Jones, Corinna Miriam

Mental Representations of Fractions and Decimals: Differences, Commonalities and Implications for Understanding

Mental representations of fractions and decimals: differences, commonalities and implications for understanding

Corinna Miriam Jones

A thesis submitted to the University of Huddersfield
in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

The University of Huddersfield
June 2017

Abstract

The purpose of this thesis is to seek evidence of commonalities in the mental representations of fractions and decimals between zero and one. The focus is on the mental representations of non-familiar fractions and decimals in adults. In addition, individual differences in the extent of common fraction and decimal mental representations are explored and their links to mathematical understanding of numbers between zero and one.

For whole numbers, number comparison tasks have found evidence of an ordered, magnitude mental representation known as the mental number line through which the magnitude of a whole number is automatically processed. This evidence consists of phenomena such as the distance effect and SNARC effect. Here, indications of a similar magnitude representation common to both fractions and decimals are sought through a task in which a fraction is compared with a decimal.

Substantial evidence of a distance effect is presented but not a SNARC effect, indicating that fractions and decimals can have mental representations containing or accessing a common magnitude but that this magnitude is not automatically processed.

In addition, two emergent phenomena are reported. The first is an effect of location which is contrasted with the size effect in whole numbers and a previously reported anchor-point effect. The second is a larger-stimulus effect which is an indication of differences in the mental representations of fractions and decimals. These effects are explored in two additional, simple magnitude and location tasks.

Furthermore, success but not speed within the comparison task is linked to strength of the distance effect for individuals. Therefore the number comparison task is repeated in series with a test designed to uncover common misconceptions of fractions and decimals. Patterns with the individual differences in responses to the test and comparison task are explored.

By making links between the features and commonalities of individuals’ mental representations of fractions and decimals and quality of their understanding, this research hopes to be of value to mathematical educators.

Contents

Word count 47 970

List of tables

List of figures

Chapter 1 Introduction and literature review

1.1 Prologue

The research contained within this thesis was inspired by my experiences as a teacher of mathematics. I teach in a sixth-form college whose students have finished their compulsory primary and secondary schooling and have performed well enough to pursue academic pre-university courses in their chosen subjects. By definition, these are at least reasonably able and motivated students. Many of them choose to study mathematics further and are successful in that pursuit.

However, there is always a significant proportion of these able students who have failed to reach the minimum standard in mathematics by age sixteen. When teaching, especially the poorest performing of these students, I am often struck by how their mathematical reasoning at the higher level tasks of algebra and problem solving is undermined by very weak and faulty proportional reasoning; by a complete lack of understanding of the meaning of fractions and decimals (between zero and one) and the links between these concepts. These are not students with an inability or unwillingness to learn but they have failed to successfully learn these key ideas.

Thus I embarked on this research to gain some understanding for myself of how people do and do not understand numbers between zero and one. I hoped to find out something useful that might inform teachers of mathematics and planners of education or at least highlight areas of further study that might be fruitful.

1.2 Definition of terms

This document concerns human understanding of small numbers between zero and one. Many different symbolic representations of these numbers are used by people in their everyday life and work. Due to the flexible nature of English, it is necessary to define precisely the meaning of the terms that are utilised throughout this thesis. The two main symbolic representations of number referred to within are fraction and decimal.

The term **fraction** always refers to a number between zero and one (not inclusive), that consists of two whole numbers separated by a horizontal or oblique line.

The term **decimal** always refers to a number between zero and one (again, not inclusive), that consists of a horizontal string of numbers beginning with a zero and a decimal point. An n digit decimal number has n digits following the decimal point.

In addition, the term **fractional numbers** is used to encompass both of the above symbolic representations. Whereas **proportion** signifies any value between zero and one with or without accompanying symbolic representation including, but not exclusively, fractions, decimals or percentages, both in symbols and in words. Proportional reasoning is therefore the knowledge and skills associated with the use of proportions.

The meaning of the term **mental representation** to which I will be referring within this thesis is the cognitive structure within the brain that acts as an internal model of our understanding and experience of an outside concept; the outside concept in this case being numbers between zero and one and their symbolic representation by way of fraction and decimal notation. The mental representation is the mechanism through which our brains interpret external stimuli and make judgements about them. As the detail of internal knowledge structures cannot be directly viewed, it is necessary to design experiments whose results allow us to infer the nature of these structures.

1.3 Review of the current literature

The purpose of this chapter is to bring together the relevant literature from education and cognitive science in order to establish the research basis for this thesis. There are three parts to the chapter.

The first part is based on research into education. It briefly describes the common problems and misconceptions that students have with respect to learning fractional numbers. It goes on to discuss how these misconceptions and poor understanding of fractional numbers create problems for students trying to gain access to higher mathematical knowledge. This demonstrates why research into understanding of fractional numbers is so important.

Next, this first part highlights how the environment in which one learns about fractional numbers and proportion affects the types of proficiencies as well as the types of misconceptions that one acquires. Thus implying that the emphasis placed on the subject matter by the teacher or learning environment affects the mental representation formed.

The second part of this chapter is a short summary of the relevant research into mental representations of number. It starts with a look at the beginnings of this discipline, that is, research into whole number representations. There is a summary of some of the tasks used to gain insight into the hidden structures of the mind and the inferences made from them.

Then, the current literature with respect to the extension of these techniques to the study of mental representations of fractional numbers is explored. The findings common with and different from those for whole numbers are dealt with as well as issues specific to the representations of fractional numbers.

The third and final part of this chapter brings together all of the above to give an account of the gaps in the current research that are tackled in this thesis.

1.3.1 What education research has to say about mental representations of fractions and decimals

This section is a short summary of educational research relevant to the topic of mental representations of fractions and decimals.

1.3.1.1 Problems and misconceptions

There has been a considerable body of research into mathematics education. Within this research it has been demonstrated many times that children particularly struggle to learn the skills and knowledge associated with fractional numbers, e.g. Behr, Lesh, Post & Silver (1983), Dickson, Brown & Gibson (1984), Kerslake (1986) or more recently Gabriel, Coche, Szucs, Carette, Rey & Content (2013).

Very much the same errors in procedures but also in conceptual knowledge were made by the children in all of these studies. In addition, these errors were found to persist throughout the school years – up to the age of 15 years by Dickson et al. (1984) and to 17 years by Behr et al. (1983). The study of Hasemann (1981) found that older school children between the ages of 12 and 15 had generally learned to successfully use mechanical methods for calculations involving fractions but could demonstrate very little of the conceptual understanding that underpinned these methods. Furthermore, Koch & Li (1996) found that even college students with 12 years of mathematical learning behind them still focussed overly on surface knowledge and failed to make the relevant conceptual connections between their areas of rational knowledge.

It seems that fractional numbers concepts are universally challenging to acquire and process. This is of concern to educators because fractions, decimals and proportional reasoning in general play a key part in many aspects of our daily life from cooking to managing our finances.

1.3.1.2 Fraction understanding and progress in learning higher mathematics

In addition to the extensive scope for misunderstanding and misconceptions in the subject, it might also be asked whether there are other reasons why poor understanding of fractions and decimals should be a key area of concern for mathematics educators. Anecdotally, teachers 'know' that an individual student's progress, in areas of mathematics such as algebra, indices and calculus, can be hampered by inconsistent and faulty knowledge of fractions, decimals and proportional reasoning generally. Indeed, Behr et al. (1983) comment on this as being one of their main motivations for conducting research into the improvement of the learning of proportional concepts and reasoning.

There is some empirical evidence to support teachers' belief of a connection between a student's skills with proportion and their readiness to learn higher mathematics. Siegler, Thompson & Schneider (2011) found that understanding of fraction magnitudes was the most reliable indicator of future progress at higher mathematics for school-aged children but not that the former directly affected the latter.

Additionally, Booth & Newton (2012) established a causal link. They conducted a school-based study into progression from learning about number to learning about algebra. They were able to find direct evidence that, more than any other factor, a poor understanding of the magnitude of fractions affects the readiness of school children to develop an understanding of algebra. This points to specifically *magnitude* understanding of fractions as a key area of focus when trying to help students progress mathematically.

1.3.1.3 How knowledge is acquired affects mental representations

Instinctively one would think that how the mathematics of fractional numbers and proportional reasoning is acquired would have an impact on the nature of the underlying cognitive structures formed. There is some evidence to support this instinct.

Schliemann & Carraher (1992) studied unschooled children in Africa and Brazil with specific jobs that required proportional reasoning and compared their skills with those of their peers in US schools. What the researchers found was that in the skills and context specific to the jobs they held, the unschooled children vastly out-performed their American counterparts. However, they had little or no understanding or skills outside of what was required by their jobs and no ability to generalise their skills to other situations. Whereas, the American children had broader, if shallower knowledge of proportion and were much more flexible in their ability to apply it.

What is shown by Schliemann & Carraher (1992) is indeed that "*Proportional reasoning could not occur at all if there were not cognitive structures available for sustaining the representation and comparison of ratios*"(pp. 70). Also, that the milieu in which you learn about proportion and the emphasis placed upon it affects the development of these cognitive structures. Part of these cognitive structures must be the mental representations of fractional numbers and proportion generally.

The study of Resnick et al. (1989) reinforces this conclusion. They compared the procedural errors and misconceptions made by children in Israel, France and the USA. In these three countries, the order in which the concepts of integers, fractions and decimals were introduced and the emphasis placed upon them were different. What Resnick et al. found was that they could differentiate between the types of misconception held by children in the different countries. They highlighted the fact that the different styles of approach to teaching decimals and fractions had had an impact on the types of mental representations that had been formed in the children's minds. Hence the researchers postulated the importance of finding out which mental representations are formed in the heads of competent mathematicians and using this information to inform educators how to design courses which engender functional mental representations of fractional numbers.

1.3.2 Mental representations of numbers

This section starts with a review of the evidence relating to the 'mental number line' for whole numbers. The mental number line theory postulates that the mental representation of whole numbers is an analogue to a spatial number line. This implies that the mental representations of whole numbers are arranged in a size-ordered array. Indeed, de Hevia & Spelke (2009) found a special link between numerical and spatial concepts both in children and adults.

The next part of the section is devoted to the evidence for and against the extension of the mental number line to fractional numbers. This encompasses research seeking automatic and deliberate responses to fractional stimuli, incorporating evidence from cognitive neuroscience.

Then additional phenomena related to the position of fractions on the number line will be explored; in particular, the concept of anchoring and its possible relevance to mental representations of fractions.

1.3.2.1 The mental number line for whole numbers

Evidence for a mental number line includes experimental data on both automatic and deliberate responses to magnitude data. An automatic response is typically sought by an experiment in which the magnitude of the number is not relevant to the task at hand. If a magnitude-related pattern is found within the responses to the task, this implies that the magnitude of a number is automatically processed even when not relevant. Furthermore, this pattern indicates that magnitude is an inextricable part of the mental representation of number. However, these tasks do not generally generate any further insight into the *structure* of any magnitude-based mental representation of number.

