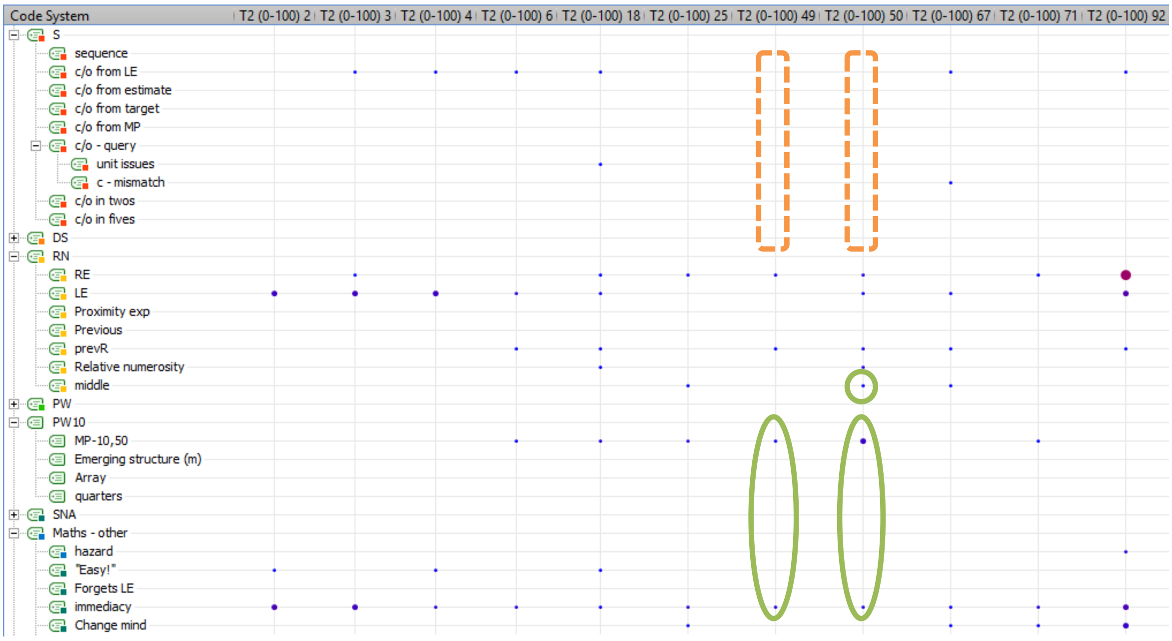


Figure 178 T2 0-100, R1, R2, R4 and R5

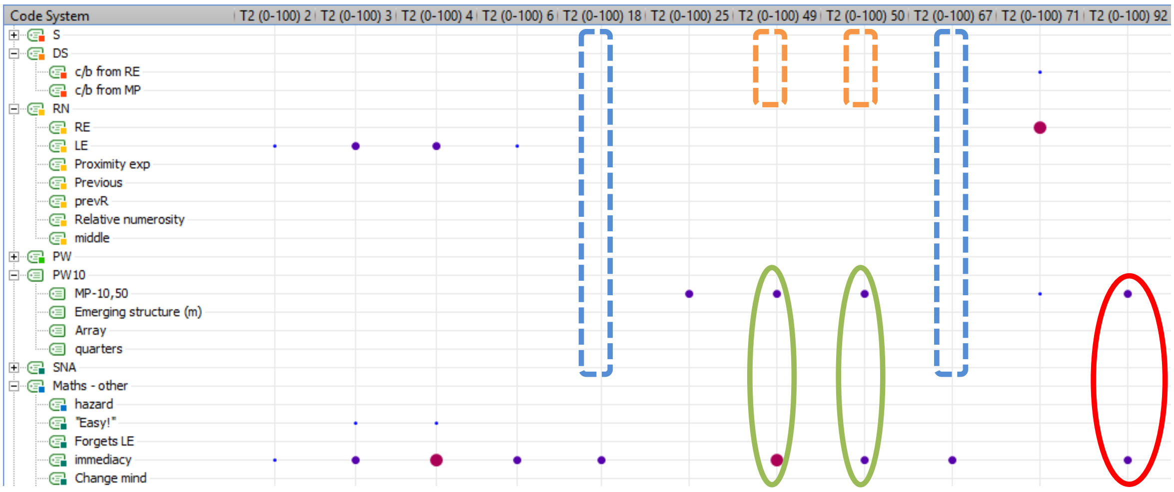


Figure 179 T2 0-100, R3

With respect to the second question, there are two observations that could help explain the particular inaccuracy of estimation for targets 18, 67 and 92 in R3. Firstly, target 92 appears to have been estimated using a midpoint anchor (red), and the likelihood of accurately estimating 92 from such an anchor point is low. Secondly, for targets 18 and 67 (see blue boxes) there were no representations of number structure at all, even references to the endpoints of the given line, only an immediate estimate of each.

The following graph of PE (percentage error) allows examination of over- and under-estimation patterns. The initial over-estimation followed by accuracy around the midpoint, and then under-estimation near the right endpoint, is as predicted by a proportion-judgment account of number line estimation. This pattern is in contrast to the patterns of over- and under-estimation in other number ranges.

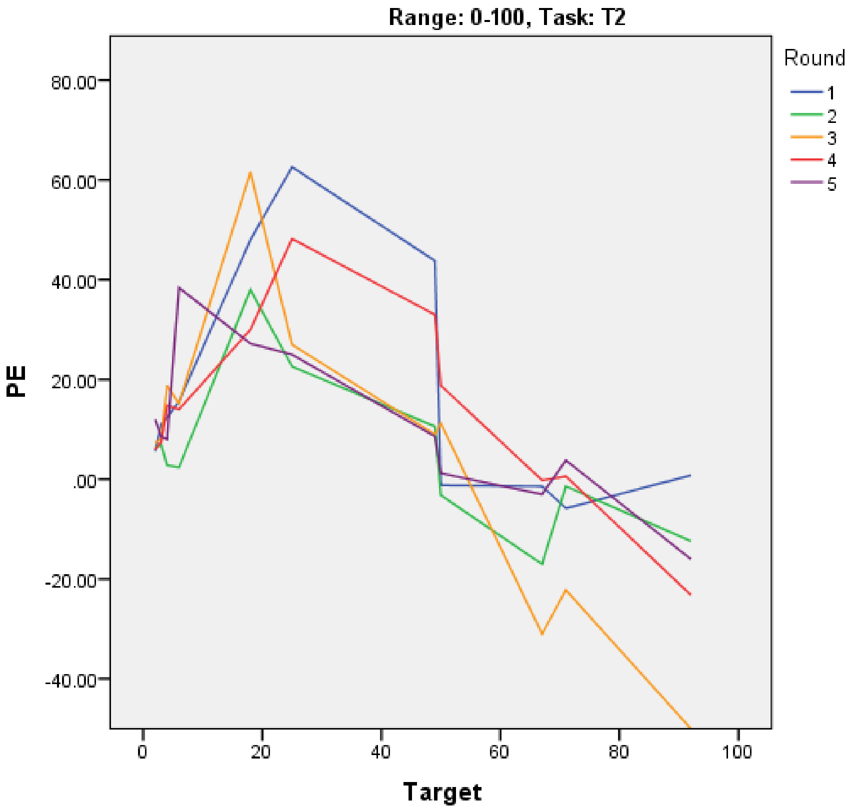


Figure 180 PE of target estimates, T2 0-100

9.3.2.6 T4, range 0-100

The following graph shows the PAE of Catharina's T4 estimates for range 0-100:

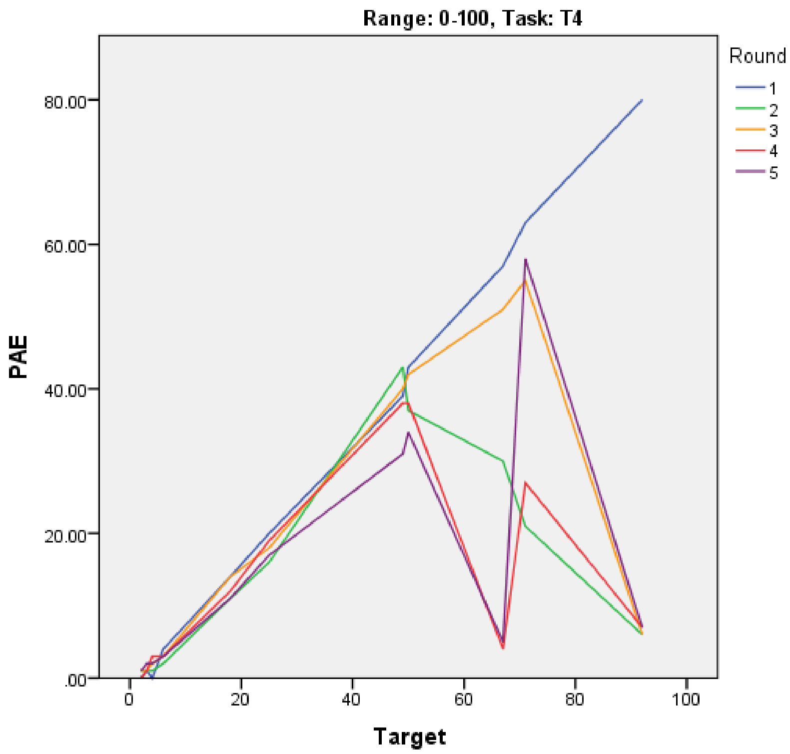


Figure 181 PAE of target estimates, T4 0-100

The questions arising are:

- What was different about estimates in R1 compared to R2-R5?
- Is there qualitative evidence for the use of a right endpoint anchor for R2-R5?

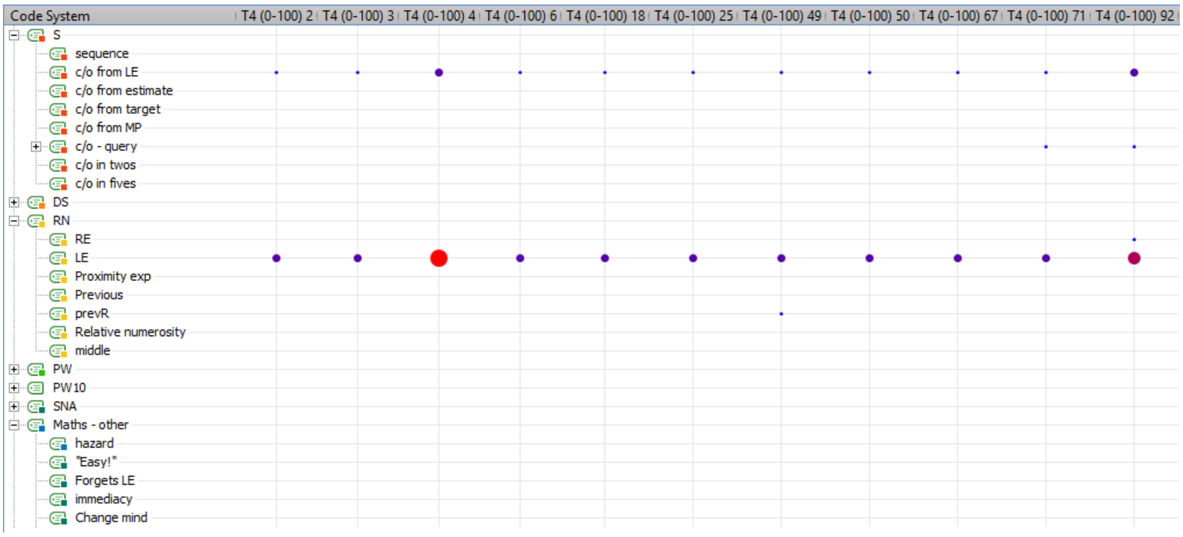


Figure 182 T4 0-100, R1

The chart above reveals that in R1, Catharina used count-on from the left endpoint for every trial, and referred to the right endpoint only once (for target 92). This offers a likely explanation for estimation error proportionally increasing with target magnitude.

In R2, R3, R4 and R5, Catharina again made frequent use of count-on from left endpoint strategies, and made frequent references to the left endpoint. However, she also referred to the right endpoint in all but two trials, and additionally included count-back from right endpoint strategies for the three highest targets (orange), suggesting some flexibility in response to the individual trial’s conditions.

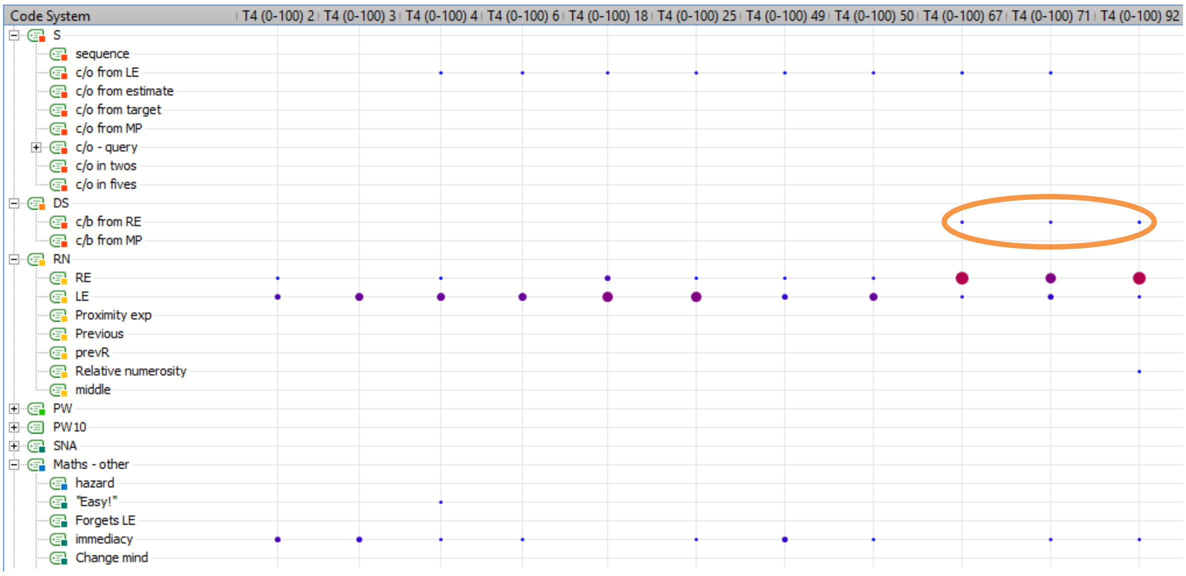


Figure 183 T4 0-100, R2 - R5

The graph below shows the PE of Catharina's estimates in this task. It reveals that the error in R1 estimations was entirely underestimation, and this follows from Catharina's use of count-on from left endpoint. The error increases linearly with target magnitude, which reflects that Catharina used a regular unit size, that was too large to achieve a linear representation for the range 0-100; with these large units, Catharina reached each target rocket with fewer unit steps from the left endpoint than the linear representation would have revealed.