More light can be thrown upon such structures by use of tasks that elicit a deliberate response to the magnitude of numerical stimuli. By studying patterns of responses to such tasks, inferences can be drawn about how magnitude representations of number are interrelated within the mind.

1.3.2.1.1 Mental representations of single-digit whole numbers

Several studies have demonstrated that single-digit whole number magnitudes are unconsciously accessed, even when irrelevant to the task in hand. Chiou, Wu, Tzeng, Hung & Chang (2012) for example, found that grip aperture can be affected by the presence of numerical labels. When participants were given an object to grasp which was labelled with a `5', their grip aperture was larger when it was accompanied by a smaller number (e.g., 2) than by a larger number (e.g., 8). Also Lindemann, Abolafia, Girardi & Bekkering (2007) similarly showed that large grip apertures were more quickly formed in response to larger number and small grip apertures more quickly primed by small numbers.

Another example of unconscious activation of numerical magnitude is the size (in)congruity effect (SCE or SiCE). When people are given the task of selecting the *physically* larger of two digits it can be observed that they are faster if that digit is also the *numerically* largest of the pair (Henik & Tzelgov, 1982).

The most commonly deliberate task used to seek evidence of a mental number line is the number comparison task. In a number comparison task the larger (or smaller) number must be selected from a pair of numbers presented usually simultaneously but sometimes sequentially. The distance effect, (Moyer & Landauer, 1967), is the phenomenon that response times for this comparison task increase as the numerical distance between the pair of numbers is reduced.

This distance effect is taken as evidence for a mental number line because it is thought that when the brain receives input from two numbers close to one another in magnitude, the mental representations of these two numbers are also close to one another. The consequent interference slows down the comparison response. The distance effect can also be found in spatial comparison judgements Johnson (1939), further reinforcing the theory of a mental number line analogous to a spatial number line.

The interpretation of the distance effect as evidence of a mental number line, is not entirely unchallenged. Cohen (2010) asserts that an alternative explanation for the effect could be the similarity of the visual input of numbers that are close in size. Goldfarb, Henik, Rubinsten, Bloch-David & Gertner (2011) investigated this assertion by seeking the distance effect in both an automatic number matching task in which number magnitude was an irrelevance and a deliberate number magnitude comparison task. They found that the similarity of the visual input of the two numbers can explain the distance effect found in the automatic task but not the deliberate one. This outcome of Goldfarb et al. indicates that a number comparison task might be the best choice for an experiment seeking evidence of a mental number line for fractional numbers.

Another phenomenon that occurs within the number comparison task, found by Moyer & Landauer (1967) is sometimes known as the 'size effect'. It is that if the distance between stimuli is controlled then response latencies are smaller for smaller stimuli than they are for larger stimuli. This is interpreted as further evidence of mental representations of number that mirrors a spatial, magnitude ordered number line. This size effect result was also found for whole numbers, both positive and negative, by Ganor-Stern & Tzelgov (2008).

The number comparison task has additionally been used to elicit an automatic response to a deliberate task. The spatial-numerical association of response codes (SNARC) effect has been found in single-digit number comparison tasks (Dehaene, Bossini & Giraux, 1993). The SNARC effect is that when people are presented with two numbers to compare, they are faster to select the larger number if it is on the right. This result demonstrates not an unconscious activation of irrelevant magnitude data but instead an activation of irrelevant spatial data in response to magnitude. It implies a right-left alignment to a mental number line, reinforcing the idea that numbers have a mental representation analogous to a spatial number line.

The experiments detailed above involved responses to single digit number stimuli only. Whilst some support has been found for an extension to the mental number line theory to two digit whole numbers, experimental results are mixed. It is most important that these two digit numbers can be shown to be represented holistically rather than as two separate digits in order for the evidence to imply a continuation of the mental number line beyond single digits.

1.3.2.1.2 Mental representations of two-digit whole numbers

Indeed, Dehaene, Dupoux & Mehler (1990) did find a distance effect for two-digit numbers with some discontinuities at the decade breaks. Unlike in Moyer & Landauer's (1967) experiment in which the two single-digit numbers were presented simultaneously, Dehaene et al. (1990), used a target-stimulus paradigm for their number comparison task. There were three experiments with three targets numbers (55, 65 & 66), stated at the start of the experiment against which numerical stimuli were compared. They also ran an experiment in which the digits of the stimulus were presented separately. Though their results were somewhat mixed they concluded that the best explanation for the distance effect they found was that the holistic magnitude of the target stimulus had been internalised and the size of the stimuli were accessed in order to make the comparison.

In contrast with this result, Zhang & Wang (2005) did not find evidence for holistic magnitude processing in a two-digit comparison task in which the target and stimulus were presented simultaneously. They found instead that the best explanation for their results was that the two digits of the numbers were being compared in parallel. This parallel approach means that the single digits in the tens position are first considered. If a judgement cannot be made on these first digits, only then are the units digits compared.

Taken together, these two studies could be taken to imply that holistic mental representations of two digit numbers are accessible but they are only used when the task at hand requires them. Conversely, Fitousi & Algom (2006) and Ganor-Stern, Tzelgov & Ellenbogen (2007) found the size congruity effect (SCE) present in certain arrays of two digit numbers. Additionally, Zhou, Chen, Chen & Dong (2008) found a SNARC effect present in a two-digit number matching task. These results imply that the holistic magnitude of two digit numbers might be automatically accessed even when irrelevant to the task at hand.

So there is strong evidence for a mental number line for whole numbers, particularly single-digit numbers. Magnitudes of single-digit numbers are automatically accessed even when not relevant to the task at hand. Holistic magnitude mental representations of two-digit numbers are available if required and are sometimes accessed automatically.

1.3.2.2 Mental representations of fractional numbers

The question now arises is whether this mental number line is available as a mental representation of fractional numbers. The current literature on mental representations of fractional numbers demonstrates, in general, the same mixed, task-dependent results as for two-digit numbers.

1.3.2.2.1 Fraction comparison studies

For example, Bonato, Fabbri, Umiltà & Zorzi (2007) carried out a number comparison task using pairs of fractions with magnitudes between zero and one and found no distance effect based upon the holistic magnitude of the fractions but instead, distance effects based on the numerator or denominator magnitudes. They concluded that no holistic magnitude representation for fractions exists but that instead, the whole number magnitudes of the components of the fractions, (numerators and denominators), are processed separately. This result is similar to that of Zhang & Wang (2005) above for two digit numbers in that numerators were compared with numerators or denominators with denominators just as tens were compared with tens then units with units.

The experiment of Bonato et al. (2007) has been criticised however, on the grounds that the pairs of fractions used with the experiment *could* all be compared by considering the numerator or denominator alone (e.g., 1/3 with 1/5). Thus these results do not refute the possibility of holistic magnitude internal representations of fractions but instead they highlight the probable primacy of single digit numbers over other numbers.

Hence, subsequent number comparison studies ensured that pairs of fractions could not be compared using simply numerator or denominator components alone. These studies consequently did produce evidence of a holistic magnitude distance effect for fractions e.g. Meert, Grégoire & Noël (2009), Schneider & Siegler (2010), Sprute & Temple (2011).

Combining these studies to one in which fractions both could and could not be compared componentially, Obersteiner, Van Dooren, Van Hoof & Verschaffel (2013) found an intuitive bias towards using whole number component techniques to make magnitude comparisons of fractions when possible, even amongst experts. Yet this was switched to a holistic magnitude representation when the task required. In addition, Faulkenberry & Pierce (2011) found a distance effect for fraction comparisons that was mediated by task-dependent strategy.

There is also evidence from cognitive neuroscience to support the existence of mental representations for the holistic magnitude of fractions. A significant distance effect was found in the neural adaptation study of Jacob & Nieder (2009). Adaptation in this context means presentation with a stimulus for a sufficient length of time for the activity associated with its presentation to subside. This activity can be observed via fMRI as increased blood flow in the relevant part of the brain. So, participants were *adapted* to the fraction 1/6. A new fractional number was then presented. Jacob & Nieder found that the time taken for the participants' brains to adapt to the new stimulus was a function of the numerical difference between the new stimulus and 1/6.

The results of these distance effect studies for fractions mirror those for two-digit whole numbers. Unless the comparison requires holistic representation, only the mental representations of the whole number components of fractions are generally accessed. Yet holistic magnitude representations can be formed and used when necessary.

1.3.2.2.2 Automatic magnitude responses to fractions

Unlike for two-digit whole numbers, attempts to reproduce unconscious/automatic activation of magnitudes for fractions similar to those of whole numbers have proved entirely unsuccessful. Bonato et al. (2007) sought, but did not find the SNARC effect in their fraction comparison tasks. However, this could well be due to the nature of the fractions they used and the consequent strategies adopted by their participants.

Kallai & Tzelgov (2009) found that although the SiCE persisted between pairs of numbers in which one was a proper fraction (between zero and one) and the other was a whole number, it did not occur in physical size comparison tasks in which both stimuli were fractions. They concluded that there is one internal magnitude representation that encompasses all numbers less than one.

This conclusion is probably too strong because if it were true then the distance effect that has been confirmed for fractions (by e.g., Faulkenberry & Pierce, 2011, Meert et al., 2009, Schneider & Siegler, 2010 and Sprute & Temple, 2011), would be unexplained. A more plausible explanation could be that because processing fractions requires additional effort and therefore time. So there is less opportunity for such automatic processes to have a significant effect on the larger response times. To support this explanation it can be noted that these studies found *average* response times for fraction comparison tasks that were in the region of 700ms to 1300ms. This contrasts with the average response times for whole number comparisons of Moyer & Landauer (1967) which were between 500ms and 650ms.

1.3.3 Differences between fractional numbers and whole numbers

1.3.3.1 Anchor points

Fractional numbers, unlike whole numbers are bounded, lying as they are between zero and one.
As mentioned above, there are discontinuities detectable (decade breaks) in the mental number line for two-digit numbers. There is some evidence that similar discontinuities around zero, one and half might exist within the mental representations of fractional numbers.

In previous studies it has been found that the numbers zero, one and (to a lesser extent) half act as reference or anchor points on the number line which can affect performance on estimation (see Hollands & Dyre (2000) for a review). This implies the numbers zero, one and half are influential 'anchor points' for magnitude estimation of small numbers, maybe in the same way that the decades are for two-digit numbers.

For example, Varey, Mellers & Birnbaum (1990) found that when estimating the size of a proportion of black to white or white to black dots present in an array, participants produced an 'inverse ogival' pattern of results. This is to say, they overestimated proportions between zero and half and underestimated proportions between half and one. Additionally, Cohen, Ferrell, & Johnson (2002) found that there was bias in estimation of small numbers around zero, half and one.