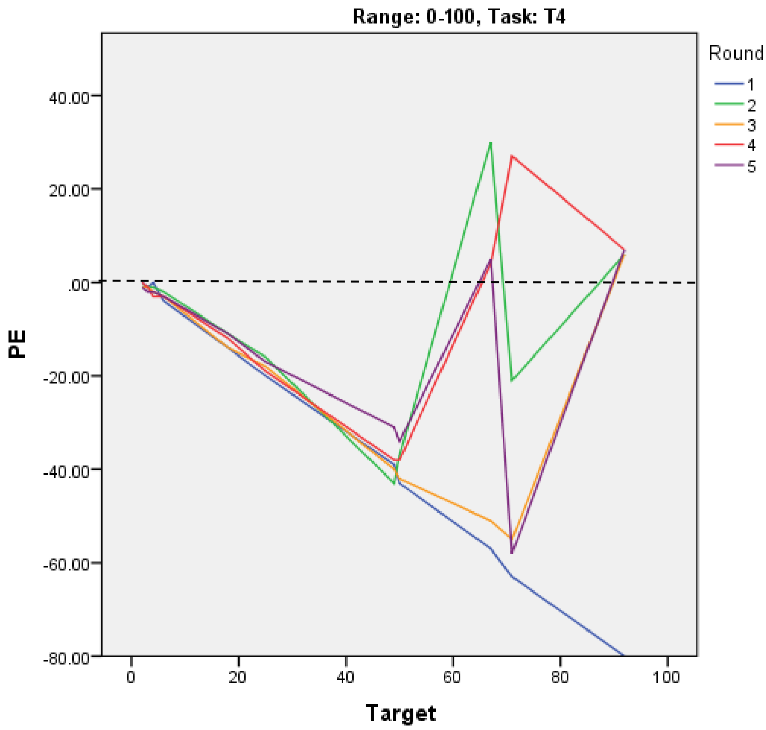


Figure 184 PE of target estimates, T4 0-100

This research question has drawn out several important findings. In this case study, there are numerous points at which there is a very plausible link between Catharina's estimation results and the observed representations and strategies for that estimation. This is particularly true of targets close to the right and left endpoints: there is good evidence that Catharina made frequent use of the endpoints as 'anchor' points in her estimation of nearby targets, and on these occasions her estimations were usually markedly more accurate than for other targets in the same range.

Catharina's case also provides particularly interesting evidence on the use of count strategies. Since Catharina varied the unit size of count strategies between interview rounds, but tended to use a very consistent unit size within rounds and often applied count strategies to an entire set of trials, Catharina's case provides especially rich evidence.

Patterns of over- and under-estimation can be convincingly linked to the count strategies that Catharina used. Her results emphasise again how the same strategy (count from left endpoint) can have notably more or less accurate outcomes depending on the precise details, in this case, unit size. Her use of count strategies also revealed a particular loyalty or attachment to them, which was also a very notable finding in Marta's case. The example below shows Catharina's T4 estimation of target 7, in range 0-10 during R2:

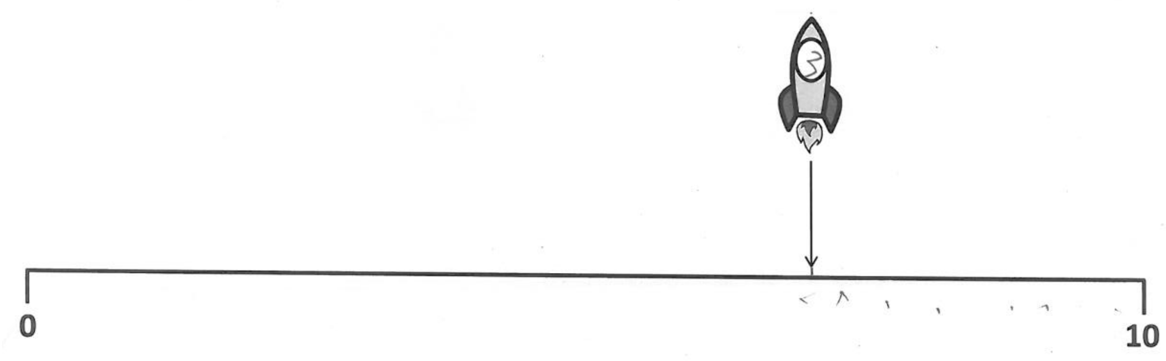


Figure 185 Catharina T4 R2

The interpretation of the above is that Catharina chose the right endpoint as point of reference for the target, since it was the closest anchor point to the blank rocket. However, by counting back with inappropriately scaled unit 'hops', Catharina reached an estimate that was implausible by her earlier reasoning (namely "this target is somewhat close to ten"). She nevertheless abided by this answer.

9.4 RQ3: How do young children's cognitive representations of number change during their first year of formal schooling?

This question involves examining the changes evidenced by both the qualitative and quantitative estimation data collected from Catharina's interviews.

The diagram below summarises the aspects of number structure that Catharina represented in each round of interviews. As in previous sections, the representations and strategies are grouped according to Resnick (1983)'s developmental theory of number understanding. The size and colour of the circles represents frequencies relative to the same category in other rounds, not to frequencies in other categories. The data are presented in this form in order to best examine the changes over time within categories.

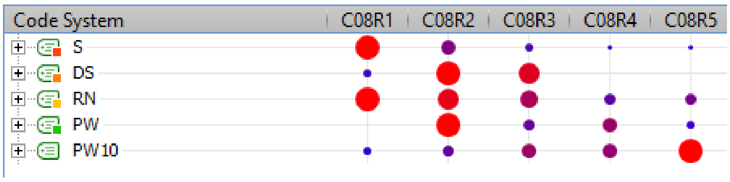


Figure 186 Number structure represented by round

The first point to be noted from this chart is that Catharina represented most categories of number structure in every interview round. The exception was decreasing sequence structure, which was not represented at all in R4 or R5. The second important point to be noted is that the chart reveals an overall trend of decreasing direct representation of the least sophisticated number structure (increasing sequence structure) at the same time as increasing direct representation of the most sophisticated number structure (part-whole structure, including base-ten). Representation of relative numerosity also decreased overall across the five interview rounds, but the change was less marked than for increasing sequence structure. The remaining two categories of number structure show mixed patterns of increasing and decreasing frequency over time. This overall trend is the same as that observed in Patrick's case, and as in Patrick's case, the hypothesis put forward to explain the trend is that as Catharina gained confidence and fluency with more sophisticated number structure, it was no longer necessary or efficient to directly represent the more basic aspects of number structure as frequently as she did initially.

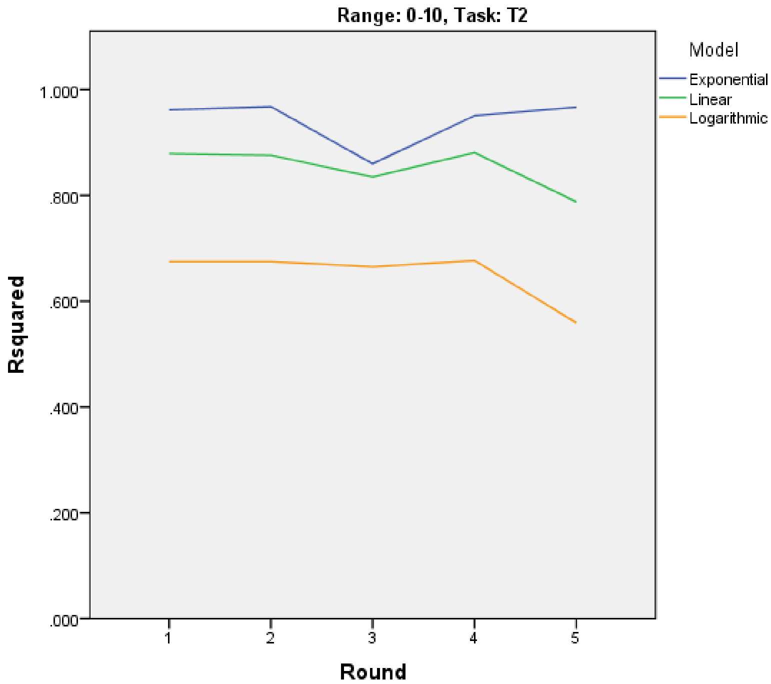
9.4.1 In what ways does evidence support or not support the log-linear hypothesis?

To recap the predictions of the logarithmic-linear shift account of number line estimation, it expects earlier estimations to be best fit by a logarithmic model in the case of number-to-position estimations (T2) and an exponential model in the case of position-to-number estimations (T4). Later estimations are expected to be increasingly better fit by linear models in both types of task. Furthermore, the shift from a logarithmic (or exponential) to linear model for a given number range may be expected to occur suddenly. The shift is expected to occur earlier for smaller, more familiar number ranges (i.e. 0-10) than larger, less familiar number ranges (i.e. 0-100).

Catharina's case provides very limited support for the logarithmic-linear shift account of number line estimation. The ANOVA analysis of linear accuracy (PAE) reported in RQ1c revealed that

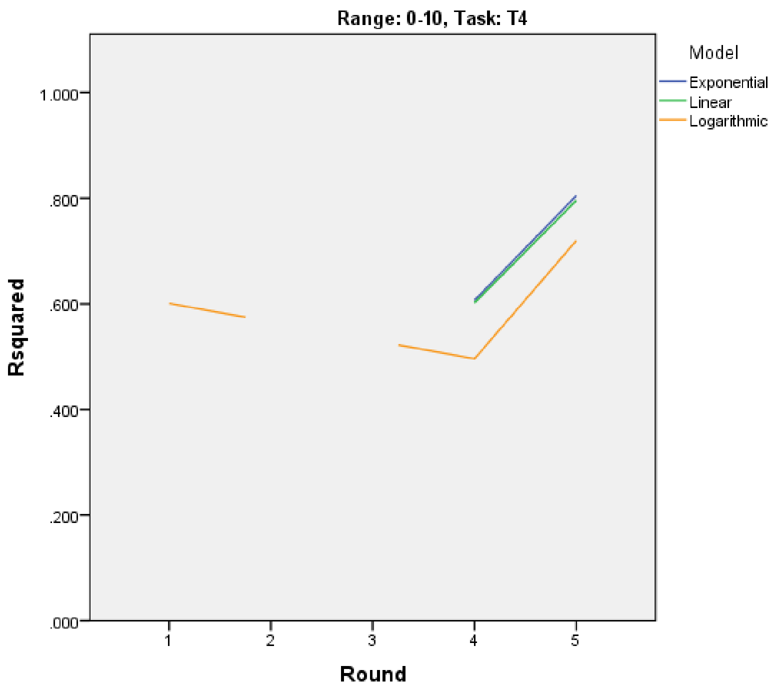
Catharina's number line estimations did become overall more linearly accurate with each interview round. However, in terms of more detailed view, Catharina's estimations did not tend to fit the patterns predicted by the log-linear shift account.

The following series of graphs plots the model fits (in the form of R^2 measure for each model) for linear, logarithmic and exponential models for each task condition in each round. For each task condition, a short comment explains whether the model fits are in line with the predictions of the log-linear account. Models that did not reach significance (at 5%) are excluded.



These model fits are not in line with the log-linear shift account: the logarithmic and linear models both provided a poorer fit to the data than the exponential model, and later estimations are in fact fit less well by a linear model than earlier ones.

Figure 187



These model fits are again not in line with the log-linear shift account, which would expect an exponential initially, and later on linear model, to provide the best fit for this task condition.

Figure 188

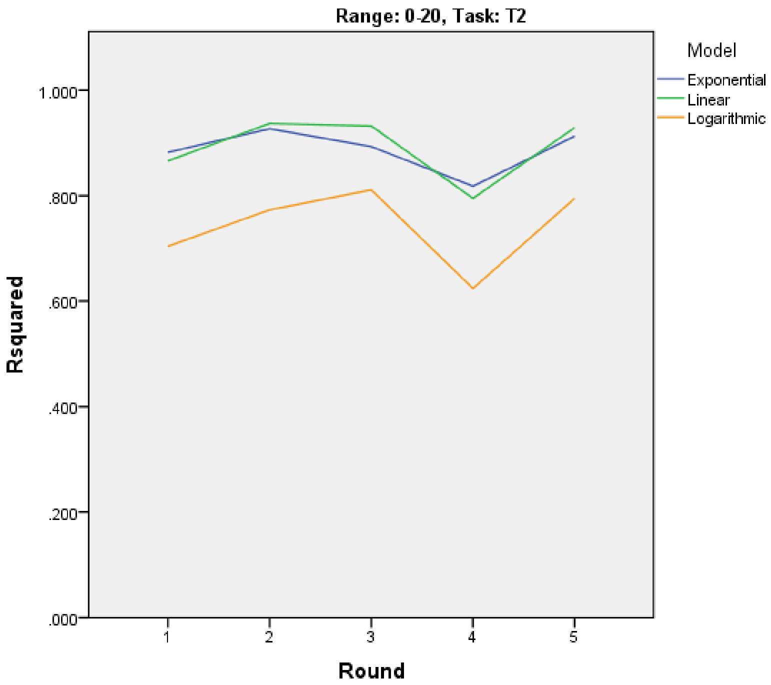


Figure 189

These model fits are not in line with the log-linear shift account, which expects a logarithmic and then linear model to provide the best fit.

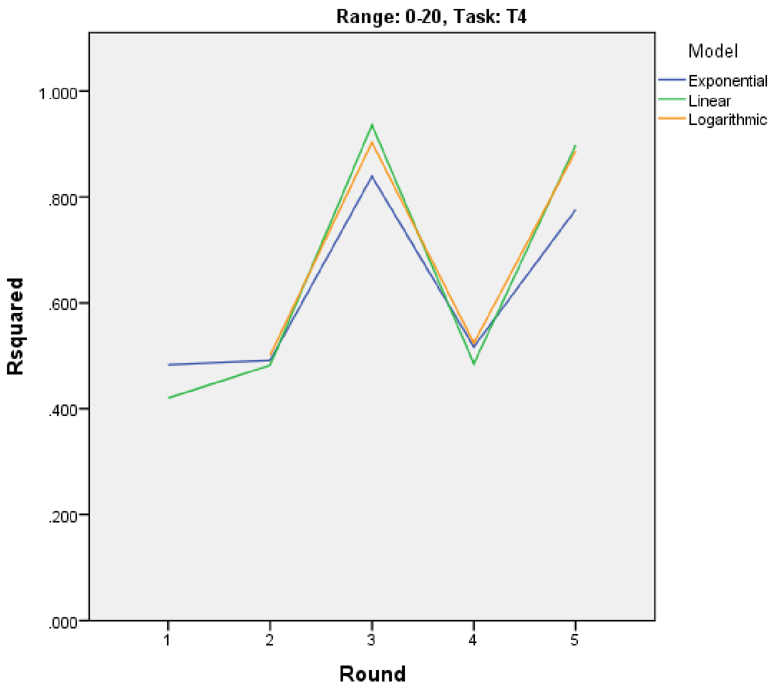
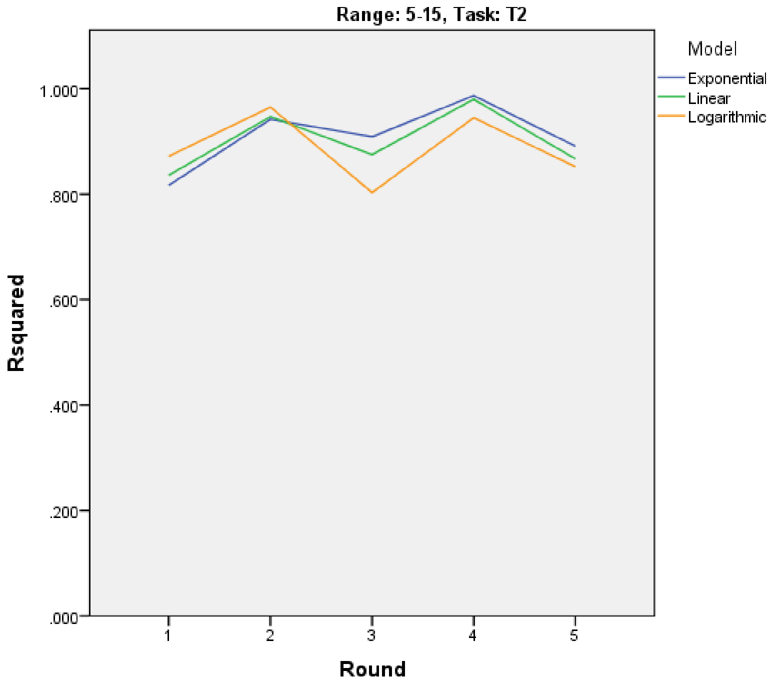


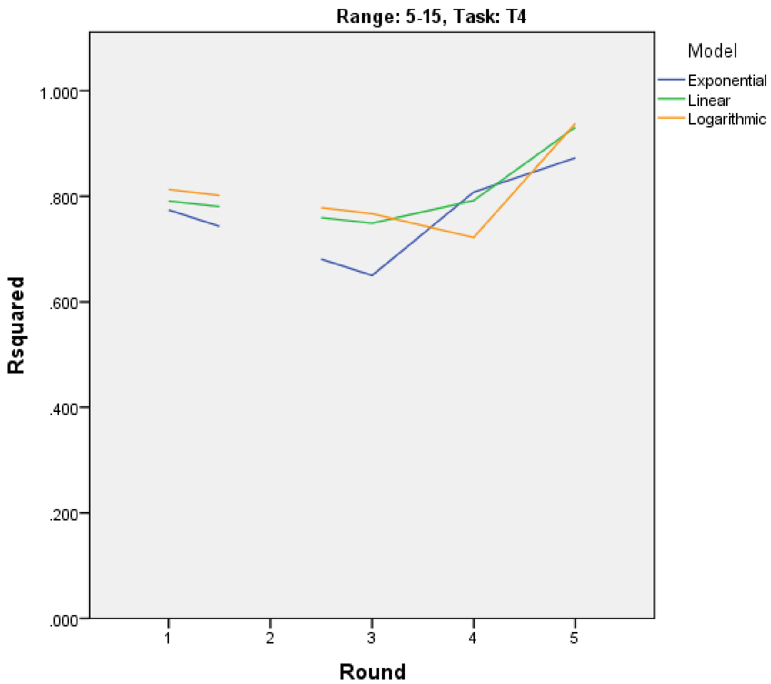
Figure 190

These model fits are not in line with the log-linear shift account, and in this case, there is very little difference between model fits at all.



These model fits are not in line with the log-linear shift account. Again, there is very little difference between model fits.

Figure 191



These model fits are not in line with the log-linear shift account. Whilst a linear model provides increasingly good fit, the same can be said of both other models too.

Figure 192

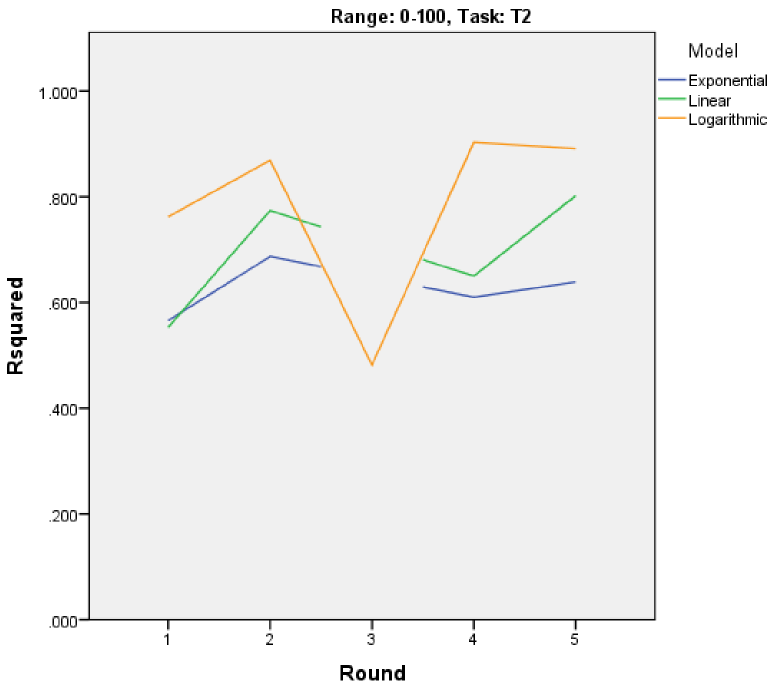


Figure 193

A logarithmic model provides the best fit for this data, which is in line with log-linear predictions for T2 estimations. There is also increasingly good fit from the linear model, though it does not occur suddenly and does not overtake the logarithmic model in terms of fit.

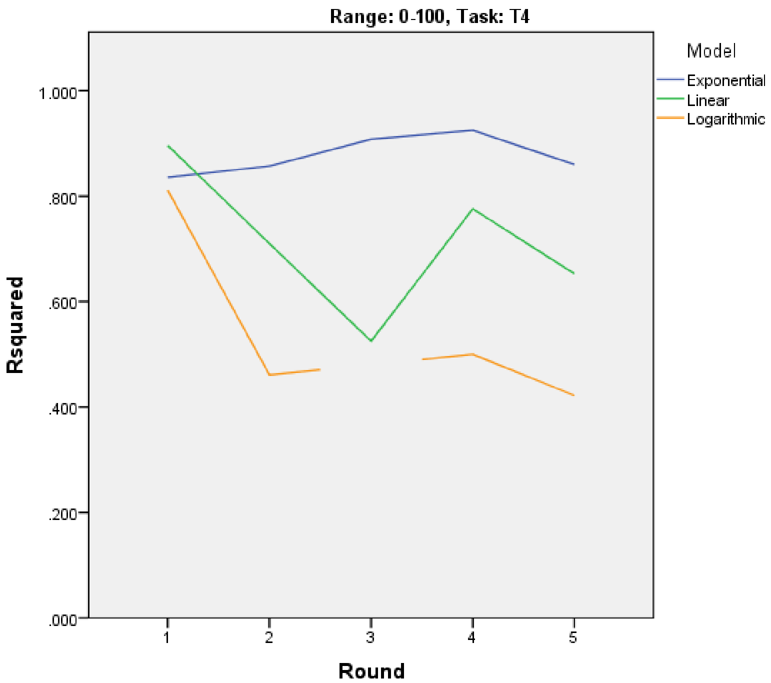


Figure 194

The overall dominance of the exponential model, especially compared to the logarithmic model, is as the log-linear shift account predicts for T4 estimations. However, the decrease in model fit for the linear model does not fit into the explanations of the log-linear account.

In summary, Catharina's case does not provide good support for the log-linear hypothesis. The extent of the convincing agreement between the two is the significant increase in linearity of number line estimations, but this is a widely documented phenomenon and is not uniquely predicted by the log-linear shift account.

9.4.1.1 *Alternative accounts*

In RQ2b, plotting PAE by target number demonstrated that Catharina's estimations were in general most accurate for targets near to anchor points, and least accurate furthest away from anchor points, as predicted by proportional judgment accounts of number line estimation. This question also showed that the anchor points used by Catharina varied. In some cases (e.g. T2, range 5-15, R2) there is good evidence that Catharina only took account of the left endpoint. In the majority of cases (e.g. T2, range 0-10, all rounds) Catharina took account of both left and right endpoints in her estimations. In some task conditions (e.g. T2, range 0-100, R5), Catharina also appeared to use a midpoint as an anchor point, and achieved good linear accuracy for estimates close to this. The order in which Catharina included anchor points is in line with the specific predictions of Ashcraft and Moore (2012): the left endpoint is included as an anchor point always, then the right endpoint, and then a midpoint.

9.4.2 What is the intra-child variability of children's numerical magnitude representations in estimation tasks at different times?

A key measure of the variability of children's number representations is variance of estimation error (PAE) in different task conditions at different times.

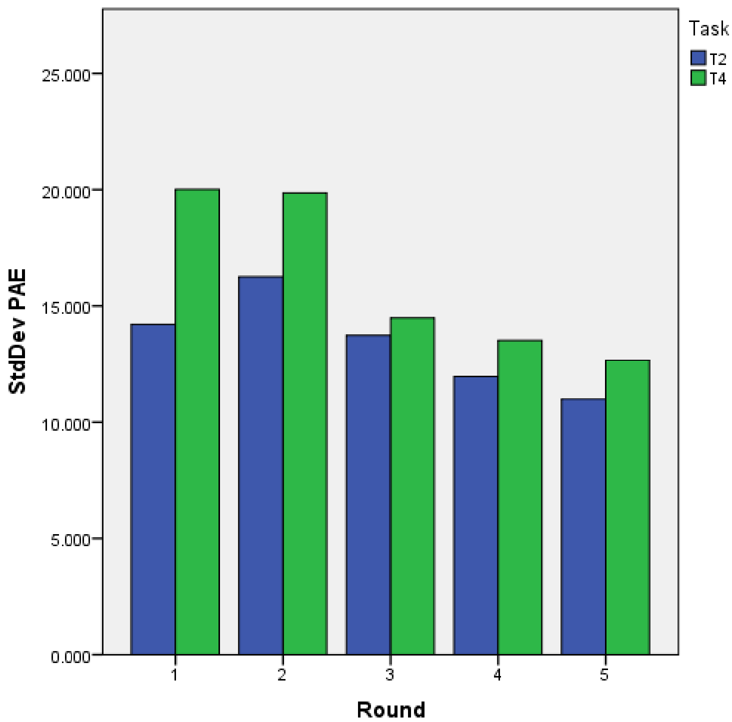


Figure 195 Standard deviation in PAE for T2 and T4 by round

This graph above shows that in each interview round, the standard deviation of PAE was higher for T4 estimates compared to T2 estimates. The variability of estimation accuracy decreased over the course of the five interviews, apart from a small increase in R2, and this mirrors the pattern of linear estimation error (see Figure 151).

9.4.3 Can trajectories or patterns of change be deduced, in terms of changes in how children cognitively represent number?

Catharina's case illustrates two overall trends across the course of the five interviews, both of which were broadly expected and also seen in Patrick's case study. The first, as discussed in RQ3a, is that Catharina represented more sophisticated aspects of number structure more frequently, and represented less sophisticated aspects of number structure less frequently (see Figure 186). As in Patrick's case, there is no evidence that this change occurred suddenly; it appeared gradually between each round of interviews.

The second trend, as explained in RQ1c, is that the linear accuracy of number line estimations increased significantly. As for other children however, Catharina's estimates did not become more linearly accurate at the same rate for all task conditions. There was a significant difference between the trend for T2 compared to T4, and even starker differences between the different number ranges (see Figure 152 for a visual representation of the different patterns observed).

This case study concludes the four chapters presenting the findings of this study. The findings at both the group and individual case levels have been presented and discussed in relation to the research questions and relevant literature. The next chapter will draw together the emerging conclusions contributed by the group findings and each of the case studies, to present the overall conclusions of the research.

Chapter 10 Summary and conclusions

10.1 Overview

This study set out to develop new depth of understanding of children's representations of number during their first year of formal schooling, through a longitudinal and multimodal research design not seen in previous research. The research examined children's cognitive representations of number, above all of number structure, as they engaged with key interview tasks taken from existing research. The study aimed to complement previous research by using a theoretical framework and research design led by Duval's inclusive theory of representation. This enabled the study to focus upon an existing mathematical phenomenon (representation), and to understand children's responses to different tasks, expressed using different modes of communication, all in terms of representations of children's developing conceptions of number. With this original design, the study was able to compare children's responses to tasks from both imagistic representation and cognitive science traditions, in contrast to previous research which investigated these areas separately. In the case of the number line estimation tasks taken from the cognitive science tradition, the theoretical framework and research design of this study offered an original perspective. As in previous studies, quantitative data on children's number line estimations was still collected, but so too was qualitative data on the representations children made whilst carrying out estimations: both data types concern children's cognitive representations of number. The study therefore brought new forms of evidence to bear on existing hypotheses about children's responses to number line estimation tasks.

This study sought to illuminate two particular aspects of children's representations of number. First was the connection between the aspects of number structure children represented in different tasks at different times. This was an original research question arising from previously separate research strands. Secondly and perhaps most importantly was the connection between the results of children's number line estimations, and the number structure they represented during the estimation. This was considered a particularly important original contribution of the present research, since previous studies had repeatedly hypothesised about children's estimation strategies without in fact examining children's task responses.

The previous chapters presented and discussed the research findings in two stages, firstly the group-level findings, and secondly the findings of three in-depth case studies. In each case, the findings were presented and discussed in relation to specific research questions and the relevant existing literature. This chapter draws together the overall conclusions of the study in relation to the research questions and research aims, showing the original contributions of the study and how they relate to previous research. This chapter also summarises the significance of the conclusions, methodological and theoretical reflections on the study, and the implications for educational practice and future research.

10.2 Conclusions

10.2.1 In what ways do children appear to cognitively represent number during the different interview tasks used in this research?

The findings of this study showed that children represented number in all the modes for which data was collected: speech/sound, gesture, gaze, and inscriptions. The children represented number structure from increasing sequence structure (considered the simplest structure) up to and including part-whole structure, including aspects of base-ten structure (considered the most mathematically sophisticated aspect of number structure examined).

The imagistic representations children produced in T1 were diverse. Just as found by Thomas et al. (2002), the responses to T1 included representations composed of pictorial components that represented no number structure, but association with quantity, such as Beatrice’s “fishes” response (Figure 22), but also representations composed of numerals in array form, with commentary to explain the internal structure, such as Patrick’s R3 response (Figure 49). This finding is firmly in line with Nunes and Bryant's (2009) theory of number concept development, which holds that children aged five are in the midst of a long process of making connections between counting knowledge, knowledge of quantities and knowledge of relations, a process that lasts well into the primary school years.

Representations during the number line estimation tasks of this study were dominated by those which connected the target number of the trial to features of the printed number line. A typical example was a point to the left endpoint followed by unit ‘hops’ of the pointing finger, forming a count-on procedure along the line. No previous research had investigated children's number line estimations in terms of representations. However, this finding is evidence in support of the diverse quantitative studies (Ashcraft and Moore, 2012; Barth and Paladino, 2011; White and Szucs, 2012) and eye-tracking studies (Heine et al., 2010; Schneider et al., 2008) hypothesising frequent use of endpoint anchors and count-on strategies, as captured at a broad level by the categorisation of strategies by Petitto (1990). Pictorial aspects were uncommon outside of T1, but children’s spontaneous representations did occasionally include them, for example Patrick’s relation of line lengths to the sky, or multiple children’s relation of target numbers to the ages of family members. The finding that responses to number line estimation tasks included such spontaneous representations is an original one not prefigured by existing research.

The findings of this study show that the main connection between number representations in different tasks was the level of complexity of number structure, according to Resnick's (1983) framework, that each child represented. The specific number representations of T1 were not frequently repeated during later tasks. However, children whose T1 representations had included little structure or solely increasing sequence structure also showed fewer representations of more complex number structure during T2 and T4. Conversely, children who represented more sophisticated number structure in their T1 representations also tended to represent more sophisticated number structure in later tasks. These broad groupings also coincided with children's teacher-assessed mathematical attainment groups. Connection between the structure of T1 representations and more advanced mathematical experience was documented by Thomas et al. (2002). However, since no previous research had investigated the representations made during number line estimations, the connection to these is necessarily an original contribution of this study.

In terms of differences between T2 and T4, both the group level results and case studies demonstrate that children tended to produce more representations of number structure in T2 than in T4. One hypothesis to explain this finding is that in T4, the blank rocket functioned as an 'answer box', leading children to focus (perhaps too quickly) on 'writing the answer' rather than on the process of conversion (in T4, from a spatial representation to a numeral or number word). This is in contrast to T2, where the conversion of representation required a physical *action* from children, and specifically an action (sticking the rocket sticker somewhere on the line) that was more unusual than 'writing a number in a blank space'.