These magnitude estimation investigations point toward a possible discontinuous or, at least, non-linear nature to the mental number line for fractional numbers with maybe zero, half and one acting as anchor points against which to judge magnitude. In other word, points on the number line in relation to which fuzzy magnitude representations of other proportions are judged. Classically Tversky & Kahneman (1974) demonstrated that people make judgements that are skewed towards initially presented bounds for estimation (the anchors). They concluded that people use the anchors as starting points for their estimation. They then utilise a heuristic by which they adjust their estimations away from the anchors leading to the observed skew in judgements. Since then many studies have demonstrated the influence of anchor points on judgements of number, probability and spatial proportion but other mechanisms by which the observed effect takes place have been proposed (see Furnham & Boo (2011) for a review).

There is however scope for more conclusive investigation of the importance of anchor points specifically in the mental representations of fractional numbers.

1.3.3.2 Symbolic representations

A further difference between whole numbers and proportions is the range of external symbolic representations each has in common usage. Though some bilingual people might use more than one language or number system, the majority of people have only two regularly used symbolic representations of whole numbers – words and numerals. Numbers between zero and one have a multitude of symbols that used in everyday life such as words, numerator/denominator fractions (of which there are infinite equivalent versions), ratios, decimals and percentages. Nonetheless, there are very few studies into the mental representation of proportion that have looked beyond fractions. Most of the existing literature concerns studies of magnitude estimation or conversion rather than magnitude comparison tasks.

One that *has* utilised a comparison task to compare different external representations is the neural adaptation study of Jacob & Nieder (2009), detailed above. Their distance effect result was found to be consistent across all combinations of stimuli and presentation: both as fractions and decimals, in symbols and in spoken words. This does seem to imply common magnitude representations across symbolic input.

However, on the face of it, the study involving both relative frequency fractions (1/x) and decimals conducted by Cohen et al. (2002) might be interpreted as leading to the opposite conclusion.

In their fifth experiment, participants were given the task of converting between relative frequencies and decimals or vice-versa. Participants' responses were extremely inaccurate (on average 13% correct) leading Cohen et al. to conclude that there is no single internal magnitude representation for very small numbers. However the stimuli were *extremely* unfamiliar, small quantities (e.g., 1/63) making the conversion process very arithmetically taxing.

The results of Cohen et al. can certainly be taken to indicate that there are limits to any concrete, common internal representations for fractions and decimals. Indeed, if they had instead looked only at conversions between 0.5 and half or 0.2 and one fifth they might have made entirely the opposite conclusion. If a broad understanding is to be shaped of human abilities to form functional, common mental representations of fractions and decimals, great consideration must be given to the limits of this ability. Experimental stimuli must be chosen to gain the greatest insight into these limits.

One last study to be considered in this section is that of Iuculano & Butterworth (2011) which seems to indicate that, for familiar examples of fractions and decimals at least, some commonality in magnitude understanding but also some differences. They conducted two number estimation tasks; one in which numbers had to be placed on a number line and another in which the value of marks on the number line had to be given a value. They found no difference in the linear pattern of accuracy of estimation between decimals and whole numbers. However, for fraction estimations, they found the pattern of responses was task-dependent with a non-linear pattern of response for the second estimation task for adults.

Only very commonly used fractions and decimals such as half and quarter were used as stimuli in this experiment and the differences found between fractions and decimals were subtle. This study seems to indicate some commonality but also some subtle differences in the magnitude representations of fractions and decimals. However, as fractions and decimals are not directly compared in the task, the limits of these commonalities and differences are not clear.

1.4 Summary of the findings and questions remaining

The above literature highlights the need to understand the mental representations of fractional numbers. It has been established that teaching emphasis can have an effect on understanding and misconceptions. Therefore, if we know what mental representations are available and even those that are more functionally useful, this will aid educators in guiding the understanding of students of mathematics.

So far it seems fairly well established that fractions can have a magnitude mental representation akin to that of the mental number line for whole numbers. What is not so clear is whether this is shared with or linked to a mental number line for decimals. Could a distance or even a SNARC effect be found for a magnitude comparison tasks between these two?

What is more, little account has been taken of the position of small numbers within the zero to one range. Is the mental number line for fractional numbers linear?

Chapter 2 Experiment one

2.1 Introduction

This chapter is a commentary on the first of the experiments carried out for the purpose of this thesis. It starts with a summary of the justification of the experiment. That is, how it was intended to extend and elaborate on existing findings on the cognitive processing of small numbers. Then follows details of the design of the experimental task and stimuli; with reference to how they were devised to addresses the intention of the experiment.

The second part of the chapter covers the experimental procedure. Then the third section contains the results of the experiment along with the methodology of the analysis.

Finally, the last section of the chapter is a consideration of the implications of the results of the experiment and areas for further investigation.

2.2 Justification and design

The previous chapter reviewed the evidence that supports the theory that there is a human capacity for holistic magnitude mental representations of fractions that is somewhat akin to the mental number line found for whole numbers. This evidence comes mainly from the distance effect found in the fraction comparison experiments of Faulkenberry & Pierce (2011), Meert et al. (2009), Obersteiner et al. (2013), Schneider & Siegler (2010) and Sprute & Temple (2011).

However, the previous chapter also showed that it has not yet been established whether this mental number line for fractions is related to the mental representations of other symbolic representations of proportion such as decimals. The question remains whether cognitive structures exist that allow us to translate between these magnitudes even for less familiar examples of proportion? The study of Iuculano & Butterworth (2011) seems to imply only subtle differences between the mental representations of fractions and decimals. However, they did not require participants to compare fractions and decimals directly. Instead they compared the outcomes of fraction and decimal estimation tasks performed separately. Moreover, they used a relatively restricted set of examples of fractions and decimals.

In contrast, the study of Cohen et al. (2002) implies no commonality at all between mental magnitude representations of fractions and decimals. Though this study did require participants to translate directly between fractions and decimals, the stimuli were exceedingly unfamiliar and very small indeed, ranging in size between 0.0001 and 0.01. A lack of ability to translate directly and exactly between extremely unfamiliar fractions and decimals does not necessarily imply there is no shared magnitude understanding albeit approximate in nature.

Taking these contrasting experimental approaches and results into account, a magnitude comparison task was designed directly comparing the magnitude of a fraction with that of a decimal. As the distance effect was the key result being sought, stimulus pairs were controlled for the magnitude of the distance between the fraction and the decimal.

In addition, relatively unfamiliar fraction and decimal stimuli were selected for the task for two reasons. Firstly, the use of very familiar fractions such as half and quarters and tenths would introduce the possibly confounding factor of familiarity. There is limited availability of familiar fractional numbers and including familiarity as a factor would be problematic as it would be difficult to quantify.

The other reason for the use of relatively unfamiliar fractional stimuli is to control for any effect of position within the zero to one range. The use of fractional stimuli with denominators of 11, 13, 15, 17 and 19 allows for stimulus pairs in many positions across the zero to one range. There are two effects of position that might influence the response times in a magnitude comparison task.

Firstly, Moyer & Landauer (1967) found a size effect for whole numbers in their single digit number comparison tasks. That is, if for the same distance between stimuli, response latencies are longer for larger numbers. This result was found also for two digit numbers by e.g. Dehaene et al. (1990); Nuerk, Weger & Willmes (2001). It is possible that even within the small range of numbers between zero to one, response times are longer for larger stimulus pairs than for smaller stimulus pairs with the same distance between them.

Secondly, a possible influence of *anchor points* might be present on approximate size estimation of fractions and decimals. It has been established that judgements of probabilities (which are also numbers between zero and one) are, like so many other judgements, subject to the anchoring effect (e.g. Chapman & Johnson, 1999). If size judgements of proportions are affected by anchoring, then it is possible that size comparison judgements are affected also.

Traditionally, in tests of anchoring, the anchors used are explicit rather than implicit. If people are informed that they are to make magnitude judgements about numbers between zero and one, these two ends of the scale might be considered explicit anchors. Yet, bearing in mind that anchors are known to influence the final judgements made such that estimates are closer to anchor points than they should be, the inverse ogival pattern of estimation errors for numbers between zero and one demonstrated by Cohen et al. (2002) implies that the likely anchor points for judgements of proportion are zero, half and one.

Thus it was necessary that response times for each distance between stimuli be measured over a full range of positions between zero and one. This should to eliminate any effect due to position interfering with detection of the distance effect. Furthermore, by selecting stimulus pairs in this way it is also possible to test whether there is an effect of position on response times in addition to the distance effect. That is, to test whether an equivalent of the size effect exists for numbers between zero and one, irrespective of their symbolic representation.

In addition to distance between stimuli and their position within the zero to one range, two further factors were built into the experiment. Stimuli were balanced for both left-right position of the largest stimulus and for whether the largest stimulus was the fraction or the decimal. Balancing for both of these factors allowed for an increased number of stimuli. The factor of whether the larger stimulus was a fraction or a decimal was not expected to have any influence on the results but was included for balance and as a means to double the number of stimuli presented.

Balancing the stimuli for both left-right position of the largest stimulus not only allowed for an increased number of stimuli but it also was intended to eliminate any confound in the results due to the SNARC effect. In addition, it provided a means of detecting a SNARC effect if present. A SNARC effect would be found if stimulus pairs in which the larger stimulus was on the right had significantly smaller response times than those on which the larger stimulus was on the left. However, the detection of a SNARC effect was not anticipated. This is because it would imply that the mental magnitude representations of *both* unfamiliar fractions and decimals are automatically processed. Further, that these are automatically accessed as magnitudes upon a common scale. As detailed in the last chapter, there is little evidence from previous studies to suggest this automatic magnitude response would be found.

The physical appearances of the two numbers being compared were quite dissimilar – one had a vertical display with a horizontal line dividing two numbers and the other, a horizontal display starting with '0.' followed by three digits. Although there were, by chance some similarities between the digits of some of the stimulus pairs this was not in any systematic manner. Thus the challenge of Cohen (2010) to the interpretation of the distance effect is answered. Any influence of physical similarity of stimuli on response latencies would not account for the distance effect being detected.

With all of these factors taken into account, any significant distance effect found would indicate that, to some extent, people have a functioning magnitude representation that has commonalities between the two symbolic representations of unfamiliar fractions and decimals. This result was anticipated but the strength of the result was expected to be possibly mediated by the position of the pairs of stimuli within the zero to one range.

2.3 Method

2.3.1 Participants

Thirty-one healthy adults (20 women and 11 men) aged between 20 and 64 years (M=36.7, SD=12.3) participated in the study. Seven were students at the University of Huddersfield. The remaining 24 were (non-mathematics) teaching and support staff at a sixth form college. One participant chose not to complete the experiment and was not included in any analysis.

2.3.2 Stimuli

Stimuli consisted of two relatively unfamiliar numbers between 0 and 1. That is, one fraction presented in its simplest terms with denominator of 11, 13, 15, 17 or 19 and one three-digit decimal number. The decimals were generated by adding and subtracting the distances 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5 from each fraction and rounding to three decimal places, ignoring pairs in which the decimal fell outside of the range 0 to 1. This resulted in a list of 582 possible stimulus pairs; 291 in which the decimal was the larger number and 291 in which it was the smaller.

To control for and identify any possible influence of the anchor points 0, 1 and 0.5, for each distance, the stimuli were grouped according to their position within the zero to one range The details are outlined below with figure 2.1 for illustration.