The group-level findings showed that T4 estimations were overall significantly less linearly accurate than T2 estimations. This is compatible with the above interpretation of differences between children's behaviour during the two tasks, but it is not clear whether it is in line with the only previous research to have investigated number-to-position and position-to-number estimations in the same study, namely Siegler and Opfer (2003) . Siegler and Opfer (2003) do not report direct comparisons of number-to-position and position-to-number estimates; they report only the results of fitting linear, logarithmic and exponential models to the median estimates of each age group investigated, meaning that this task difference is nowhere directly discussed in the literature. Their results for 2nd graders (aged 7-8), the youngest participants in the study and thus closest in age to the participants of this study, showed no overall difference in linear model fit between the cohort's median estimates in position-to-number and number-to-position estimations. In addition to difference in action required by the two tasks, another factor very likely to have contributed to systematically lower PAE in T2 is that the task provides both numbered endpoints, and a numbered target to be placed in between, which places an upper bound on the PAE for each trial (the trial cannot possibly have greater than 100% PAE). In contrast, in T4, a participant is free to write any estimate, and so if for example only acknowledging the left endpoint as a bound, may in theory produce an estimate with any value PAE. If T4 estimation trials were carried out as the estimation of a target *within a range* as intended (which evidence suggests was the case for the majority of trials) then PAE is limited as in T2, but there were clear examples (see Marta's case) where the task as completed by the child was effectively open-ended.

A finding to emphasise is that although the frequency and proportion of representations of different categories of number structure varied between children, all categories of number structure were represented by all children during the study. The case studies of Patrick and Marta illustrate the two extremes of the group, and the observations of their strategies and representations across the interview rounds (e.g. Figure 75 and Figure 135) reveal that, for both children, the whole range of number structures were indeed represented. This means, firstly, that even with no specific direction from the researcher, the numerical-spatial features of number line estimation tasks allowed or encouraged children to represent multiple aspects of number structure during conversion of representations of number. This is an original contribution to the literature on number line estimation. Secondly, the finding emphasises that all children, including those less highly-attaining in classroom mathematics, incorporated these aspects of number structure into their conception of some of the study's target numbers.

10.2.2 What strategies can be identified in children's interactions with number line estimation tasks?

All the strategies hypothesised in previous research were documented in this study. These include count-on strategies, and use of the line endpoints and partitions of the range as 'anchor' points. Using extensive video data, this study was able to categorise children's strategies more precisely than earlier research (e.g. Petitto, 1990), resulting in the following list of observed strategies:

- Reference to potential anchor point
    - Those included in the task environment (left and right endpoints)
    - Those created or visualised by child
        - Midpoint
        - Quarter-point
        - Three-quarter point
    - Previous trials of the task
        - Within the same range
        - Within a different range

- Counting strategies
    - Count-on from LE to estimate
    - Count-on from estimate to anchor point (e.g. RE)
    - Count-back from RE
    - Count-back from midpoint
- Judgment using relative numerosities: particularly clear when estimate, or further strategies, appeared to be influenced the relative numerosity of target and task environment.

The three case studies contributed very strong evidence of connection between children's representations and the resulting estimation value or position. Investigation of this connection had not been attempted by previous research, so the findings relating to this research question form a set of original contributions. Each case study offers numerous examples where striking changes in linear estimation accuracy can be connected to specific changes in number structure represented: both for different targets within a range (e.g. target 8 estimated with high PAE, whereas target 9 estimated with very low PAE) and for the same target in different rounds (e.g. target 50 estimated with high PAE in R1 and R2, followed by low PAE in R3-R5).

Furthermore, findings demonstrate that the connection extended beyond individual target examples, in two ways. The first extension is where the distribution of estimates for a whole set of trials can be convincingly explained by the observed representations and strategies. A typical example of this is Catharina's T2 estimations in range 5-15 during R2 (Figure 173 - Figure 176). The second extension is in terms of a correlation between representation of more complex number structure and higher linear accuracy at the level of whole interviews. This was documented in the findings both within children's own interviews (for example Patrick, Marta and Catharina's changes over the five interview rounds) and between different children.

There was good evidence that children's representations and use of estimation strategies varied according to both the mathematical features of the trial (the number range, and the relation between the target number and the range given) and the child's own mathematical experience and confidence. This finding strongly supports White and Szucs' (2012, p. 9) proposal that "specific numbers could exhibit unique behaviours as a function of the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy", a suggestion not previously supported by evidence of children's task interactions.

In terms of variation according to mathematical features of trials, the findings repeatedly demonstrated representations and strategies oriented towards the nearest anchor point to a target. It is particularly important to note that this was true even in Marta's case, albeit to a lesser extent than for Patrick and Catharina. Charts displaying observed representations and strategies according to the nearest anchor point of the target (see Figure 154) showed variation in line with mathematical expectation.

This finding is in broad agreement with all previous studies hypothesising the use of anchor points (Ashcraft and Moore, 2012; Barth and Paladino, 2011; Ebersbach et al., 2008; Heine et al., 2010; Schneider et al., 2008; White and Szucs, 2012). However, the findings of this study differ from those of previous research in terms of the extent of anchor point use in children aged only 5-6 (Year 1). Though White and Szucs (2012) found good overall evidence for the use of anchor points, they did not find clear evidence for anchor-point use in their Year 1 participants. Similarly, Ashcraft and Moore (2012) hypothesised that their youngest participants (1st grade, ages 6-7) oriented their estimates using the left endpoint of the line only, with other anchor points only later introduced. It seems plausible to conclude that this difference in findings largely reflects a key methodological difference between the present study and these two quantitative examples. The present study examined children's representations of anchor points separately from the results of estimation trials. In the two cited studies, anchor point use was *inferred* from estimates with lower linear residuals and lower variance, a method which can only detect children's use of anchor points on the occasions when it leads to more linearly accurate estimates. The only previous study to collect qualitative data on children's strategies again did not find evidence of midpoint [the only anchor point beyond the left endpoint investigated] use in the youngest participants (1st grade, ages 6-7) (Petitto, 1990, p. 70). In this case, methodological differences again offer plausible reasons for substantial differences – the limitations of Petitto's study, explored in the literature review, mean that many representations captured in the present study could not be expected to have been recorded using Petitto's methods.

In terms of variation between children, children assessed by the teacher as lower-attaining in mathematics represented more complex number structure less frequently, and employed the estimation strategy of counting-on from the left endpoint more frequently. Once again, this corresponds to hypotheses in the literature, which suggest this as a 'basic' strategy (Ashcraft and Moore, 2012; Petitto, 1990; White and Szucs, 2012). Marta's case provides clear examples of trials in which, as Slusser et al. (2013) for example suggest, the number line task as completed by the child is effectively open-ended, since the right endpoint is either noted and dismissed, or not considered at all. The findings show that the count-on from left endpoint strategy did frequently correspond to lower estimation accuracy. However, there were also examples of linearly accurate results, and it is important not to categorise the strategy itself as fundamentally inaccurate.

The context of use of count strategies was important. Some children, for example Patrick and Zoe, represented aspects of number-space ‘fit’ that showed adaptation of unit size depending on the number range, a finding not identified or suggested by previous research with this task. For all children however, whether or not number-space ‘fit’ ideas were represented, count-on strategies were not equally successful in all number ranges. Children consistently achieved more linearly accurate representations in the range 0-20, including when count-on strategies were used, and this study concludes that this is likely connected to children considering the required unit size for a linear representation of 0-20 on an A4 page to be a ‘natural’ unit size. As the findings discuss, since children expressed both the idea that there was *too much space* to represent 0-10 and *not enough space* to represent 0-100, it seems possible that the classroom resources in the children's classroom, which overwhelmingly used the length of a piece of A4 paper to represent ranges 0-20 or 0-30, (ranges 0-10 and 0-100 were represented in other formats within the classroom) had accustomed children to perceiving this scale as ‘natural’ and to using it as their default scale. The finding that children in this study estimated more accurately in the range 0-20 than in ranges 0-10 or 0-100 is contrary to expectations based on previous literature (e.g. Berteletti et al., 2010; Slusser et al., 2013).

This study found that children were more attached to count-on strategies than their knowledge of number structure could explain. In all three case studies, but particularly Marta’s, there was evidence of children employing count-on from the left endpoint even after an initial response based on other aspects of number structure. A very typical example would be an initial response where the child moved the target rocket towards the right endpoint or midpoint based on the relative numerosity of target and endpoints. In most cases, the subsequent count strategy then led to a less linearly accurate response than the initial response. This finding was not anticipated from existing research, but seems to point towards children regarding the count-sequence structure as fundamentally reliable, which is plausible given the children’s educational experiences with number.

The study was not designed to investigate proportional-reasoning accounts of number line estimation (see Ashcraft and Moore, 2012; Barth and Paladino, 2011; Rouder and Geary, 2014), but examination of linear estimation error by individual target allowed many of the predictions of these accounts to be examined. The findings show that only estimations in the range 0-100 followed the patterns of over- and under-estimation (following S-shaped curves) predicted by proportional reasoning models (see Patrick, Figure 70, Marta, Figure 131, and Catharina, Figure

180). In addition, estimation patterns on lower ranges frequently followed the *opposite* pattern to that predicted (see Catharina, Figure 162 and Figure 168).

Overall, the patterns of over- and under-estimation error observed were more successfully explained by appeal to children's representations of number structure and estimation strategies than by proportional reasoning models in the literature. The comparison of children's representations of number structure with their number line estimations is an original contribution of this study, and this finding was not anticipated by previous studies in the literature. The explanation of observed estimation errors takes the following form: firstly, children tended to use strategies oriented towards the left endpoint for lower target numbers, and strategies oriented towards the right endpoint for higher target numbers. Secondly, the study found that children were most accurate at estimating in the range 0-20; that the implicit unit size of their strategies in the range 0-10 was too small; and that the implicit unit size of their strategies in the range 0-100 was too large. The combination of these findings explains over- followed by under-estimation for T2 range 0-100, and under- followed by over-estimation for T4 range 0-100 (the two findings which *are* compatible with a proportional reasoning account) but *also* explains under-followed by over-estimation for T2 range 0-10, and over-followed by under-estimation for T2 range 0-10 (two patterns directly *opposite* to that predicted by a proportional reasoning account).

10.2.3 How do young children's cognitive representations of number change during their first year of formal schooling?

Children represented increasingly sophisticated aspects of number structure with increasing frequency, as predicted by accounts of number concept development (Nunes and Bryant, 2009; Resnick, 1983). In addition, they represented the more basic aspects of number structure with *decreasing* frequency. As previously discussed, this second finding does not follow necessarily from the first; children could represent increasingly sophisticated number structure whilst retaining the representation of less sophisticated structure, but they did not. The hypothesis of this study is that as children gained confidence with more sophisticated number structure, they no longer found it efficient to represent number sequence as frequently as at the start of the year. It is important to remember that in addition to increasing general mathematical experience, the children in this research gained repeated experience with the interview tasks, allowing plenty of opportunities for them to develop their task responses over time.

Children's number line estimations more closely resembled linear representations over time, a finding directly in line with the literature in this field. The median estimates of the group became significantly more linear, proportionally with interview round (see group RQ1c). The individual case studies also document increases in linear estimation accuracy with interview round.

The findings show particularly strong evidence for the step-wise inclusion of anchor points as an important factor in increasing linear accuracy of estimation. This has been proposed multiple times (Ashcraft and Moore, 2012; Slusser et al., 2013) from quantitative analysis of estimation variation, and the present study now contributes original qualitative evidence from children's task interactions. The findings support the idea that the left endpoint of the line is always a site of high linear accuracy, and is then joined by accurate estimates close to the right endpoint, and then the midpoint. In the case of Patrick, there is also evidence for the subsequent inclusion of quarter and three-quarter points as anchor points.

The stage-by-stage inclusion of anchor points is found in overall conflicting accounts, with Ashcraft and Moore (2012) concluding that their data overall support the interpretation of number line estimation as reflecting the mental representation of number (with stage-by-stage incorporation of anchor points), whilst Slusser et al. (2013) conclude that their data support the view of number line estimation as proportional reasoning (with stage-by-stage incorporation of anchor points). In this second account, the stage-by-stage incorporation of anchor points is one of two factors leading to improved estimation accuracy: the inclusion of further anchor points changes the mathematical model from unbounded, to one-cycle, to two-cycle power function, whilst a parameter β present in each of these models (see Slusser et al., 2013) indexes an overall bias, which with age and experience approaches 1 (“perfect” linear estimation).

The present study was not designed to compare the above two accounts, but the findings do present ideas that pertain to both. Firstly, this study *was* designed to investigate the log-linear shift account, in which number-line estimation reflects children’s mental representation of number (Siegler et al., 2009), and concludes that the log-linear shift does not describe the present findings well. Secondly, the current findings do suggest that another factor in addition to anchor points is involved in the increasing linearity of children’s estimations. An original suggestion from these findings would be that the parameter β, indexing bias, could be involved in reflecting the degree to which children both appreciate a need to scale unit size, helping to make error relative to the anchor point lower, and are able to carry out the spatial representation that they intend (with both perceptual and motor error expected to decrease with age).

The findings document only weak support for the log-linear shift account. The group’s median estimates were best fit by models predicted by this account only for ranges 0-20 and 0-100, and even for these ranges, the change in estimations observed was not well-described by a log-linear shift in children’s estimations. The case studies again show only weak support for the theory. The only aspect clearly supported by the findings is that children’s estimations progressively resemble a linear distribution, and this, as previously noted, is not unique to the log-linear shift account.

The findings of this study support the central premise of Siegler's overlapping waves theory, which is that at any given age, children know and use "a variety of approaches (i.e., strategies, rules, or representations) that compete with one another for use, with each approach being more or less adaptive depending on the problem and situation" (Opfer and Siegler, 2007, p. 170; Siegler, 1996). However, the findings do not support the theory as applied to number line estimation in Siegler's own work (see Opfer and Siegler, 2007). This application of overlapping waves theory functions within the interpretation of number line estimation as revealing individuals' internal or mental representation of a number range. Within this interpretation, overlapping waves theory then describes the mechanism of change from a logarithmic to linear representation of each range considered. The findings of this study fail to support the idea that number line estimation reflects an internal or mental representation of number ranges in this transparent way. Instead, the findings strongly support the idea that individual estimation trials vary with trial-specific factors. In this understanding of number line estimation, the "variety of approaches (i.e. strategies, rules, or representations)" posited by overlapping waves theory are certainly present, but at the level of children's representations of individual numbers which they bring to bear on estimation trials with trial-specific factors. The multiple representations are not a logarithmic or linear distribution of numbers on the range 0-10, for example, but instead the multiple representations children have of target numbers, for example "nine is nine steps on from the left of the number line", "nine is one hop back from ten", and "nine is about halfway between 0 and 20".

10.3 Significance of findings

The study provides important qualitative evidence of children responding to number line estimation tasks in the ways hypothesised, but not actually investigated, by many previous studies in the field. The study provides a new form of evidence in support of conclusions from eye-tracking and novel statistical analyses about the strategies that children adopt in number-line estimation tasks, which hold that children incorporate progressively more ‘anchor points’ into their estimations and adapt their strategies based on not only their age and mathematical experience, but the mathematical features of the estimation trial.

The study provides a unique longitudinal data set, which has allowed the examination of individual children’s estimations – both qualitatively and quantitatively – not seen in previous studies. Close examination of number line estimations in this longitudinal data set does not reveal abrupt changes of representations of number ranges. Instead, the data shows how children's interactions with number line estimation tasks corroborate accounts of number concept development that emphasise connection-forming (Nunes and Bryant, 2009).

The study agrees with findings from the vast body of literature showing that children's estimations more closely resemble linear distribution with age and experience. The original research design of this study provided a newly in-depth view of the estimation processes that go along with this, and a close-grained look at how estimation trials themselves change. The evidence more strongly supports this description ("more closely resemble") than a log-linear hypothesis for children's number-line representations changing based on a 'shift' in their internal representation of number (e.g. Siegler and Booth, 2004). The significance of this study in this respect is that the depth of evidence on children's number line representations has afforded a more critical comparison of these two descriptions of increasing linearity. The conclusion of this study is that describing the change in terms of resembling or approximating a linear function (see Siegler and Thompson, 2014) does not merely avoid the incorrect assumption of the log-linear account (that the estimation results transparently reflect children's 'mental number line') but in fact has more (positive) scientific accuracy. The conclusion of this study is that multiple factors contribute to changes in children's estimations, and that increasingly linear estimations *result* from these changes together.

The research design provides a significantly new and original layer of evidence even within strategies already hypothesised or identified. Because of the multimodal data collected, the analysis not only, for example, identifies children using relative numerosity and a right endpoint anchor, but also observes how they do this. In an example from Zoe's R4 interview (see section 6.4.2.3) the study is able to show that Zoe left a (too large) space between 71 and 100 because of reckoning on the decades that needed to fit in between these two points. In previous studies in the literature, this level of detail was not possible: the qualitative data collected by Petitto (1990) comes closest, but was restricted to a much coarser level of detail (categorising children's responses to each trial into one of just four categories).

The study further challenges the idea that interpreting children's number line representations as components of their representation of a whole range of numbers is meaningful. White and Szucs (2012) concluded that their findings, suggesting trial-specific variation in estimation, shed reasonable doubt on this idea and highlighted it as a priority for investigation. The findings of this study clearly demonstrate the estimation of individual target numbers to be influenced by trial-specific factors, and thus support White and Szucs' position that it is not meaningful to think of individual estimation trials as outputs or results of an underlying 'representation of the range *n* to *m*'. This study is therefore in agreement with recent quantitative studies which have also directed their focus to individual estimation trials (e.g. Ashcraft and Moore, 2012) and found statistical evidence of trial-specific variation. As with other findings, the significance of this study remains that it is unique in offering the longitudinal and qualitative evidence on this point.

Although the findings in this study emerged from focus upon individual cases, there are good reasons to believe that they are not unique to these children studied. The sample was taken from a typical South of England primary school with mixed intake, and the sample of children included the full range of mathematical attainment in the class as assessed by the classroom teacher. For other year one pupils following the National Curriculum in England, it is believed that similar responses to the interview tasks, and similar changes throughout the school year, would be found. This study does not of course offer data on responses to number line estimation tasks in older children and adults, or on children of the same age (5-6 years) experiencing significantly different education.

10.4 Theoretical and methodological reflections

Duval's theory of representation provided a foundation for viewing children's responses to diverse research tasks as representations of their emerging conceptions of number. This was particularly valuable for number-line estimation tasks; this theoretical framework allowed the study to focus upon the mathematical process of converting between representations and to consider with an open mind all of children's representations during each trials. Instead of beginning with the categorisation of responses as correct/incorrect, or with the counting of pre-identified strategies, the theoretical framework led towards a methodology that captured a far more rounded view of children's mathematical activity.

Recording children's estimation processes over one school year provided a unique set of qualitative data that allowed changes in their representation of number structure and estimation strategies to be compared with changes in the linear accuracy of their estimations. The longitudinal aspect of the research design was crucial for the analysis and findings of this study.

The methodology, very strongly shaped by the demands of the theoretical framework, allowed the study to provide new forms of evidence on children's representations of number, and for this reason was a key part of the study's overall value and contribution. Whilst it was a highly time-consuming approach, it provided extremely interesting findings in response to the research questions, and I believe the application of the methodology to other samples – particularly older children at a different point in their mathematical education – would reveal similarly interesting aspects of their representation of number.

The theoretical framework and methodology were particularly valuable for allowing an in-depth exploration of children's representations of number in light of the highly complex 'unknowns' in this field. As the literature review explored in depth, the themes to which this research relates are problematic: definitions of number and concepts of number; and how the mind grasps, visualises or processes number, are complex and contentious topics. Furthermore, the contributions of different research traditions are frequently not reconciled.

The methodology does not allow the explanatory power of log-linear, proportional reasoning and strategy-based explanations of number line estimation data to be compared, but it was not designed to do so. It is possible that the current qualitative aspect of the methodology could be adapted into a categorisation scheme and its power to statistically explain results tested. The current study has provided groundwork that could enable this however, converting a qualitative analysis into a scheme that could successfully categorise children's behaviour would be difficult. Furthermore, the very numerous existing quantitative studies comparing (simpler) number-line estimation accounts have shown strictly limited success over the years.

10.5 Implications for practice

The conclusions of this study suggest several implications for educational practice. Firstly and most importantly, this study emphasised that young children represented emerging understanding of number structure and even measure. Evidence showed that children represented structure even when their resulting number line estimations were linearly inaccurate, and this was the case for all children in the study, not only the children considered 'high attaining'. Some children explicitly drew on knowledge of halves and quarters, but *all* children represented developing aspects of number structure that went beyond increasing and decreasing number sequences. The question implied for educational practice is whether these early and developing ideas about number structure are sufficiently appreciated and encouraged. Given the existing research on children's development of what is termed 'number sense', it is clear that connection *between* early number concepts is central to 'understanding' number, and this study emphasises that it would be false to delay mathematical discussion of relative numerosity and emerging ideas of proportion (in favour of the important early sequence structure) because children of this age group already represent such ideas in their interaction with a task that asked them to engage with spatial representation of numbers.

A second implication is that use of number line estimation tasks to investigate or assess children's emerging understanding of number would require extremely detailed (research-level) observation or mathematical dialogue with the child before the task could provide meaningful feedback. This study concludes that there are multiple ways for children to achieve linearly accurate representation on a number line, and multiple ways in which number line estimations can end up linearly inaccurate, and hence in order to connect the task to children's emerging number understanding requires significant effort.

A further implication is that for the same reason that number line estimation tasks are complex as tools of assessment, they are rich in terms of mathematical possibility, specifically in terms of emerging understanding of number structure. Particularly in light of the first point in this section – the representation of emerging number structure that the young children in this study showed – a question for educational practice is whether sufficient use of this simple task concept is made in classrooms for younger children. Even without direction from the researcher, the task context led or enabled children to represent number structure. In particular, the moments of triumph – using the 'shortcut' of relative numerosity to accurately position target 18 in the range 0-20 – and contradiction – particularly when count strategies contradicted relative numerosity and 'what was possible' – suggest obvious opportunities for mathematical discussion in the classroom.

A final implication of the study is to ask about the impact of two very different usages of the number line in current year one classrooms. The first is the frequent appearance of ordering tasks, where the linear spacing of numbers is already provided (either with blank boxes or numbers on equally sized pieces of card). This task reinforces sequence order, but no other number structure, and may indeed contribute to difficulties progressing beyond understanding numbers as discrete items in a sequence. The difficulties posed to later learning by the "whole number bias" (see for example Siegler et al., 2011) make this a serious concern. The current research relates to this strand of research by demonstrating the more structurally interesting mathematical ideas brought into play by the number line representation task, as well as points at which children encountered contradictions after reliance on sequence structure representation.

The second number line usage that the current study relates to is the use of the empty number line in addition and subtraction problems. In the light of the number structure and relationships brought to the foreground by the tasks in this study, the question is to what extent the empty number line method encourages an unhelpfully abstract sense of decades and units. The method encourages children to decompose numbers using part-whole structure, but then uses the number line to represent the parts in a way that divorces the spatial representation from the relative size of the components. To take a concrete example, although children are encouraged to use a 'big' jump to represent ten and a 'little' jump to represent one unit, a set of four wobbly 'jumps' along an empty number line could be used to represent 31 (3 tens + 1 unit), but also 13 (1 ten + 3 units), or indeed 40 (4 tens). Whilst acknowledging the excellent potential for encouraging children to decompose numbers using part-whole structure, the method as modelled in the classroom seems to have high potential for reducing meaning to 'hopping along the number line'. The findings of this study show that even in the range 0-10, children's sense of the relative numerosity of numbers is not yet robust. The present study has not examined the influence of empty number line usage on young children, but the conclusions of the study point strongly towards encouraging spatial representation of number in the classroom to respect relative numerosity.

10.6 Considerations for future research

Directions in which the present study could be strengthened or extended have been noted at various points. The study has confirmed strong hypotheses about ways in which young children respond to number line estimation tasks, but there are important areas which remain for further research.

One highly important area would be to apply the in-depth analysis of the present methodology to an adaption of the task-based interview that included an intervention by a teacher-researcher. Whilst previous research (Thompson and Opfer, 2010) has examined the impact of feedback (specifically, relating estimations to a 'correct' linear representation), the data collected focused only on children's quantitative estimation results. I believe that applying the mixed-methods analysis used in the current study would reveal the changes (if any) in children's representation of number at a far richer level. It would be particularly interesting to stop short of the 'correcting' feedback in Thompson and Opfer's study, and first examine the effect of designed questioning that engaged children in mathematical dialogue about where numbers 'belonged' on the number line representation. On one hand it is a strength of the current study that focused entirely on what children did, alone, with only the prompts of the task context and the very general support and encouragement of the researcher. On the other hand, it is likely that children's incorporation of number structure into the task could have increased given targeted encouragement, and research into this as an educational process is what this research did *not* do.

Related to the above point and to the well-documented connection between linear accuracy of estimation and mathematical attainment, an obvious route for future research would be to use the qualitative findings of the present study to design a teacher-researcher intervention for children, particularly those with currently low mathematical attainment. The aim of this would be to examine the short and long term effects of an intervention focused on linear representation *based* on the current findings about children's existing representation of number structure whilst carrying out estimation tasks. Parallel research into the existing task responses of older children, particularly after teaching on measure and proportionality, and then the potential of number line intervention, would also be illuminating.

References

Adelman, C., Kemmis, S., & Jenkins, D. (1980). Rethinking Case Study: Notes from the Second Cambridge Conference. In H. Simons (Ed.), *Towards a Science of the Singular: Essays About Case Study in Educational Research and Evaluation*. Norwich: Centre for Applied Research in Education, University of East Anglia.

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating Mental Representations through Speech and Gesture. *Psychological Science, 10*(4), 327-333.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal About a Child's State of Mind. *Cognitive Psychology, 25*, 468-523.

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive Processes of Numerical Estimation in Children. *Journal of Experimental Child Psychology, 111*(2), 246-267.

Aunio, P., & Niemivirta, M. (2010). Predicting Children's Mathematical Performance in Grade One by Early Numeracy. *Learning and Individual Differences, 20*(5), 427-435.

Badiou, A. (2008). *Number and Numbers* (R. Mackay, Trans.). Cambridge: Polity Press.

Barner, D., Thalwitz, D., Wood, J., Yang, S.-J., & Carey, S. (2007). On the Relation between the Acquisition of Singular–Plural Morpho-Syntax and the Conceptual Distinction between One and More Than One. *Developmental Science, 10*(3), 365-373.

Barth, H., Baron, A., Spelke, E., & Carey, S. (2009). Children's Multiplicative Transformations of Discrete and Continuous Quantities. *Journal of Experimental Child Psychology, 103*(4), 441-454.

Barth, H., & Paladino, A. (2011). The Development of Numerical Estimation: Evidence against a Representational Shift. *Developmental Science, 14*(1), 125-135.

Berch, D. (2005). Making Sense of Number Sense: Implications for Children with Mathematical Disabilities. *Journal of Learning Disabilities, 38*(4), 333-339.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical Estimation in Preschoolers. *Developmental Psychology, 46*(2), 545-551.

Booth, J. L., & Siegler, R. S. (2006). Developmental and Individual Differences in Pure Numerical Estimation. *Developmental Psychology, 41*, 189-201.

Booth, J. L., & Siegler, R. S. (2008). Numerical Magnitude Representations Influence Arithmetic Learning. *Child Development, 79*(4), 1016-1031.

Box, K., & Scott, P. (2004). Early Concepts of Number and Counting. *Australian Mathematics Teacher, 60*(4), 2-6.

Cantlon, J. F., Cordes, S., Libertus, M. E., & Brannon, E. M. (2009). Comment on "Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures". *Science, 323*(2).

Cantlon, J. F., Safford, K. E., & Brannon, E. M. (2010). Spontaneous Analog Number Representations in 3-Year-Old Children. *Developmental Science, 13*(2), 289-297.

Carey, S. (2004). Bootstrapping and the Origin of Concepts. *Daedalus, 133*(1), 59-69.

Carey, S., & Sarnecka, B. W. (2006). The Development of Human Conceptual Representations. In M. H. Johnson & Y. Munakata (Eds.), *Processes of Change in Brain and Cognitive Development: Attention and Performance Xxi* (pp. 473-496): Oxford University Press.

Case, R., & Okamoto, Y. (1996). The Role of Central Conceptual Structures in the Development of Children's Thought. *Monographs of the Society for Research in Child Development, 61*(1-2).

Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language, 35*, 2658.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*: The MIT press.

Chu, M., & Kita, S. (2011). The Nature of Gestures' Beneficial Role in Spatial Problem Solving. *Journal of Experimental Psychology, 140*(1), 102-116.

Clark, A., & Chambers, D. (1998). The Extended Mind. *Analysis, 58*, 7-19.

Clements, D., & Sarama, J. (2009). *Learning and Teaching Early Math: The Learning Trajectories Approach*. New York: Routledge.
Cohen, D. J. (2009). Integers Do Not Automatically Activate Their Quantity Representation. *Psychonomic Bulletin & Review, 16*(2), 332-336.
Cohen Kadosh, R., & Walsh, V. (2009). Numerical Representation in the Parietal Lobes: Abstract or Not Abstract? *Behavioral and Brain Sciences, 32*(3-4), 313.
Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (7th ed.). Abingdon: Routledge.
Condry, K. F., & Spelke, E. S. (2008). The Development of Language and Abstract Concepts: The Case of Natural Number. *Journal of Experimental Psychology: General, 137*(1), 22-38.
De Cruz, H. (2012). How Do Spatial Representations Enhance Cognitive Numerical Processing? *Cognitive Processing, 13*(1 Supplement), 137-140.
De Smedt, B., & Gilmore, C. K. (2011). Defective Number Module or Impaired Access? Numerical Magnitude Processing in First Graders with Mathematical Difficulties. *Journal of Experimental Child Psychology, 108*(2), 278-292.
De Smedt, B., & Verschaffel, L. (2009). *Cognitive Neuroscience Meets Mathematics Education.* Paper presented at the Cognitive Neuroscience Meets Mathematics Education, Brugge.
De Smedt, B., Verschaffel, L., & Ghesquiere, P. (2009). The Predictive Value of Numerical Magnitude Comparison for Individual Differences in Mathematics Achievement. *Journal of Experimental Child Psychology, 103*(4), 469-479.
Dehaene, S. (1997). *The Number Sense*. New York: Oxford University Press.
Dehaene, S. (2001). Précis of the Number Sense. *Mind & Language, 16*(1), 16-36.
Dehaene, S. (2003). The Neural Basis of the Weber–Fechner Law: A Logarithmic Mental Number Line. *Trends in Cognitive Sciences, 7*(4), 145-147.
Dehaene, S., & Akhavein, R. (1995). Attention, Automaticity and Levels of Representation in Number Processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 314-326.
Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General, 122*(3), 371-396.
Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract Representations of Numbers in the Animal and Human Brain. *Trends in Neuroscience, 21*(8), 355-361.
Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science, 320*(1217).
DeWindt-King, A., & Goldin, G. A. (2003). Children's Visual Imagery: Aspects of Cognitive Representation in Solving Problems with Fractions. *Mediterranean Journal for Research in Mathematics Education, 2*(1), 1-42.
Duval, R. (1995). *Sémiosis Et Pensée Humaine*. Bern: Peter Lang.
Duval, R. (1999). *Representation, Vision and Visualization: Cognitive Functions in Mathematical Thinking. Basic Issues for Learning.* Paper presented at the 21st Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Cuernavaca, Morelos, México.
Duval, R. (2000). *Basic Issues for Research in Mathematics Education.* Paper presented at the 24th Conference of the International Group for the Psychology of Mathematics Education (PME-24), Hiroshima, Japan.
Duval, R. (2002). Representation, Vision and Visualization: Cognitive Functions in Mathematical Thinking. Basic Issues for Learning. In F. Hitt (Ed.), *Representations and Mathematics Visualization* (pp. 311-336): Departamento de Matematica Educativa.
Duval, R. (2006). A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics. *Educational Studies in Mathematics, 61*(1/2), 103-131.
Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The Relationship between the Shape of the Mental Number Line and Familiarity with Numbers in 5- to 9-