For distances 0.05, 0.1 and 0.2 there were seven positions presented:

A. The smaller stimulus below 0.1
B. Both stimuli between 0.1 and 0.4
C. The larger stimulus between 0.4 and 0.5
D. The stimuli on either side of 0.5
E. The smaller stimulus between 0.5 and 0.6
F. Both stimuli between 0.6 and 0.9
G. The larger stimulus above 0.9

It was not possible for distances above 0.2 to be presented in all seven positions. Therefore for distances 0.3 and 0.4 there were three positions presented:

H. Both stimuli below 0.5
I. The stimuli on either side of 0.5
J. Both stimuli above 0.5

For the distance of 0.5, all stimulus pairs would be either side of 0.5. So the following three positions were presented:

K. The smaller stimulus below 0.1
L. Both stimuli between 0.1 and 0.9
M. The larger stimulus above 0.9

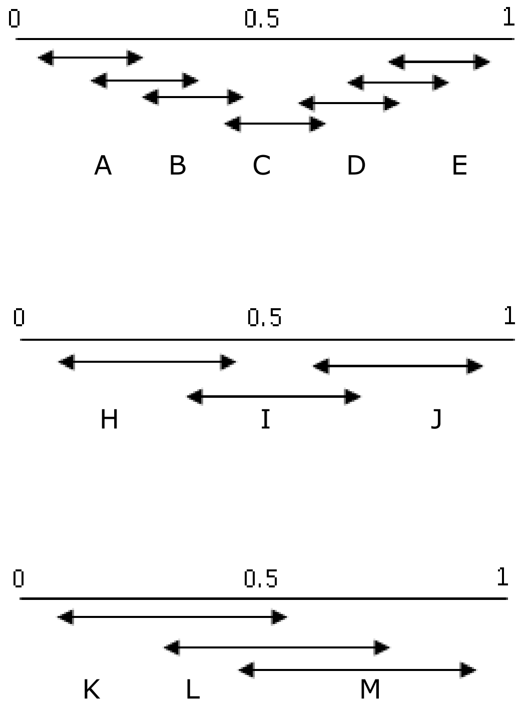

Figure 2.1 Illustration of the levels of the position factor

The 582 possible stimulus pairs were thus split into 60 groups; two for each of the 30 combinations of distance and position; one in which the decimal was the larger and one in which it was the smaller. Two stimulus pairs were randomly chosen from each of the 60 groups. In order to detect any evidence of the SNARC effect, one of these was presented with the larger number on the right and the other with the larger number on the left. Thus there were a total of 120 stimuli.

The fraction was presented vertically with a horizontal line separating the numerator and denominator, each of which had one or two digits. The decimal fraction was presented as a zero followed by a decimal point and three digits to the right of the decimal point. Stimuli were displayed in black type on a white background.

2.3.3 Procedure

The experiment was conducted using SuperLab® 4.0 stimulus presentation software in the participant's place of study or work in a quiet, well-lit room. Participants were instructed that within each trial they had to decide which of the two numbers presented was the largest and to press the 'A' key if it was the number on the left and the 'L' key if it was the number on the right. They were informed that both speed and accuracy of response were important. In addition, it was made clear that some of the tasks were expected to appear very easy and some extremely difficult and the purpose of the experiment was to find out what factors made the task more difficult.

A practice block of four stimuli preceded the experimental blocks. Participants were given feedback on their accuracy on the practice stimuli and were allowed to ask questions if they did not understand the procedure. The 120 experimental stimuli were then presented in random order in three blocks of 40 with no further feedback on accuracy nor opportunity to ask questions.
Participants were given the opportunity to take a break between the blocks. All participants were presented with the same stimuli.

Response times and accuracy were recorded by the program.

Following the experimental blocks the participants were questioned on the strategies and reasoning they had used to complete the tasks. To facilitate their explanation they were presented with three further stimuli each with a distance between the fraction and decimal of 0.2 and asked to choose the largest number and explain their reasoning. These explanations were recorded on film.

2.4 Results

2.4.1 Pre-analysis data processing

In the post-task interview, two participants said they had chosen the smaller stimulus instead of the larger. Their results bore this out with one giving only 14 correct responses and the other 4 (out of 120). The interview took place immediately after the task and they were able to give correct verbal responses to the three further stimuli. Therefore their responses were included in the analysis with the error responses recoded.

Taking this into account, the number of errors per subject ranged from 3 to 45 (out of 120). The worst of these, with 45 errors, had a probability of only 0.00392 of achieving this result or worse by guesswork alone (B(120,0.5)). Thus all participants can be considered to have performed better than chance at the task and none were excluded from the analysis.

2.4.2.1 Skew

Using the method of Crawley (2005), it was found that the RT data were highly positively skewed ($\gamma = 3.31$, $p < .001$). This skew can be seen in figure 2.2. Therefore, initial exploratory graphs used the median as a measure of central tendency.

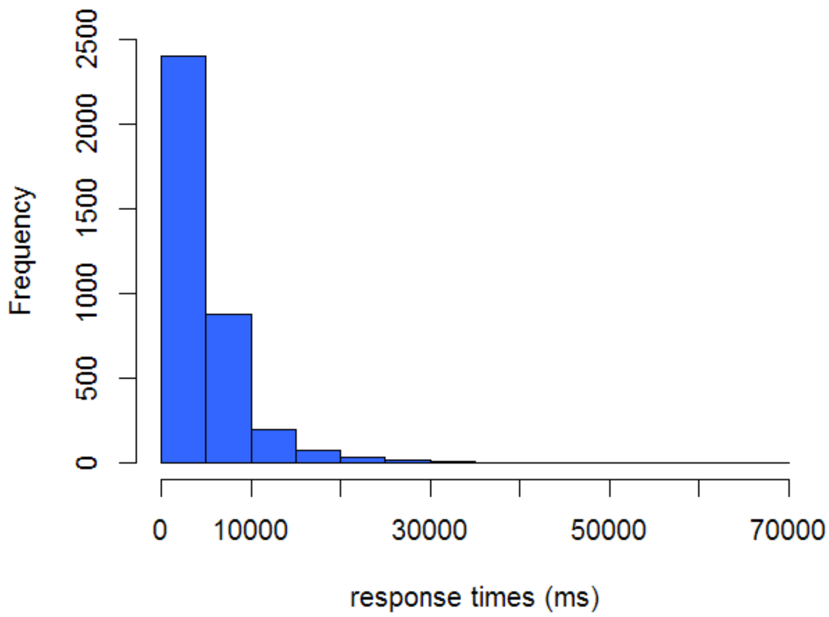

Figure 2.2 Histogram showing distribution of response times for experiment 1

2.4.2.2 Distance effect

The graph of median RT against distance (figure 2.3) demonstrates a distance effect such that RTs decreased as distance between stimuli increased. The effect appears to have been approximately linear for distances between 0.05 and 0.4. However there was little difference in median RT between distances 0.4 and 0.5.

The Welford function was used by Moyer & Landauer (1967) to link their results for the numerical distance effect to previous distance effects for physical size. It is a linear model using log(larger stimulus size/distance between stimuli) as a predictor for RTs. To test the fit of the Welford function to the experiment one data, the mean average RT was calculated for each stimulus pair.

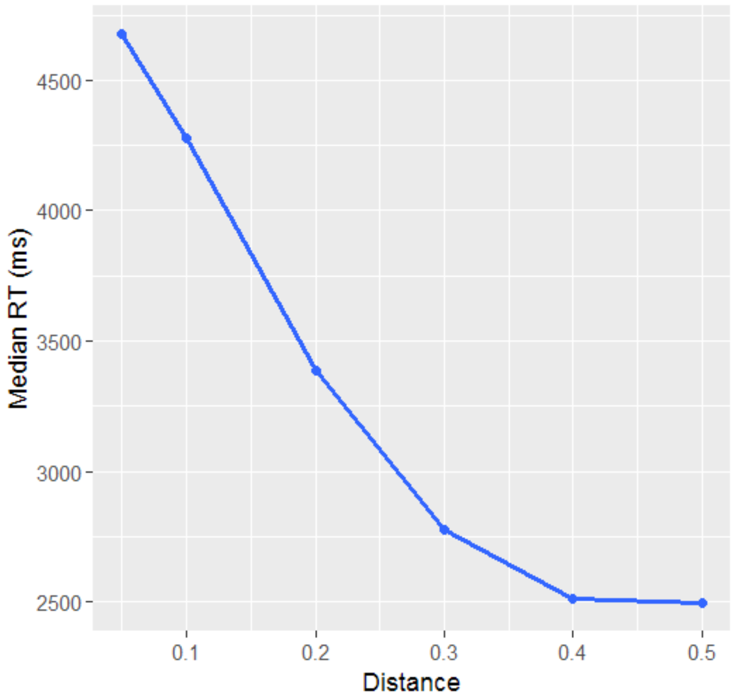

Figure 2.3 Graph of median RT against distance for experiment one

A small but significant association was found (r_s = .425, p < .001) between log(larger stimulus size/distance between stimuli) and mRT, indicating a reasonable fit of these data to the Welford function. This implies a distance effect was present that, to some extent, mirrors that found for physical magnitude comparison. However, the association between simply distance and mRT was much larger in size (r_s = -.681, p < .001).

The association of the absolute difference between the numerator of the fraction with the first decimal place of the decimal was also tested. This difference between the numerator and first decimal is not independent of the distance between stimuli. There was some concern that any distance effect found might actually be due to participants comparing these two values rather than the holistic magnitudes of the stimuli. However, though significantly associated, the association was not as strong as either of the associations above (r_s = -.363, p < .001).

2.4.2.3 Position factor

As detailed in section 2.2 stimuli pairs were placed in different sets of positions for the different distances. There were seven positions for distances 0.05, 0.1 and 0.2 and three positions for the larger distances. I judged that if no effect of position could be observed, or if an effect was found only for certain positions, this might simplify somewhat the analysis of the RTs. In particular, if some of the seven positions A to G for the smaller distances of 0.05, 0.1 and 0.2 could be combined, the RT (and error) analysis could be simplified.

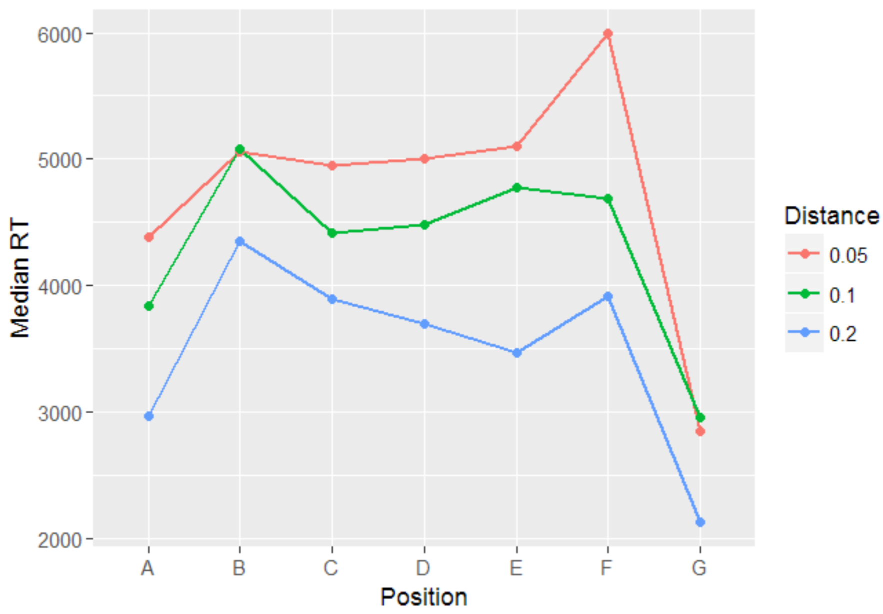

Figure 2.4 Graph of median RT against position by distance for experiment one

The graph of median RT against position for the smaller distances, (fig 2.4), indicated that there might not be a significant difference between the median RTs for all of the positions A to G, particularly the middle positions B to F. In light of this, analysis was carried out to test, first for a significant effect of position on RTs; then for a significant difference between each pair of positions for the distances 0.05, 0.1 and 0.2.

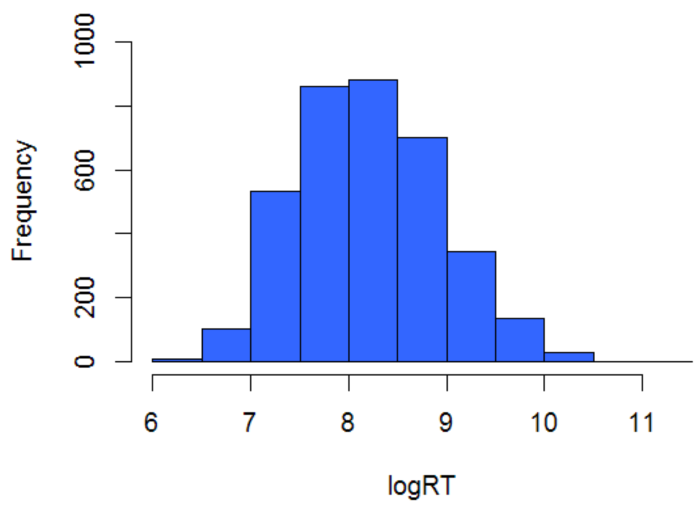

Figure 2.5 Histogram showing distribution of logRT for experiment 1

For the purpose of this analysis, and all further parametric analysis, a natural logarithm (log) transform was applied to the RT data, (see fig 2.5). This resulted in a much less skewed distribution, (γ = 0.289, p = .386), which a Q-Q plot shows to be approximately normal, (fig 2.6).

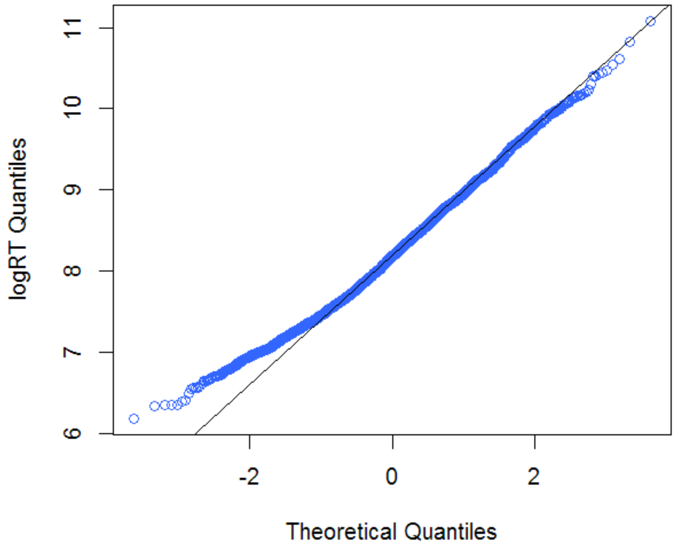

Figure 2.6 Q-Q plot of logRT quantiles against theoretical normal quantiles for experiment one

First, log RTs for each participant were averaged over the seven positions A to G for each of the distances 0.05, 0.1 and 0.2. Then an ANOVA was carried out for the factor of position for distances 0.05, 0.1 and 0.2 combined, taking onto account the factor of participant to minimise the effect of individual differences. This revealed a significant effect of position $F(6,420) = 25.18$, $p < .001$.

Secondly, a pairwise (Bonferroni corrected) t-test between positions A to G was carried out for this data set. The significance results are shown in table 2.1.

	A	B	C	D	E	F
B	< .001	-	-	-	-	-
C	1	.094	-	-	-	-
D	.259	.185	1	-	-	-
E	.047	.743	1	1	-	-
F	.122	1	1	1	1	-
G	< .001	< .001	< .001	< .001	< .001	< .001

Table 2.1 Significance results for pairwise t-test between positions A to G for distances 0.05, 0.1 and 0.2 for experiment one

This test revealed that position G (the larger number above 0.9) stood out as being significantly different from the other positions with position A (the smaller number below 0.1) being significantly different from positions B (where both stimuli are between 0.1 and 0.4) and position E (where the smaller number is between 0.5 and 0.6). Other positions did not differ significantly in terms of mean log RT.

Due to these findings, the factor of position was re-coded to a new factor called location at three levels. These three levels being:

near _zero the smaller stimulus below 0.1;

near_one the larger stimulus above 0.9;

middle all other positions.

This re-coding was easily applied to distance 0.5 for which the re-coding matched the original definitions of the K,L and M levels of the position factor.

More problematic were distances 0.3 and 0.4. The original definitions of the H,I and J levels of the position factor did not match the definitions of the new near_zero, middle and near_one levels of the location factor. For distance 0.4 it was found, however, that the stimuli chosen for levels H,I & J also fit the definitions for levels near_zero, middle and near_one respectively. Stimuli for distance 0.3 did not transfer so neatly to the new location factor. In the re-coding, three of the four position H stimuli fell into the middle location as did two of the four position J stimuli.

Despite the slight unbalancing of the data caused by this re-coding, it was decided to perform analysis on the full data set as long as this did not conflict with the requirements of the analytical methods used.

2.4.2.4 Mixed linear modelling for logRT

Initial testing of the null model (for logRTs) with no predictors, against a baseline model including random intercepts for participants, showed significant individual differences (L Ratio 1151, $p < .001$). Hence I decided that the use of a mixed linear modelling approach which accounted for individual's baseline speed of response would be the most appropriate method of analysis of the RT data. Averaging across individuals and performing a standard analysis of variance would have been an easier but less rigorous approach.

Therefore, using the R statistical package nlme, a mixed linear model was applied to the logRT data with subject as the random effect. The maximum likelihood (ML) method of estimation was utilised to allow for the calculation of likelihood ratios as the model was built up (see Field, Miles & Field 2012). There were four potential fixed factors: distance, location, left or right position of the larger stimulus (*largelr*), included to detect a SNARC effect and whether the larger stimulus was a fraction or decimal (*largerstim*). The mixed linear modelling method can accommodate missing data at some combinations of factor-levels.

2.4.2.4.1 Individual testing of potential fixed factors

First, the four possible fixed factors were added separately to the random intercepts only (baseline) model. An ANOVA test was then applied to detect improvements of each individual factor on its own to the fit of the model (see table 2.2). The factors of distance and location, as predicted, made a significant improvement to the baseline model but the factor largelr did not. Thus no SNARC effect was detected and the factor largelr was excluded from any further analysis. Incidentally, this exclusion improved the balance of the design. It left only one combination of levels of the significant factors (distance: 0.3, location: near_zero, largerstim: decimal) with no data.

Factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Baseline	3	6785.7	6804.2	-3389.8		
Distance	4	6390.6	6415.4	-3191.3	397.1	<.001
Location	5	6464.1	6495.0	-3227.0	325.6	<.001
Largerstim	4	6711.2	6736.0	-3351.6	76.46	<.001
Largelr	4	6786.0	6810.7	-3389.0	1.70	.192

Table 2.2 Results of ANOVA comparisons between baseline linear model and linear models including single factors for experiment one

A somewhat surprising result was the significant improvement made to the model fit by the inclusion of largerstim as a fixed effect. That implied logRTs differed significantly when the larger of the stimulus pair was a fraction from when it was a decimal. However, with factors added to the model in isolation, it was possible that this outcome was just an artefact resulting from the other factors and the stimuli chosen.

2.4.2.4.2 Building the model – single factors

The mixed linear model was then built up in stages by adding the fixed factors in turn. An ANOVA test was applied to test for an improvement in the model (see table 2.3). The addition of location and largerstim to the distance only model significantly improved the fit of the model.

Additional factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Distance	4	6390.6	6415.4	-3191.3		
Location	6	6189.5	6226.7	-3088.8	205.1	<.001
Largerstim	7	6103.3	6146.6	-3044.6	88.2	<.001

Table 2.3 Results of ANOVA comparisons between versions of the linear model as single factors are added for experiment one

A summary of the mixed effects model for the single factors can be seen in table 2.4.

Factor/level	b (95% CI)	SE	df	t-value	p
distance	-1.122 (-1.249, -0.994)	0.065	3566	-17.20	<.001
location: near_zero → middle	0.140 (0.090, 0.190)	0.025	3566	5.50	<.001
location: middle → near_one	-0.223 (-0.283, -0.162)	0.031	3566	-7.22	<.001
largerstim: decimal → fraction	-0.175 (-0.211, -0.138)	0.018	3566	-9.45	<.001

Table 2.4 Summary of the linear model including all significant single factors for experiment one

Within this linear model, the b-value for distance implies that for every **0.1** increase in distance, there was a reduction of 0.112 in logRT (or a 10.6% reduction in RT). This distance effect was significant (p < .001).

There were significant (p < .001) changes in average logRT with the shift in stimulus location between zero and one. The b-value (0.140) between locations near_zero and middle was smaller in size than that between middle and near_one (-0.223). In general, location near_one response times were fastest and middle response times were slowest. These were, on average, around 15.0% longer than those for the near_zero location and 25.0% longer than those for the near_one location. Though this effect was significant, it was small in size.

The significant (p < .001) effect found for largerstim had a b-value of -0.175. This implies that RTs for stimulus pairs in which the decimal was the larger number were, on average, around 16.1% longer than for those in which the fraction was the larger number.

2.4.2.4.3 Building the model – interactions

Interactions were added to the model to see if they would improve the fit to the logRT data. The addition of an interaction between distance and location significantly improved the fit of the model. The addition of an interaction between distance and largerstim made a marginally significant improvement to the model. However, the addition of an interaction between location and largerstim also significantly improved the model. Therefore the three-factor interaction between distance, location and largerstim was added to the model and again, a significant improvement was found. The results of the ANOVA comparisons can be seen in table 2.5.