Year Old Children: Evidence for a Segmented Linear Model. *Journal of Experimental Child Psychology, 99*(1), 1-17.
Ernest, P. (1996). Varieties of Constructivism: A Framework for Comparison. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin & B. Greer (Eds.), *Theories of Mathematical Learning* (pp. 335-350). Mahwah, NJ: Lawrence Erlbaum Associates.
Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of Different Types of Numerical Magnitude Representations to Each Other and to Mathematics Achievement. *Journal of Experimental Child Psychology, 123*, 53-72.
Feigenson, L., & Carey, S. (2003). Tracking Individuals Via Object-Files: Evidence from Infants' Manual Search. *Developmental Science, 6*(5), 568-584.
Feigenson, L., & Carey, S. (2005). On the Limits of Infants' Quantification of Small Object Arrays. *Cognition, 97*(3), 295-313.
Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The Importance of Magnitude Information in Numerical Processing: Evidence from the Snarc Effect. *Mathematical Cognition, 2*(1), 95-110.
Fias, W., & Fischer, M. H. (2005). Spatial Representation of Number. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 43-54). New York: Psychology Press.
Fuson, K., & Hall, J. W. (1983). The Acquisition of Early Number Word Meanings: A Conceptual Analysis and Review. In H. P. Ginsburg (Ed.), *The Development of Mathematical Thinking* (pp. 50-109). New York: Academic Press.
Fuson, K. C., Richards, J., & Briars, D. J. (1982). The Acquisition and Elaboration of the Number Word Sequence *Children's Logical and Mathematical Cognition* (pp. 33-92): Springer.
Gallistel, C. R., & Gelman, R. (1992). Preverbal and Verbal Counting and Computation. *Cognition, 44*(1-2), 43-74.
Galton, F. (1880). Visualised Numerals. *Nature, 21*, 252-256 and 494-495.
Garber, P., Alibali, M. W., & Goldin-Meadow, S. (1998). Knowledge Conveyed in Gesture Is Not Tied to the Hands. *Child Development, 69*, 75-84.
Garber, P., & Goldin-Meadow, S. (2002). Gesture Offers Insight into Problem-Solving in Adults and Children. *Cognitive Science, 26*(6), 817-831.
Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of Number Line Representations in Children with Mathematical Learning Disability. *Developmental Neuropsychology, 33*(3), 277-299.
Gelman, R., & Butterworth, B. (2005). Number and Language: How Are They Related? *Trends in Cognitive Sciences, 9*(1), 6-10.
Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.
Gelman, R., & Gallistel, C. R. (2004). Language and the Origin of Numerical Concepts. *Science, 306*, 441.
Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early Identification and Interventions for Students with Mathematics Difficulties. *Journal of Learning Disabilities, 38*(4), 293-304.
Ginsburg, H., Kossan, N. E., Schwartz, R., & Swanson, D. (1983). Protocol Methods in Research on Mathematical Thinking. In H. P. Ginsburg (Ed.), *The Development of Mathematical Thinking* (pp. 7-47). New York: Academic Press.
Ginsburg, H. P. (1981). The Clinical Interview in Psychological Research on Mathematical Thinking: Aims, Rationales, Techniques. *For the Learning of Mathematics, 1*(3), 4-11.
Goldin, G. A. (1992). On the Developing of a Unified Model for the Psychology of Mathematics Learning and Problem Solving. In W. Geeslin & K. Graham (Eds.), *Proceedings of the 16th Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 235-261).
Goldin, G. A. (2002). Representation in Mathematical Learning and Problem Solving. In L. D. English (Ed.), *Handbook of International Research in Mathematics Education* (pp. 197). New York: Psychology Press.

Goldin, G. A. (2008). Perspectives on Representation in Mathematical Learning and Problem Solving. In L. D. English (Ed.), *Handbook of International Research in Mathematics Education, Second Edition* (2nd ed., pp. 176-201). New York: Routledge.
Goldin, G. A., & Kaput, J. J. (1996). A Joint Perspective on the Idea of Representation in Learning and Doing Mathematics. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin & B. Greer (Eds.), *Theories of Mathematical Learning* (pp. 397-430). Mahwah, NJ: Lawrence Erlbaum Associates.
Greeno, J. G. (1991). Number Sense as Situated Knowing in a Conceptual Domain. *Journal for Research in Mathematics Education, 22*(3), 170-218.
Guttenplan, S. (2005). Mental Representation. In S. Guttenplan (Ed.), A Companion to the Philosophy of Mind (pp. 441). Malden, MA: Blackwell Publishing.
Halberda, J., & Feigenson, L. (2008). Developmental Change in the Acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-Year-Olds and Adults. *Developmental Psychology, 44*(5), 1457-1465.
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual Differences in Non-Verbal Number Acuity Correlate with Maths Achievement. *Nature, 455*, 665-669.
Heine, A., Thaler, V., Tamm, S., Hawelka, S., Schneider, M., Torbeyns, J., . . . Jacobsa, A. M. (2010). What the Eyes Already 'Know': Using Eye Movement Measurement to Tap into Children's Implicit Numerical Magnitude Representations. *Infant and Child Development, 19*, 175-186.
Hitt, F. (2002). *Representations and Mathematics Visualization*: Departamento de Matematica Educativa.
Ho, C. S.-H., & Fuson, K. C. (1998). Children's Knowledge of Teen Quantities as Tens and Ones; Comparisons of Chinese, British, and American Kindergartners. *Journal of Educational Psychology, 90*(3), 536-544.
Hollands, J. G., & Dyre, B. P. (2000). Bias in Proportion Judgments: The Cyclical Power Model. *Psychological Review, 107*(3), 500-524.
Holloway, I. D., & Ansari, D. (2009). Mapping Numerical Magnitudes onto Symbols: The Numerical Distance Effect and Individual Differences in Children's Mathematics Achievement. *Journal of Experimental Child Psychology, 103*(1), 17-29.
Howden, H. (1989). Teaching Number Sense. *Arithmetic Teacher, 36*(6), 6-11.
Howell, S. C., & Kemp, C. R. (2010). Assessing Preschool Number Sense: Skills Demonstrated by Children Prior to School Entry. *Educational Psychology, 30*(4), 411-429.
Hubbard, E. M., Ranzini, M., Piazza, M., & Dehaene, S. (2009). What Information Is Critical to Elicit Interference in Number-Form Synaesthesia? *Cortex, 45*(10), 1200-1216.
Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). *Non-Verbal Number Acuity Correlates with Symbolic Mathematics Achievement: But Only in Children*.
Izard, V., & Dehaene, S. (2008). Calibrating the Mental Number Line. *Cognition, 106*(3), 1221-1247.
Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number Sense Growth in Children: A Longitudinal Investigation of Children at Risk for Mathematics Difficulties. *Child Development, 77*, 153-175.
Kaminski, E. (1997). Teacher Education Students' Number Sense: Initial Expectations. *Mathematics Education Research Journal, 9*(2), 225-235.
Kaput, J. J. (1987). Representation Systems and Mathematics. In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics* (pp. 19-26). Hillsdale, NJ: Lawrence Erlbaum Associates.
Kendon, A. (1988). How Gestures Can Become Like Words. In F. Poyatos (Ed.), *Cross-Cultural Perspectives in Nonverbal Communication* (pp. 131-141). Toronto: Hogrefe.
Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.
Krajcsi, A., & Palatinus, Z. (2004). Subitising: Analogue Magnitude System or Object File System? *Perception 33: ECVP Abstract Supplement*.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., . . . von Aster, M. (2011). Mental Number Line Training in Children with Developmental Dyscalculia. *Neuroimage, 57*(3), 782-795.

Kucian, K., & Kaufmann, L. (2009). A Developmental Model of Number Representation. *Behavioral and Brain Sciences, 32*, 340-342.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental Dyscalculia and Basic Numerical Capacities: A Study of 8--9-Year-Old Students. *Cognition, 93*(2), 99-125.

Larkin, J. H., & Simon, H. A. (1987). Why a Diagram Is (Sometimes) Worth Ten Thousand Words. *Cognitive Science, 11*, 65-99.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a Big Number? Correlational and Causal Connections among Numerical Categorization, Number Line Estimation, and Numerical Magnitude Comparison. *Child Development, 76*, 1723-1743.

Le Corre, M., & Carey, S. (2007). One, Two, Three, Four, Nothing More: An Investigation of the Conceptual Sources of the Verbal Counting Principles. *Cognition, 105*(2), 395-438.

Le Corre, M., & Carey, S. (2008). Why the Verbal Counting Principles Are Constructed out of Representations of Small Sets of Individuals: A Reply to Gallistel. *Cognition, 107*(2), 650-662.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool Acuity of the Approximate Number System Correlates with School Math Ability. *Developmental Science, 14*(6), 1292-1300.

Longo, G., & Viarouge, A. (2010). Mathematical Intuition and the Cognitive Roots of Mathematical Concepts. *Topoi, 29*(1), 15-27.

Lovell, K. (1965). *The Growth of Basic Mathematical and Scientific Concepts in Children*. London: University of London Press.

Marcou, A., & Gagatsis, A. (2003). A Theoretical Taxonomy of External Systems of Representation in the Learning and Understanding of Mathematics. In A. Gagatsis & I. Elia (Eds.), *Representations and Geometrical Models in the Learning of Mathematics* (Vol. 1, pp. 171-178). Nicosia: Intercollege Press (in Greek).

Mason, J. (1992). Doing and Construing Mathematics in Screen-Space. In B. Southwell, B. Perry & K. Owens (Eds.), *Proceedings of the Fifteenth Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 1-17). Nepean, Australia: University of Western Sydney.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago: The University of Chicago Press.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative Development in Infancy and Early Childhood*. New York: Oxford University Press.

Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2009). Children's Early Mental Number Line: Logarithmic or Decomposed Linear? *Journal of Experimental Child Psychology, 103*(4), 503-515.

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early Place-Value Understanding as a Precursor for Later Arithmetic Performance--a Longitudinal Study on Numerical Development. *Research in Developmental Disabilities: A Multidisciplinary Journal, 32*(5), 1837-1851.

Mulligan, J., Mitchelmore, M., & Prescott, A. (2006). Integrating Concepts and Processes in Early Mathematics: The Australian Pattern and Structure Mathematics Awareness Project (Pasmap). In J. Novotná, Moraová, H., Krátká, M. & Stehlíková, N. (Ed.), Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education (Vol. 4, pp. 209-216). Prague: PME.

Mundy, E., & Gilmore, C. K. (2009). Children's Mapping between Symbolic and Nonsymbolic Representations of Number. *Journal of Experimental Child Psychology, 103*(4), 490-502.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Newman, R. S., & Berger, C. F. (1984). Children's Numerical Estimation: Flexibility in the Use of Counting. *Journal of Educational Psychology, 76*(1), 55-64.

Nuerk, H.-C., Kaufmann, L., Zoppoth, S., & Willmes, K. (2004). On the Development of the Mental Number Line: More, Less, or Never Holistic with Increasing Age? *Developmental Psychology, 40*(6), 1199-1211.

Nunes, T., & Bryant, P. (2009). *Paper 2: Understanding Whole Numbers*. London: Nuffield Foundation.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How Training on Exact or Approximate Mental Representations of Number Can Enhance First-Grade Students' Basic Number Processing and Arithmetic Skills. *Learning and Instruction, 23*(0), 125-135.

Opfer, J., & Siegler, R. S. (2007). Representational Change and Children's Numerical Estimation. *Cognitive Psychology, 55*, 169-195.

Opfer, J. E., & DeVries, J. M. (2008). Representational Change and Magnitude Estimation: Why Young Children Can Make More Accurate Salary Comparisons Than Adults. *Cognition, 108*(3), 843-849.

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The Powers of Noise-Fitting: Reply to Barth and Paladino. *Developmental Science, 14*(5), 1194-1204.

Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early Development of Spatial-Numeric Associations: Evidence from Spatial and Quantitative Performance of Preschoolers. *Developmental Science, 13*(5), 761-771.

Oxford English Dictionary online. (2012). "Conception, N.", from http://www.oed.com/view/Entry/38137?redirectedFrom=conception&

Patro, K., Nuerk, H.-C., Cress, U., & Haman, M. (2014). How Number-Space Relationships Are Assessed before Formal Schooling: A Taxonomy Proposal. [Mini Review]. *Frontiers in Psychology, 5*.

Perry, M., Breckinridge Church, R., & Goldin-Meadow, S. (1988). Transitional Knowledge in the Acquisition of Concepts. *Cognitive Development, 3*(4), 359-400.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1992). Is Gesture-Speech Mismatch a General Index of Transitional Knowledge? *Cognitive Development, 7*(1), 109-122.

Petitto, A. L. (1990). Development of Numberline and Measurement Concepts. *Cognition and Instruction, 7*(1), 55-78.

Piaget, J. (1969a). *The Child's Conception of Number*. London: Routledge & Kegan Paul.

Piaget, J. (1969b). *Science of Education and the Psychology of the Child*. New York: Viking Press.

Piaget, J., & Inhelder, B. (1967). *The Child's Conception of Space*. New York: Norton.

Piaget, J., & Inhelder, B. (1971). *Mental Imagery and the Child*. London: Routledge & Kegan Paul.

Presmeg, N. C. (1986). Visualisation in High School Mathematics. *For the Learning of Mathematics, 6*(3), 42-46.

Presmeg, N. C. (1992). Prototypes, Metaphors, Metonymies and Imaginative Rationality in High School Mathematics. *Educational Studies in Mathematics, 23*(6), 595-610.

Presmeg, N. C. (1998). Metaphoric and Metonymic Signification in Mathematics. *Journal of Mathematical Behavior, 17*(1), 25-32.

Presmeg, N. C. (2002). Preface. In F. Hitt (Ed.), *Representations and Mathematics Visualization* (pp. ix-xv): Departamento de Matematica Educativa.

Presmeg, N. C. (2006). Research on Visualization in Learning and Teaching Mathematics. In A. Gutirrez & P. Boero (Eds.), *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future* (pp. 205-235). Rotterdam: Sense.

Presmeg, N. C. (2008). An Overarching Theory for Research in Visualization in Mathematics Education *Plenary Presentation at Tsg 20 on Visualization in the Teaching and Learning of Mathematics at the 11th International Congress on Mathematical Education (Icme-11)*. Monterrey, Mexico, July 6 - 13, 2008.

Putnam, H. (1987). *The Many Faces of Realism*. LaSalle, IL: Open Court.

Ramani, G. B., & Siegler, R. S. (2008). Promoting Broad and Stable Improvements in Low-Income Children's Numerical Knowledge through Playing Number Board Games. *Child Development, 79*(2), 375-394.

Rasanen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-Assisted Intervention for Children with Low Numeracy Skills. *Cognitive Development, 24*(4), 450-472.

Resnick, L. B. (1983). A Developmental Theory of Number Understanding. In H. P. Ginsburg (Ed.), *The Development of Mathematical Thinking* (pp. 109-151). New York: Academic Press.

Rey, G. (2005). Concepts. In S. Guttenplan (Ed.), A Companion to the Philosophy of Mind (pp. 185-193). Malden, MA: Blackwell.

Rips, L. J., Asmuth, J., & Bloomfield, A. (2006). Giving the Boot to the Bootstrap: How Not to Learn the Natural Numbers. *Cognition, 101*(B51-B60).

Rips, L. J., Asmuth, J., & Bloomfield, A. (2008). Do Children Learn the Integers by Induction? *Cognition, 106*(2), 940-951.