A visual inspection of the residuals of the final linear mixed model for logRT (including all interactions between distance, location and largerstim) indicated no obvious deviation from the assumptions of normality and homoscedasticity (see figure 2.7).

Interaction added	df	AIC	BIC	Log Likelihood	L.Ratio	p
No Interaction	7	6103.3	6146.6	-3044.6		
distance/location	9	6092.3	6148.0	-3037.1	15.03	<.001
distance/largerstim	10	6090.5	6152.4	-3035.3	3.75	.053
location/largerstim	12	6080.4	6154.6	-3028.2	14.15	<.001
distance/location/largerstim	14	6076.8	6163.4	-3024.4	7.60	.022

Table 2.5 Results of ANOVA comparisons between versions of the linear model as interactions are added for experiment one

The fixed effects of the model including all interactions are summarised in table 2.6. It can be seen that the inclusion of the interactions in the model increased the size of the b-value for distance to -1.720. In this model, distance was the largest of the significant effects on logRT.

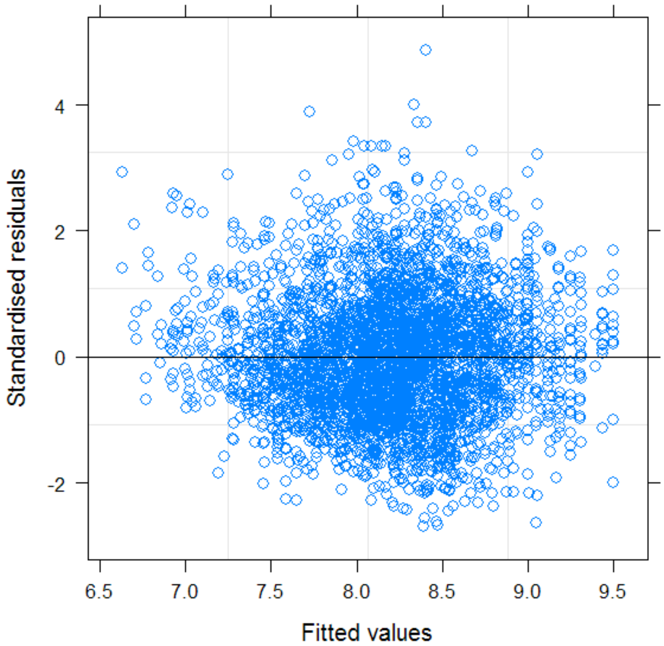

Figure 2.7 Residual plot for the final mixed linear model for experiment one

The inclusion of the interactions greatly reduced the significance of the increase in logRT between locations near_zero and middle, (p = .175). It was only within the three-factor interaction that the shift between locations near_zero and middle had a significant effect (p = .017, b = -0.760) on logRT. The three-factor interaction between distance, location and largerstim was complex and is investigated in detail later in the analysis.

The transition between locations middle and near_one continued to be highly significant, (p < .001) showing an average fall in logRT of 0.372. This effect was significantly mediated by distance (p < .001, b = 1.056), with the logRT difference between locations near_zero and middle reducing as distance increased. In addition, the shift between locations middle and near_one was also significantly affected within the three-factor interaction (*p* = .012, *b* = -0.917).

single factors	b (95% CI)	SE	df	t-value	p
distance	-1.720 (-2.080, -1.360)	0.184	3559	-9.351	<.001
location: near_zero → middle	0.085 (-0.038, 0.207)	0.062	3559	1.356	.175
location: middle → near_one	-0.372 (-0.527, -0.218)	0.079	3559	-4.719	<.001
largerstim: decimal → fraction	-0.28 (-0.434, -0.125)	0.079	3559	-3.541	<.001
two-factor interaction					
dist/location near_zero → middle	0.401 (-0.039, 0.841)	0.225	3559	1.783	.075
dist/location middle → near_one	1.056 (0.547, 1.564)	0.260	3559	4.065	<.001
distance/largerstim decimal → fraction	0.874 (0.366, 1.383)	0.260	3559	3.367	<.001
L.stim fraction/loc. near_zero → middle	0.086 (-0.087, 0.258)	0.088	3559	0.971	.332
L.stim fraction/loc. middle → near_one	0.003 (-0.215, 0.221)	0.111	3559	0.028	.977
three-factor interaction					
dist/loc near_zero → middle/largerstim decimal → fraction	-0.760 (-1.383, -0.137)	0.318	3559	-2.389	.017
dist/loc middle → near_one/largerstim decimal → fraction	-0.917 (-1.635, -0.199)	0.367	3559	-2.499	.012

Table 2.6 Summary of the linear model including all interactions for experiment one

With the inclusion of the interactions in the model, the single factor of largerstim continued to be significant, (*p* <.001) and with an increased (negative) b-value of -0.279. On average, responses were significantly faster when the larger of the stimulus pair was a fraction than when it was a decimal. This effect had a significant (*p* <.001) interaction with distance with *b* = 0.874.

Both significant two-factor interactions included the distance factor. I found that by picking apart the three-factor interaction I could clarify the nature of these interactions better.

2.4.2.4.4 Three-factor interaction

Figure 2.8 shows the three-factor interaction between distance between stimulus pairs, location of stimulus on the zero to one scale and whether the larger of the stimulus pair is a fraction or a decimal. The missing data for the combination of factor levels distance: 0.3, location: near_zero, largerstim: decimal is apparent.

It can be observed that as distance increased, the interaction between largerstim and location changed. Most noticeable was that for the larger distances of 0.4 and 0.5, there was a much smaller difference between logRTs for near_zero and middle located stimuli when the fraction was the larger of the stimulus pair than when the decimal was the larger. Fraction-larger stimuli actually having longer response latencies than decimal-larger ones when the stimulus pairs were in the near_zero location for the two largest distances. Fraction larger stimuli being consistently faster than decimal larger stimuli, on average, across all other distance and location pairings for which data was recorded.

Though, as seen earlier, (figure 2.3), there appears to have been a consistent, if not necessarily entirely linear, distance effect across all stimuli; the three-factor interaction graph (figure 2.8) demonstrates that this distance effect was not consistent across any combination of the factors location and largerstim. In particular, it was only the distances 0.05, 0.1 and 0.2 that a decrease in response latencies was consistently associated with an increase in distance between stimulus pairs for all combinations of the factors location and largerstim.

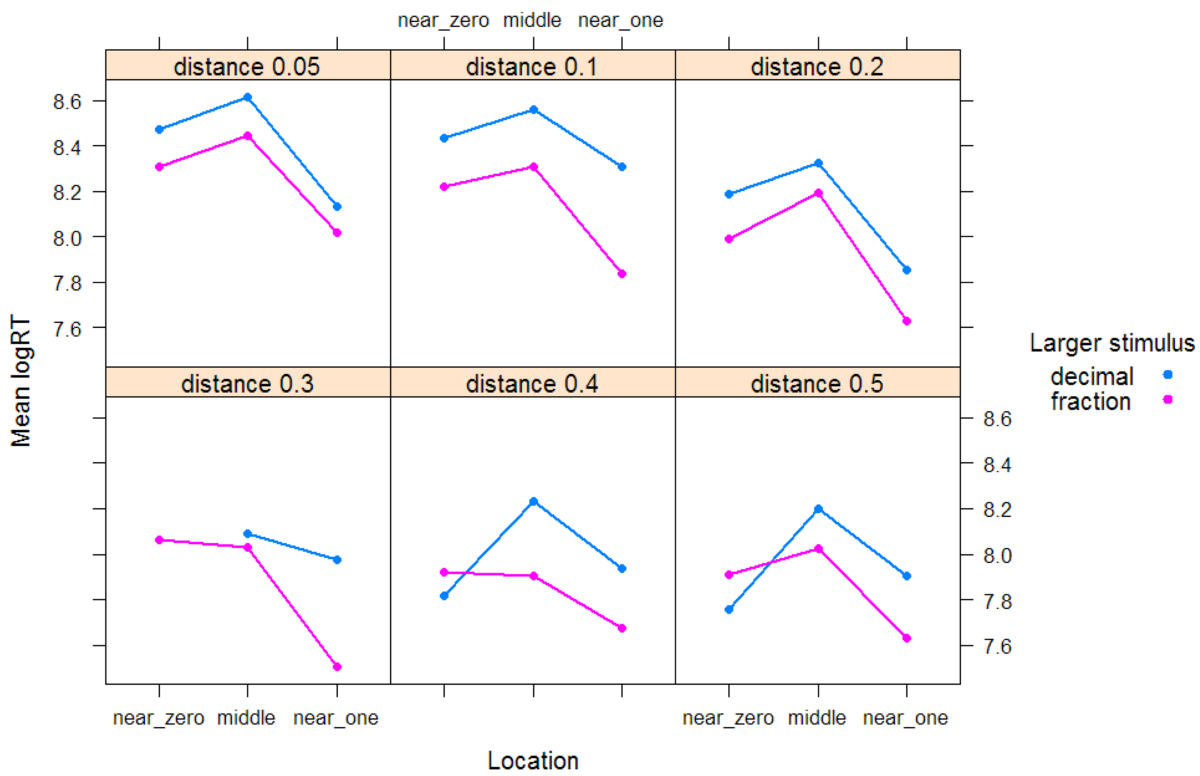

Figure 2.8 Lattice graph showing three-way interaction between factors distance, location and largerstim for experiment one

Because of this change in response patterns as distance increased, I decided it would be appropriate to consider the data for the smaller distances of 0.05, 0.1 and 0.2 separately from the larger distances of 0.3, 0.4 and 0.5.

2.4.2.4.5 Smaller distances (≤ 20)

For the smaller distances, the factors of distance, location and largerstim all made significant improvements to the model (all $p < .001$). No interactions made a significant improvement to the model (all $p > .398$). Again, the (SNARC) factor of whether the larger stimulus was on the right or the left did not significantly improve the model (p = .288).

Hence, the final model for the smaller distances contains only the three factors distance, location and largerstim with no interactions. For this model, the significant ($p < .001$) effect of distance was b = -1.921 (95%CI -2.267, -1.574) , SE = 0.177 indicating a reduction of 0.192 in logRT, (or a 17.5% decrease in RT), for each 0.1 increase in distance.

From location near_zero to location middle, there was a significant ($p < .001$) average increase in logRT of b = 0.139 (95%CI 0.076, 0.202) , SE = 0.032. This value of b indicates a 14.9% increase in RT. The significant reduction in logRT from the middle location to the near_one location indicates a drop in logRT of b = 0.306 (95%CI -0.386, -0.225) , SE = 0.041. This is equivalently a 26.4% drop in RT between the two locations.

Stimuli in which the decimal was the larger of the pair were significantly (p < .001) smaller in logRT by b = -0.198 (95%CI -0.241, -0.155) , SE = 0.022 (or 21.9% faster in RT).

2.4.2.4.6 Larger distances (≥ 30)

In building up the mixed linear model for the larger distances, the factors of distance (p = .008), location ($p < .001$) and largerstim ($p < .001$) all made significant improvements to the model. Again, the (SNARC) factor of whether the larger stimulus was on the right or the left did not significantly improve the model (p = .762). The interactions of distance with location (p = .040) and location with largerstim ($p < .001$) both made a significant improvement to the model.

In the final model for the larger distances which included interactions, there was no significant individual factor effect of distance (p = .161) or of location (p = .872 for near_zero to middle and p = .507 for middle to near_one).

However, the largerstim effect persisted. Stimulus pairs in which the fraction was larger had significantly (p = .029) lower logRTs than those in which the decimal is larger b = 0.140 (95%CI 0.015, 0.265), SE = 0.064 (or 13.1% faster in RT).

The interaction between location (near_zero to middle only) and distance was marginally significant in the final model, *p* = .086. However, the 95%CI for *b* was (-0.134, 2.016) which does not indicate a reliable effect. This result could be attributable to the paucity of near_zero data for the distance 0.3.

However the factor largerstim interacts significantly (*p* < .001) with the shift in location between near_zero and middle, *b* = -0.293 (95%CI -0.446, -0.140), *SE* = 0.078 and with the shift in location from middle to near_one (*p* < .001), *b* = -0.447 (95%CI -0.617, -0.277), *SE* = 0.087. This interaction is apparent on figure 2.8. Particularly notable is that for distances 0.4 and 0.5, stimuli in which the larger stimulus is a fraction had greater logmRT than those for which the larger stimulus is a decimal at the near_zero position – a reversal of the generally observable pattern.

In summary, at the larger distances of 0.3, 0.4 and 0.5, the largerstim effect persisted. However, the effect of distance did not. The effect of location is only apparent as an interaction with the largerstim factor; the shape of which can be seen in figure 2.8.

2.4.3 Error analysis

To identify which factors had a significant influence on errors, a logistic (binomial) regression analysis was applied to the error responses, (0 for a correct response and 1 for an incorrect response). The factor of participant was included in the analysis but results for individuals are not reported here. The coefficient and odds ratio results can be seen in table 2.7.

Source	b	SE	odds ratio (95% CI)	z value	p
distance	-4.996	0.494	0.007 (0.003, 0.018)	-10.12	<.001
location: near_zero → middle	0.753	0.174	2.123 (1.51, 2.984)	4.33	<.001
location: middle → near_one	0.264	0.218	1.302 (0.849, 1.996)	1.21	.227
largerstim: decimal → fraction	-0.424	0.106	0.655 (0.532, 0.806)	-4.00	<.001
largelr: left → right	0.072	0.105	1.074 (0.874, 1.32)	0.68	.495

Table 2.7 Results of a logistic regression analysis of error data for experiment one

Individual differences between participants had a significant effect on the number of errors ($\chi^2(29) = 227$, *p* <.001).

the percentages of incorrect responses for each distance, location and larger stimulus type can be seen in Table 2.8.

Again, there was no effect due to whether the larger stimulus was on the right or left ($\chi^2(1) = 0.47$, p =.495). However, participants were significantly more likely to make an error if the larger stimulus was a decimal than if it was a fraction ($\chi^2(1) = 16.16$, p <.001). Errors were 1.52 (1/0.655) times as likely to occur if the larger stimulus was a decimal rather than a fraction all other factors being equal. This reflects the result found for the effect of the largerstim factor on RTs and further reinforces the observation that participants found it more difficult to decide which of the stimuli was the larger when the larger one was a decimal rather than a fraction.

distance	0.05	0.1	0.2	0.3	0.4	0.5
% Errors	23.0%	18.2%	9.2%	5.8%	5.8%	4.7%

location	near_zero	middle	near_one
% Errors	6.7%	17.2%	8.8%

larger stimulus	decimal	fraction
% Errors	15.6%	11.2%

Table 2.8 Tables showing percentage of errors at levels of distance, location and largerstim factors for experiment one

Location also had a significant effect on whether an error was made ($\chi^2(2) = 27.91$, p <.001). This effect was specifically applicable to the shift from near_zero to middle locations. An error was a little over twice (2.123) as likely to occur if the stimulus was in the middle location than if it was in the near_zero position, all other factors being equal. Table 2.8 indicates that, indeed, participants were, in general, much more likely to make an error when a stimulus was in the middle position.

Distance had a highly significant influence on errors ($\chi^2(1) = 157.89$, p <.001). The odds ratio here is difficult to interpret in any very meaningful way but the fact that it is so small is implying a rapid dropping off of errors as distance increased. Table 2.8 shows that the percentage of error responses decreased rapidly between distance 0.05 and 0.2 but then almost levelled out between distances 0.3 and 0.5.

2.4.4 Verbal report of strategy

Once participants had completed all 120 stimuli they were asked to give a verbal report of their strategy with reference to three further stimulus pairs, if required. These strategies were studied for general themes and categorised into ten expressed methods (coded 1 to 10) plus three further non-method categories (coded 11, X and 0), as detailed in table 2.9. If participants mentioned several methods, all were recorded. Some reported using no method or a faulty method as well as legitimate method(s). In these cases, both the method and no method or faulty method responses were recorded. The recordings made of two of the participants became corrupted before analysis was carried out. So no data for those participants are available. The vast majority of participants, (23 of the 29 recorded), reported using two or more strategies.

Strategy	Number of participants
1. Use 0.5 or ½ as an anchor point.	20
2. Use other fractional anchor points.	13
3. Use 0 as an anchor point.	2
4. Use 1 as an anchor point.	1
5. Convert the decimal to a fraction.	3
6. Convert the fraction to a decimal.	6
7. Convert both (into e.g. percentages/tenths).	4
8. Partition of the whole (e.g. mental shading)	3
9. Multiplication/division to/from a whole.	5
10.Abstract method of multiplication/division.	1
11. Just know.	3
X. Faulty/incoherent reasoning.	3
0. No method.	3

Table 2.9 Summary of reported strategies for experiment one

There are several interesting points to note in relation to these expressed strategies and participants' responses to the number comparison task.

Firstly, the one person who reported using a solely abstract method of multiplication and division with no reference to the size or position or physical magnitude of the numbers (strategy 10) was the one participant who was unable to complete the task. Upon further questioning they said that the task was too difficult as their abstract strategy was too much to hold in their head and they knew no other strategy. This responses of this participant are an interesting indication that understanding small numbers only in an abstract sense is not sufficient for successfully using them in a flexible manner. However, they constitute a very small sample.

Secondly, the majority of participants (20 out of 31) reported making judgements by comparing the two stimulus numbers to 0.5 or one half. If just the smaller distances are considered, only position D had stimuli either side of 0.5. However, there was no evidence that participants were quicker in general in this position. On the contrary, only one of them was quickest in position D for the smaller distances. In general, it was shown in section 2.4.2.3 that position D was not significantly different in terms of mean logRT from any of the other positions but G (closest to one), which was, in fact, significantly faster. This implies that the participants were not using a comparison with 0.5 as the first choice strategy to make their judgements despite their verbal reports.

Lastly, of the three participants who gave faulty or incoherent reasoning strategies, one was the most erroneous participant (45 errors out of 120). However, one of them made only 4 errors and the other only 22 errors (ranked 25 of 30 participants). This demonstrates that an inability to verbalise their strategy was not necessarily indicative of an inability to complete the task successfully.

These two last points imply that whatever strategies participants were using, they were not necessarily consciously aware of them. As such, there is perhaps, little value in these verbal reports for drawing conclusions about the actual strategies that had been used.

Nevertheless, they might be indicative of the nature of the conscious schema of fractional numbers held by the participants. The non-abstract strategies (1 to 9) can be categorised in three ways. Strategies 1 to 4 are number line strategies that involve the comparison of the stimuli to anchor points on a number line. Whereas strategies 5 to 7 involve the use of direct conversion between different representations. Finally, strategies 8 and 9 involve relating the part (stimulus) to the "whole" with the whole not being specified as a *number line* between zero and one.

Of the 29 participants whose recordings were analysed, 22 of them stated using number line strategies (1 to 4), often in combination with other methods. Only one of the 13 participants who reported using a conversion strategy (5, 6 or 7) did not also report using a number line strategy. That was, again, the most erroneous participant. Upon further explanation, it became apparent that their conversion strategy was actually mathematically faulty, despite being correct in intent. This implies that most of the participants were aware of the spatial analogue for the ordering of the size of fractions and decimals. That is, they were aware that the mental number line concept can be extended to fractional numbers of different external representations.

For four participants the only correct reported strategy was a comparison to the whole method (8 or 9). Thinking of fractional numbers as parts of the whole is a useful and correct concept and strategies 8 and 9 are effective methods of completing the number comparison task. However, they do not relate to extending the number line (linear spatial analogue) for whole numbers to fractional numbers. Nevertheless, two of these four participants did demonstrate a significant distance effect.

2.4.5 Summary of results

2.4.5.1 Response times

The RT results were noisy with a great deal of the variance being accounted for by the random effect of individual difference.

Three fixed effects were found to have significant effects on RTs. Firstly the classic distance effect for number comparison tasks. This effect was sought by the experiment and expected to occur. Secondly, the location effect in which RTs were affected by the location of stimulus pairs in the zero-to-one interval. This effect was not entirely unexpected. Thirdly, the entirely unexpected largerstim effect in which whether the larger, (or indeed smaller), stimulus was a decimal, (or fraction), affected the RT for that stimulus pair.

At the distances of 0.05, 0.1 and 0.2, there was a significant classic distance effect with RTs increasing as the distance between stimulus pairs decreased. This was by far the greatest effect on RTs found. Also, the effect of location was significant with RTs in the middle location being largest and those in the near_one location being smallest. Finally, also for these smallest distances, the largerstim effect was significant. That is, stimulus pairs in which the larger was a decimal had larger average RTs than those in which the larger stimulus was a fraction. There were no significant interactions between these effects.

At the larger distances of 0.3, 0.4 and 0.5, the largerstim effect persisted. However, the effect of distance did not. There were no significant differences in RTs at these distances. Though those at distance 0.3 might have been smaller than those at 0.2. The effect of location was only apparent as an interaction with the largerstim factor. Specifically, the *difference* between RTs in the near_zero and middle locations was greater when the larger stimulus was a decimal than when it was a fraction.

There was no effect on response times of whether the larger stimulus was (congruently) on the right or (incongruently) on the left. Thus there was no evidence of an unconscious spatial-numerical association between size, physical location and response key location (SNARC effect).

2.4.5.2 Errors

The results for the error analysis showed a similar story to those of the RT analysis. That is, firstly a significant distance effect was present, at least over the distances of 0.05 to

0.3. That is, error rates decreased as distance increased but this effect tailed-off for the larger distances. Distance between stimuli was the most significant influence on error rates.

Error rates also demonstrated a significant location effect. They were significantly higher for the middle location than the near_zero and near_one locations. In addition, errors were significantly more likely to occur when the larger of the stimulus pair was a decimal than when it was a fraction. That is the same largerstim effect as was found for RTs.

2.4.5.3 Reported strategies

The reported strategies of the majority of participants implied that they were consciously aware of the number line analogue for fractional number magnitude and that they employed this knowledge in completing the number comparison task.