Rips, L. J., Bloomfield, A., & Asmuth, J. (2008). From Numerical Concepts to Concepts of Number. *Behavioral and Brain Sciences, 31*(06), 623-642.

Rouder, J. N., & Geary, D. C. (2014). Children's Cognitive Representation of the Mathematical Number Line. *Developmental Science, 17*(4), 525-536.

Santens, S., & Gevers, W. (2008). The Snarc Effect Does Not Imply a Mental Number Line. *Cognition, 108*(1), 263-270.

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental Number Line, Number Line Estimation, and Mathematical Achievement: Their Interrelations in Grades 5 and 6. *Journal of Educational Psychology, 101*(2), 359-372.

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., . . . Stern, E. (2008). A Validation of Eye Movements as a Measure of Elementary School Children's Developing Number Sense. *Cognitive Development, 23*(3), 409-422.

Schneider, M., & Stern, E. (2010). The Developmental Relations between Conceptual and Procedural Knowledge: A Multimethod Approach. *Developmental Psychology, 46*(1), 178-192.

Schoenfeld, A. (2008). Research Methods in (Mathematics) Education. In L. D. English (Ed.), *Handbook of International Research in Mathematics Education, Second Edition*. New York: Routledge.

Schwartz, R. (2005). Representation. In S. Guttenplan (Ed.), A Companion to the Philosophy of Mind (pp. 536-541). Malden, MA: Blackwell Publishing.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. Oxford: Oxford University Press.

Siegler, R. S. (1997). Beyond Competence-toward Development. *Cognitive Development, 12*(3), 323-332.

Siegler, R. S., & Booth, J. L. (2004). Development of Numerical Estimation in Young Children. *Child Development, 75*(2), 428-444.

Siegler, R. S., & Opfer, J. (2003). The Development of Numerical Estimation: Evidence for Multiple Representations of Numerical Quantity. *Psychological Science, 14*, 237-243.

Siegler, R. S., & Robinson, M. (1982). The Development of Numerical Understandings. In H. W. Reese & L. P. Lipsett (Eds.), *Advances in Child Development and Behavior: Vol. 16* (pp. 242-312). New York: Academic Press.

Siegler, R. S., & Thompson, C. A. (2014). Numerical Landmarks Are Useful--except When They're Not. *Journal of Experimental Child Psychology, 120*, 39-58.

Siegler, R. S., Thompson, C. A., & Opfer, J. (2009). The Logarithmic-to-Linear Shift: One Learning Sequence, Many Tasks, Many Time Scales. *Mind, Brain, and Education, 2*, 143-150.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An Integrated Theory of Whole Number and Fractions Development. *Cognitive Psychology, 62*(4), 273–296.

Slaughter, V., Kamppi, D., & Paynter, J. (2006). Toddler Subtraction with Large Sets: Further Evidence for an Analog-Magnitude Representation of Number. *Developmental Science, 9*(1), 33-39.

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental Change in Numerical Estimation. *Journal of Experimental Psychology: General, 142*(1), 193-208.

Sophian, C. (1997). Beyond Competence: The Significance of Performance for Conceptual Development. *Cognitive Development, 12*, 281-303.

Sophian, C., & Madrid, S. (2003). *Goal Sketches in Fraction Learning*. Paper presented at the 27th International Group for the Psychology of Mathematics Education Conference Held Jointly with the 25th PME-NA Conference, Honolulu, Hawaii, July 13-18. http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED501127

Sophian, C., & Wood, A. (1996). Numbers, Thoughts, and Things: The Ontology of Numbers for Children and Adults. *Cognitive Development, 11*(3), 343-356.

Sullivan, J. L., & Barner, D. (2010). Mapping Number Words to Approximate Magnitudes: Associative Learning or Structure Mapping? In S. Ohlsson & R. Catrambone (Eds.), *Cognition in Flux: Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1246-1251). Austin, TX: Cognitive Science Society.

Sullivan, J. L., Juhasz, B. J., Slattery, T. J., & Barth, H. C. (2011). Adults' Number-Line Estimation Strategies: Evidence from Eye Movements. *Psychonomic Bulletin & Review, 18*(3), 557-563.

Telegraph. (2011). People Are 'Born Bad at Maths'. *The Telegraph*. Retrieved from

Thagard, P. (2011). Cognitive Science. *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)* Retrieved 21/09/2011, 2011, from <http://plato.stanford.edu/archives/fall2011/entries/cognitive-science/>

Thomas, N. (1992). An Analysis of Children's Understanding of Numeration. In B. Southwell, B. Perry & K. Owens (Eds.), Proceedings of the fifteenth annual conference of the mathematics education research group of Australasia (pp. 521-540). Nepean, Australia: University of Western Sydney.

Thomas, N., & Mulligan, J. (1995). Dynamic Imagery in Children's Representations of Number. *Mathematics Education Research Journal, 7*(1), 5-25.

Thomas, N., Mulligan, J., & Goldin, G. A. (1994). Children's Representations of the Counting Sequence 1-100: Study and Theoretical Interpretation. In d. P. J. P. & J. F. Matos (Eds.), *Proceedings of the 18th Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol. III, pp. 1-8). Lisbon, Portugal: International Group for the Psychology of Mathematics Education.

Thomas, N., Mulligan, J., & Goldin, G. A. (2002). Children's Representation and Structural Development of the Counting Sequence 1-100. *Journal of Mathematical Behaviour, 21*(1), 117-133.

Thompson, C. A., & Opfer, J. (2010). How 15 Hundred Is Like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition. *Child Development, 81*, 1768-1786.

Thompson, C. A., & Siegler, R. S. (2010). Linear Numerical-Magnitude Representations Aid Children's Memory for Numbers. *Psychological Science, 21*(9), 1274-1281.

Thompson, P. W. (1996). Imagery and the Development of Mathematical Reasoning. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin & B. Greer (Eds.), *Theories of Mathematical Learning* (pp. 267-283). Mahwah, NJ: Lawrence Erlbaum Associates.

Thompson, P. W. (1999). *Representation and Evolution: A Discussion of Duval's and Kaput's Papers*. Paper presented at the PME-NA XXI Plenary Session.

Trafton, P. (1992). Using Number Sense to Develop Mental Computation and Computational Estimation *Challenging Children to Think When They Compute: Proceedings of a Conference Sponsored by the Centre for Mathematics and Science Education*. Brisbane, Australia: Queensland University of Technology.

Van de Walle, J., & Watkins, K. B. (1993). Early Development of Number Sense. In R. J. Jensen (Ed.), *Research Ideas for the Classroom* (pp. 127-151). New York: Macmillan.

Vergnaud, G. (1987). Conclusion. In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics* (pp. 227-232). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vinner, S., & Dreyfus, T. (1989). Images and Definitions for the Concept of Function. *Journal for Research in Mathematics Education, 20*(4), 356-366.

von Glaserfeld, E. (1987). Preliminaries to Any Theory of Representation. In C. Janvier (Ed.), *Problems of Representation in the Teaching and Learning of Mathematics* (pp. 215-226). Hillsdale, NJ: Lawrence Erlbaum Associates.

White, S. L. J., & Szucs, D. (2012). Representational Change and Strategy Use in Children's Number Line Estimation During the First Years of Primary School. *Behavioral and Brain Functions, 8*(1), 1.

Williamson, J. (2013). Young Children's Cognitive Representations of Number and Their Number Line Estimations. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 401-408). Kiel, Germany: IPN–Leibniz Institute for Science and Mathematics Education at the University of Kiel.

Wynn, K. (1992). Evidence against Empiricist Accounts of the Origins of Numerical Knowledge. *Mind and Language, 7*, 315-332.

Yin, R. K. (1994). *Case Study Research: Design and Methods* (2nd ed.). Thousand Oaks, CA: Sage.

Young, C. J., & Opfer, J. E. (2011). Psychophysics of Numerical Representation. *Zeitschrift für Psychologie / Journal of Psychology, 219*(1), 58-63.

Appendix 1

The following table has been my working document for keeping track of theoretical ideas in the thesis.

	Theoretical idea	Says what	Impact/role	Thoughts
Background	Cognitive integration (De Cruz, 2012)	"internal and external parts of numerical cognition are complementary, and need not be isomorphic" Expect "a causal, dynamic interaction between both types of processes"	One of several philosophical theories attempting to explain "how actions in the world enhance cognitive processing"	Have to contrast with the *internalist* views (Dehaene, de Cruz suggests).
Framework	Duval's theory of cognitive representation	Complex. Discussed at length in Theoretical F/W chapter.	Research problem, RQs, lit review, design and analysis.	Need to make clear the theory influences the project at all these levels.
	Presmeg's theory of visualisation	5 types of visual imagery. As adapted, 3 types of component sign.	RQs, design, analysis.	A way to interpret, classify and compare imagistic rep^{n}s.
Being tested/explored in analysis	No name – hypothesis from results (White and Szucs, 2012)	"specific numbers could exhibit unique behaviors as a function of the familiarity with the number range, proximity to either external or mental anchor points, as well as knowledge of arithmetic strategy"	A way to interpret number line estimation task results	Try interpreting using this: does it seem plausible and constructive?
	Proportion judgement	Hypothesis about how estimation task done	A way to interpret number line estimations	To be tested/considered
	Log-linear hypothesis	See lit review.	A way to interpret number line estimations	To be tested
	Overlapping waves theory (Siegler and Opfer, 2003)	Individuals have multiple rep^{n}s at a time, and their use comes and goes in waves	A way to interpret changes in uses of representations.	Try interpreting using this: does it seem plausible and constructive?

Appendix 2

The following is the diagram of Duval's classification of cognitive representations as it appears in (Duval, 1999). It is an earlier version of the same classification diagram cited in the Theoretical Framework, from Duval (2000). In substance, the two diagrams are very similar, but certain changes in Duval's labelling and examples mean that the 2000 diagram clarifies some earlier ambiguities.

Diagram of cognitive representation classification (Duval, 1999, p. 6; republished Duval, 2002, p. 315)

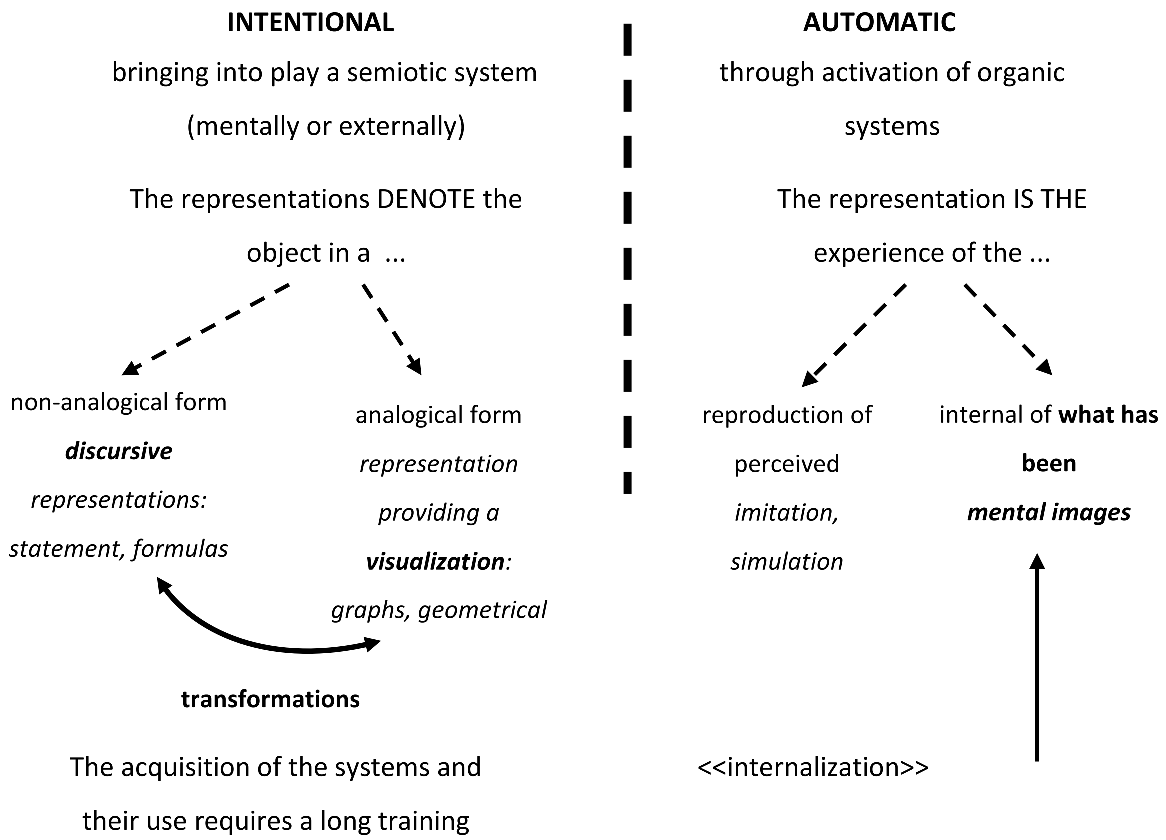


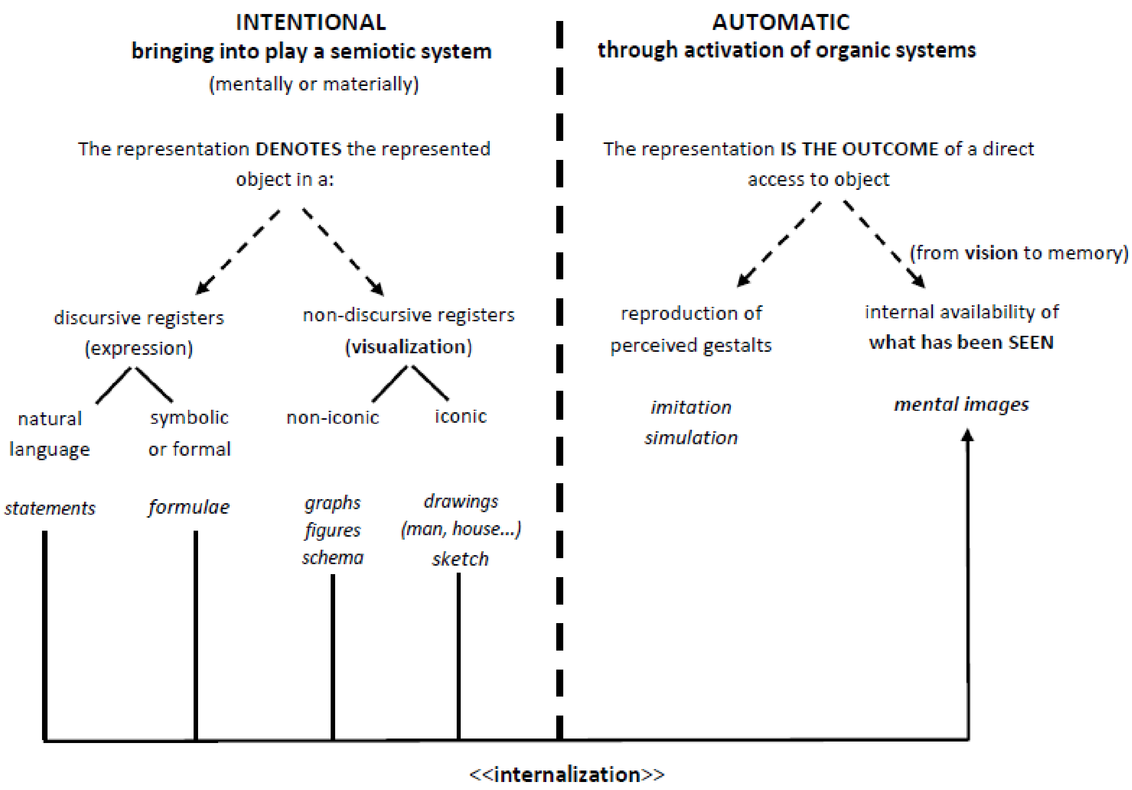
Figure 1. "Cognitive classification of conscious representations. This classification can be expanded more and includes all kinds of representations."

The key differences between the 1999 and 2000 diagrams are the following:

- In the 2000 diagram, Duval no longer labels the two classes of intentional representations (left hand side) "non-analogical" and "analogical". Instead, representations are divided into "discursive" and "non-discursive (visualization)", then further subdivided into "natural language" and "formulae" (discursive) and "non-iconic" and "iconic" (visualization). Duval provided examples for each category.

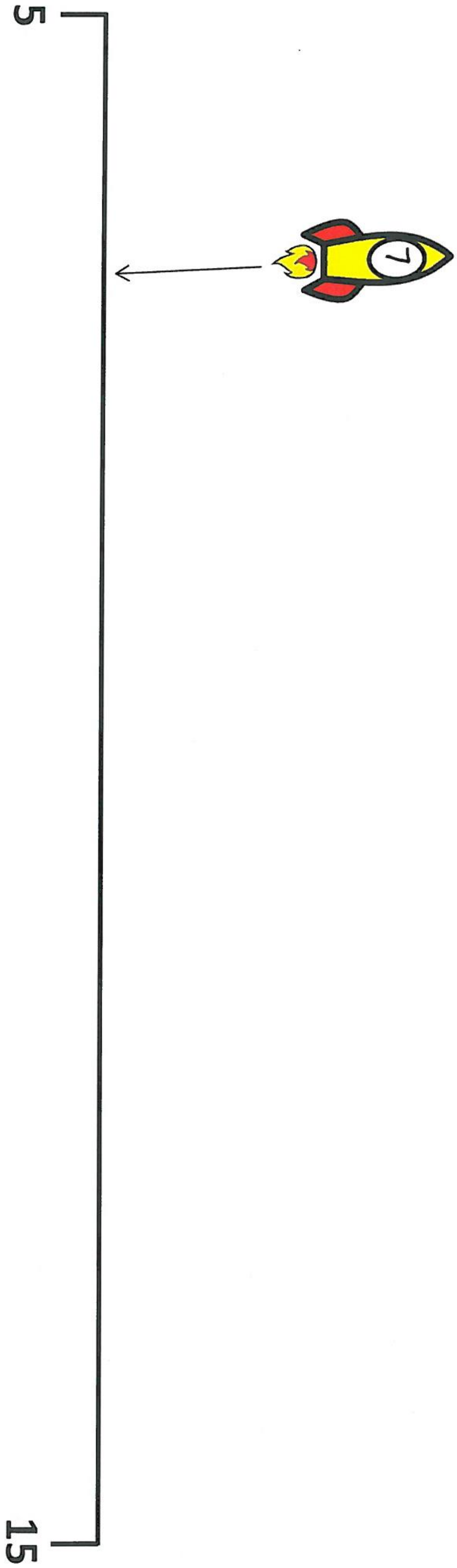
- The 1999 diagram caption explains that the diagram classifies "conscious representations", but this is not defined anywhere in the paper.
- The classification can be extended to "all kinds of representations" in the 1999 diagram: unclear what this means.
- The 2000 diagram explicitly includes among the intentional representations (left hand side) representations that do not belong to any particular formal semiotic system: "drawings (man, house)" and "sketch".
- The 2000 diagram omits the comment that "The acquisition of the systems and their use requires a long training" (Duval, 1999, p. 6). Elsewhere, Duval explains how formal semiotic systems do require extensive training, but by omitting the comment from the diagram confusion is avoided over the nature of training required or possible for non-formal representations such as iconic drawings and sketches.

<u>2000 diagram, for comparison (Duval, 2000)</u>



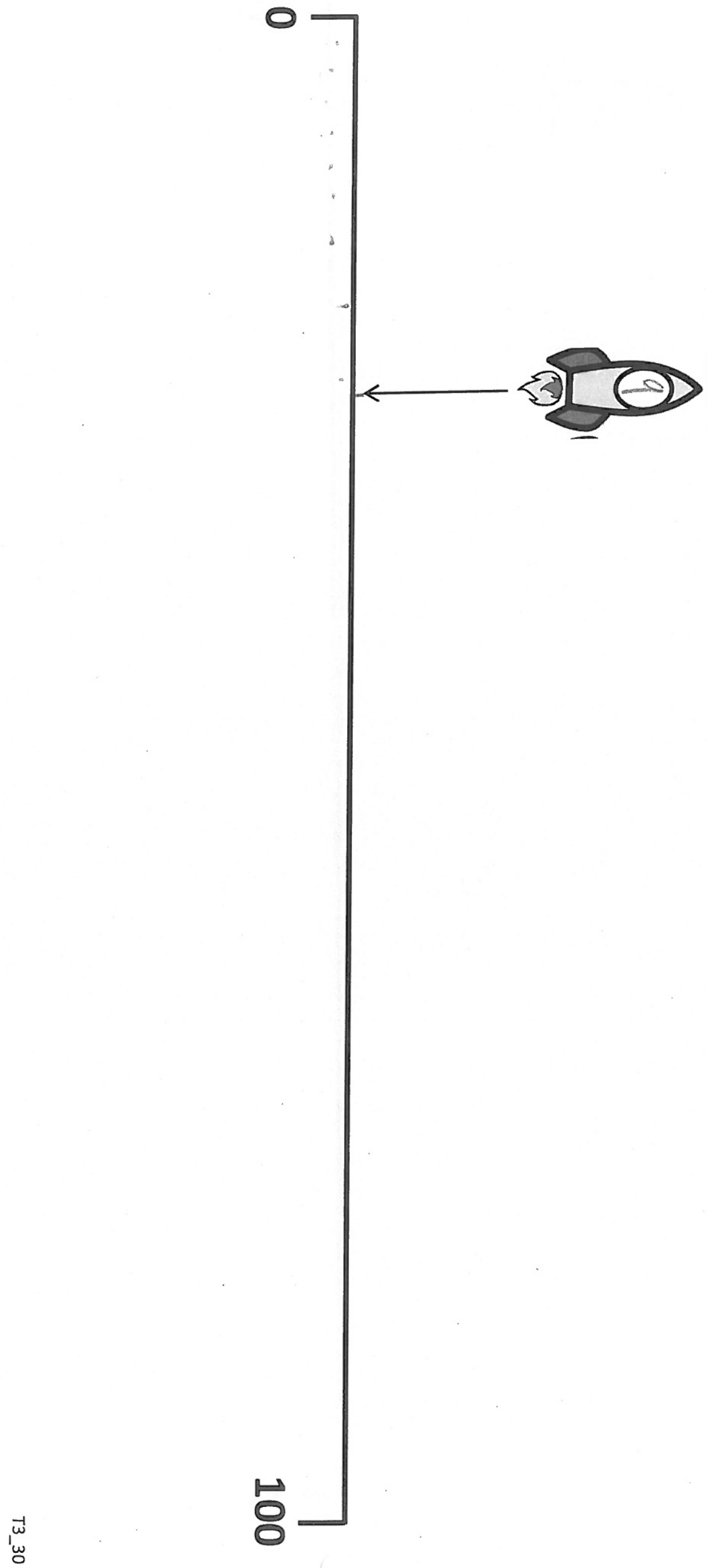
Appendix 3

Example of Task 2: number-to-position rocket estimation (final research design).



Appendix 4

Example of Task 4: position-to-number rocket estimation (final research design).



Appendix 5

Example of Task 2: number-to-position rocket estimation (as used in pilot study).

Appendix 6

In the following list of inferred strategies and task responses, an example of typical behaviour is provided for each item. As in the rest of the transcripts, the majority of these episodes show more than one strategy in operation. The examples have been chosen in order to demonstrate typical behaviour, and in typical cases strategies were combined, for example reference count-back from right endpoint was almost always combined with evidence for use of a right endpoint anchor.

The following abbreviations are used:

R**E:** right endpoint

L**E:** left endpoint

M**P:** midpoint

RH: right hand

LH: left hand

RHI: right hand index finger

LHI: left hand index finger

J: the researcher

P, A, etc. the participating children

Strategies

Reference to left endpoint anchor

Child 2, R3. T4, range 0-100, target to estimate is 2.

Researcher, J: What about that one over there?
A: [*RE, LE. RE, LE.*] Easy! One. [*Measures LE to rocket with one small jump of LH pencil point. Writes.*]
J: Oh, very good

Reference to right endpoint anchor

Child 1, R1. T2, range 0-20, target to estimate is 19.

P: I think ... it goes there ... [*Sticking rocket sticker on line next to RE*]
J: Yes?
P: Because it's really close to the twenty. [*RH traces rocket->RE, LHI points at rocket.*]

Reference to midpoint anchor

Child 11, R4. T4, range 0-10, target to be estimated is 4. Chris estimates the answer to be 5.

C: [*looks at rocket, then MP*] It's nearly in the middle. [*looks at J ... points towards MP with pencil*.]
J: Yes.

Reference to quarter-point anchor

Child 1, R4. T2, range 0-20, target to estimate is 4.

J: Number 4
P: [*Looks briefly to MP, then takes rocket quickly to left part of line. Sticks*.] Just under a quarter.

Reference to three-quarter-point anchor

Child 1, R5. T4, range 0-100, target to estimate is 67.

P: [*Turns the page to first 0-100 trial*] ... up to a hundred … I think that's 75 cos I – looks like three quarters. [*Looks from rocket, to RE, to rocket. Hands rest at page edges, pause. RH pencil to rocket … writes "75"*.]

Reference to previous trial within the same range

Child 4, R1. T2, range 5-15, target to estimate is 9.

J: Nine.
E: [*looks from proffered rocket to MP … takes immediately to MP, sticks*.]
J: Very nice.
E: Cos the ten went there. [*RHI makes a sweeping gesture through the line, to the right of the 9 rocket, marking "there", the rocket space for 10*.]

Reference to previous trial in a different range

Child 8, R1. T2, range 0-20, target to estimate is 6.

C: But didn't you give me six a minute ago? [*turning back to look*]
J: Yes, it could be the same rocket again, that's alright.
C: Ah I put it a bit closer [to the LE on this 20 line, compared to 10 line]

Count-on from left endpoint to target

Child 9, R1. T4, range 0-100, target to estimate is 18.

L: [*Eyes sweep line from LE to rocket … LHI makes 4 evenly-sized jumps rightward from LE to blank rocket. Writes estimate.*]

Count-on from target to right endpoint

Child 12, R4. T2, range 5-15, target to estimate is 14.

J: Fourteen is the first one, where does that go?
C: [*whispering*] next to 15... [*Sticks quickly next to RE*.] 14 15! [*Taps just-stuck rocket, then RE.*]

Count-on from midpoint

Child 4, R3. T2, range 0-10, target to estimate is 5.

J: Number 5.
E: ... [*glances to LE, then rocket in hand, looks to right half of line, pauses*] ... … [*takes rocket to middle area of line … RHI sweeps from near RE to right of MP, and holds … LH takes rocket to MP and holds there whilst* <u>*RHI makes 5 emphatic jumps rightward from MP to RE*</u>*. Both hands stick rocket at MP*.] [whispered] dah. [*glances to LE*]

Count-back from right endpoint

Child 12, R5. T2, range 5-15, target to estimate is 13.

J: Thirteen...
L: [*smacks lips like a fish*]...um....hmmm [*looking around the room*] ... [*RHI with sticker makes 2 evenly-sized 'hops' leftward from RE … makes one further hop to the left, and sticks rocket. Glances to previous trial, then latest estimate.*]

Count-back from midpoint

Child 5, R4. T2, range 0-20, target to estimate is 9.

J: And number 9.
Z: [*eyes sweep from LE fast to middle area of line, then MP… takes rocket to MP, then makes* <u>*one jump leftward along line from MP*</u>*. Sticks rocket. Throws arms in the air (as if in victory).*]

Judgment using relative numerosities

Child 5, R5. T4, range 0-100, target to be estimated is 92.

Z: [*looks to LE, RE then rocket. Writes answer*.] ... Ninety ... eight.

Task responses

Change of mind

Child 10, R2. T2, range 0-100, target to be estimated is 18.

J: Eighteen.
M: [*Counts-on about 18-20 small jumps from LE rightwards …. then glides further rightwards to MP and sticks rocket*.]

Immediacy

Child 2, R3. T4, range 0-100, target to be estimated is 4.

J: And this one?
A: [*looks to rocket next to LE and writes immediately*]
J: Two, good.

Hazard –initial response led child to a recognised mathematical contradiction

Child 8, R1. T4, range 0-10, target to be estimated is 8.

J: Now this rocket over here?
C: [*goes to LE, glances RE, counts-on from LE*] 1 2 3 4 5 ...[*glances to blank rocket, carries on with silent count-on*] … 13 14 uhhhoo [*reaching rocket*], there's no 15 though. [*Sits up with puzzled expression, looks to J*].
J: No ...
C: But how, can it -? [*Looks to previous rocket, then to current trial … crosses arms*]

"Easy" – explicit indication that a trial was found easy

Child 1, R4. T2, range 0-100, target to estimate is 25.

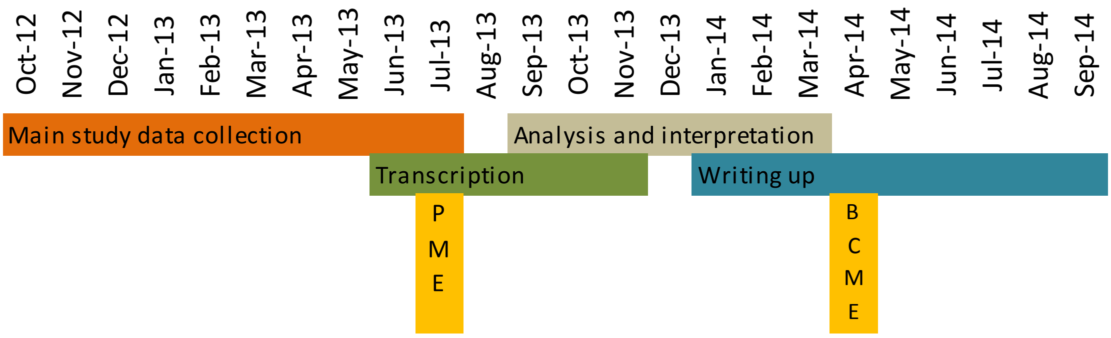
J: 25
P: That's quite easy cos it's a quarter. [T*aking rocket to left part of line, adjusting tiny amount, sticking down*.]

Appendix 7

Work plan after pilot study

This section briefly outlines the projected research progress and timeframe for completion. The main data collection is well underway, so the principal stages to be considered are data analysis and writing up. The following list and diagram show the planned timing of the research phases.

Oct 2012 – Jul 2013	Main study data collection
Jun 2013 – Nov 2013	Transcription of video data
Jul 2013	International Conference for the Psychology of Mathematics Education research report presentation (Kiel, Germany)
Sept 2013 – Mar 2014	Data analysis and interpretation
April 2014	British Congress of Mathematics Education presentation (Nottingham, UK)
Jan 2014 – Sept 2014	Writing up



The main study data collection is on schedule to be completed in early July 2013. The transcription phase will overlap with the final months of data collection since it is not necessary to wait – transcription of earlier data can and should begin sooner, to allow more time to reacquaint with the data and reflect upon it. The analysis and interpretation of data will similarly overlap with the transcription phase. Transcription is itself a stage of data analysis, of course, but in addition progress can be made on other analyses – particularly the statistical work – without needing to wait for all video interviews to be transcribed. Beginning earlier will allow more time to carry out the intended analyses, reflect upon possible interpretations and consider whether further analysis is required.

Writing will be carried out as part of the research process at all phases shown above, but it is anticipated that the remaining chapters of the thesis – in a fairly final form – will be written from January 2014 onwards.