Most participants reported comparing magnitudes by using the *anchor points* of zero, 0.5 and one. Though the location effect found within the RTs and error analysis supports the use of zero and one as anchors, there is no pattern of responses supporting the use of 0.5 as an anchor point against which judgements were made.

2.5 Discussion

One of the most remarkable outcomes of this experiment was just how few errors participants made. I had made the task deliberately difficult yet there was an error rate of only 22.9% at the very small distance of 0.05. It can also be noted that the only participant who reported using a completely abstract strategy, that made no reference to magnitude related features of the numbers, could not complete the task. Taking these two outcomes into account seriously challenges the assertion of Cohen et al. (2002) that there is no common mental magnitude representation for fractions and decimals. This common, ordered, magnitude representation may be approximate and imprecise in nature but it seems to be the only mechanism by which participants in this experiment could have been so successful a making their judgements.

Evidence of participants accessing a mental number line for fractional numbers was sought in two ways. First, the distance effect was tested for. Additionally, the stimulus were chosen so as to allow for the detection of a Stroop-like SNARC effect. The second of these is considered first.

2.5.1 SNARC effect

The factor of largelr was included in this experiment in order to detect any evidence of a SNARC effect. Largelr had virtually no detectable effect upon RTs. The SNARC effect is an unconscious effect indicating that magnitudes are accessed swiftly and unconsciously associated with a left-right spatial magnitude ordering. So the fact that this effect was not found implies either that magnitudes were not swiftly accessed or that they were not associated with a left-right spatial magnitude ordering.

In the experiments of Dehaene et al. (1993) which first highlighted the existence of a SNARC effect for whole numbers, average RTs that were between 460 and 530ms. These were for quite a different comparison task than my experiment one. Nevertheless, even at the largest distance between stimuli of 0.5, the median RT for experiment one was around 2500ms – almost five times as long. Hence it is reasonable to assume that the comparison procedure taking place that was deliberative and too lengthy to be affected by an unconscious effect such as the SNARC effect. This does not mean that the magnitudes of fractional numbers are not accessible, just that they are not automatically accessed.

Indeed, Gabriel, Szucs & Content (2013a) and (2013b) found evidence that holistic magnitude of fractions were accessed for numerical tasks but not for tasks in which the magnitude of the number was not relevant. So the lack of automatic access to fraction magnitudes does not mean that internal representations are not akin to a mental number line.

2.5.2 Distance effect

A strong and unequivocal distance effect for RTs (and errors) was found. Indeed, the effect was not dissimilar to that found by Tversky & Kahneman in 1974. The way this experiment was set up, however allowed for more in-depth analysis of the nature and limits of the effect. The distance effect was most clear for the stimuli with smaller distances between them. For these it was by far the strongest effect found. For larger distances, the effect was less concrete.

This shape of the distance effect found in experiment one (figure 2.3) is very similar to the pattern of distance effect found by Schneider & Siegler (2010) in their second target-stimulus number comparison task. That task included comparisons between pairs of fractions some of which had double-digit denominators. They too recorded a sharp fall-off in RTs as the stimuli initially moved away from the target (i.e. at smaller distances) with that decrease in RTs becoming smaller as distance from the target increased. Schneider & Siegler interpreted their result as evidence of holistic magnitude processing of their stimuli.

Conventionally, indeed, the distance effect has been thus understood as an implication that holistic magnitudes representations of both stimuli are used to make comparisons. Cohen (2010) made a challenge to this conventional interpretation. He showed that any distance effect found in his magnitude comparison task (for single digits) could be better explained by the visual similarity of the stimulus pairs than the distance between stimuli.

The stimulus pairs chosen for my experiment one were made up of one decimal with a zero & a dot followed by three digits and one fraction with one or two digits above a line with two digits below. It would be difficult to construct a meaningful way to codify the similarity between these pairs of numbers. As such, I have not demonstrated that physical similarity was not a better predictor of RTs than distance. However, it seems more plausible that participants in my experiment were translating the two numbers into (approximate) magnitude analogues for comparison and the distance effect seen was an outcome of this translation. The stimuli were chosen specifically to make it difficult for participants to do otherwise.

The diminishing of the distance effect for larger distances raises the question of whether there is a limit on the distance effect for fractional number pairs. These numbers are on a bounded continuum with zero at one and one at the other. They can only get so far apart before they hit the ends of their limits. There *was* an effect on response times of these limits, especially at the smaller distances however. I termed that phenomenon the *location effect*.

2.5.3 Location effect

In contrast to other studies of the distance effect, by controlling the distance between stimuli I was able to investigate the influence of location within the zero-to-one range on the magnitude comparison task. The main feature of the location effect found was that, at least for the smaller distances of 0.05, 0.1 and 0.2, RTs were fastest when stimuli were within 0.1 of one and next fastest when the stimuli were within 0.1 of zero and slowest when they located in the other *middle* positions.
This is not the *size effect* reported for whole numbers by Moyer & Landauer (1967); Dehaene et al. (1990); Nuerk et al. (2001). RTs did not increase as the size of stimulus pairs increased (the distance remaining the same). Instead more of an *anchor point*

effect was found. Perhaps implying that magnitudes for fractional numbers between zero and one are automatically estimated from these two anchoring points.

There was no interaction between location and distance for the smaller distances and at larger distances it was questionable also. Taking this into account, the diminution of the distance effect for larger distances might not, in fact be down to the fact that both stimuli are moving closer to the end points of zero and one as the distance between them increases.

2.5.4 Largerstim effect

The final and completely unexpected main outcome of experiment one was that RTs (and errors) were found to be significantly greater when the larger of the two stimuli presented was a decimal than when it was a fraction. I termed this outcome the *largerstim effect*.

The effect was present at all distances, though less clear-cut at the greatest distances. Also, for the smaller distances at least, the largerstim effect was consistent at all locations of the stimuli in the zero-to-one range.

Not only was this outcome unexpected it is also difficult to interpret. The only simple conclusion that can sensibly be drawn is that it indicates a difference between participants' mental representations of decimals and fractions and/or how participants processed the magnitude of fractions and decimals.

It was not completely unexpected to find evidence of such a difference. Fractions and decimals look different and though, mathematically they can represent the same types of phenomena they are generally used in different ways; fractions to represent relative frequency and decimals to represent partial quantities, for example. What was unexpected, was the nature of the difference found.

Another way of stating the largerstim effect would be that RTs times were longer (and error rates were greater) when comparing a *larger* decimal with a *smaller* fraction than when comparing a smaller decimal to a larger fraction. This was an effect that indicated not only a difference between fraction and decimals also their *comparative size*.

It is therefore possible that the instruction to choose specifically the *larger* of the two stimuli is triggering some procedure of comparison that is quicker when the decimal is smaller and the fraction is larger. Perhaps changing the response that participants have to give from choosing the larger to choosing the smaller would reverse this largerstim effect.

2.5.5 Reported strategies

Most participants reported the use of more than one strategy. In their fraction comparison experiment, Faulkenberry & Pierce (2011) similarly found that participants often used multiple strategies that were dependent on the types of fraction being compared. They had participants report strategies for every trial and so obtained much richer data on strategies than I. They encoded the strategies differently but they were generally very similar to those I recorded. However, in the light of the location effect, it is interesting that they did not report any participants using a comparison to zero or one strategy. I had one participant reporting using both and another reporting using zero only.

The reports I recorded probably reflect only the most salient strategies that participants used. This could contribute to some explanation of why so many people said they made their comparisons on whether numbers were either side of 0.5 despite the fact that the minority of comparisons could be made that way. Also, there was evidence to suggest that comparisons that *could* be made that way actually took longer. (More consideration of this point follows the results of experiments three and four.)

The fact that the great majority of participants reported strategies based on correct holistic magnitude reasoning is supportive of my conclusion that the distance effect found does reflect holistic magnitude comparison.

2.5.6 Next steps

The next chapter summarises the second experiment carried out for the purposes of this thesis. Experiment two was intended to investigate a possible reason for the largerstim effect, further investigate the limits of the distance effect for fractional numbers and find out whether the location effect could be replicated.

Chapter 3 Experiment two

3.1 Introduction

This third chapter details the second of the experiments carried out for the purpose of this thesis. Experiment two was an extension of experiment one. This chapter starts with a summary of the justification of the experiment. The changes to the design of the experimental task and stimuli are then explained. In particular, how these changes were intended to further investigate and clarify the results of experiment one.

The next part of this chapter covers the method of experiment two. Then the following section contains the results of the experiment along with the methodology of the analysis.

Finally, the last section of the chapter is a consideration of the implications of the results of the experiment and areas for further investigation.

3.2 Justification and design

The second experiment detailed in this chapter was essentially an extension of the first experiment. This consisted of two key modifications. The intention was to find out both whether the results of experiment one could be replicated/extended and to test for possible explanations of these results.

Experiment one sought to find evidence of common mental magnitude representations of fractions and decimals. Evidence was indeed found in the form of a significant distance effect in the magnitude comparison task. However, this distance effect, (decrease in RTs and errors associated with an increase in distance between stimulus pairs), appeared to diminish as distance increased. Between the two largest distances, 0.4 and 0.5, there appeared to be no difference in average RT or errors.

Therefore, the first modification to the experiment was to increase the range of distances between the two stimuli. Specifically, two larger distances between stimulus pairs were included. This was intended to investigate whether the diminishing of the distance effect between distances 0.4 and 0.5 would continue for larger distances.

As the stimuli were numbers between zero and one, a gap of only one unit, there was a limit to the practical increase in distance between the stimuli. In addition, in experiment one, the factor of location appeared to have an effect on RTs and errors. Thus it was deemed important to balance the new stimulus pairs for the three levels of the location factor – near_zero (smallest number between 0 and 0.1), near_one (largest number between 0.9 and 1) and middle (both numbers between 0.1 and 0.9).

These considerations meant that the two additional distances included were 0.6 and 0.7. Twenty-four new and distinct stimulus pairs were created in the same manner as the stimuli of experiment one. The 24 new stimulus pairs encompassed every combination of the two new distances, the three levels of the location factor, the two levels of the largerstim factor and the two levels of the largerlr factor. This was done to allowed for consistency and replication within the experiment, despite the fact that the factor largelr did not demonstrate any effect on mRT in experiment one.

The second modification was included as an attempt to explain the unexpected largerstim effect observed in experiment one. This is the effect that average RTs and error rates were significantly greater when the larger stimulus of the pair was a decimal than when it was a fraction. As discussed in the previous chapter, this outcome could possibly have been caused by the response that participants were asked to give during the experiment. They were asked to choose specifically the *larger* of the two numbers presented.

Therefore, the second modification to the experiment was to include an additional between participants factor of *response* with two conditions. One group of participants was asked to respond by identifying the larger number of the stimulus pair, as in experiment one (the *response-larger* condition) and the other group was asked to respond by identifying the smaller number (the *response-smaller* condition). If the participants in the two groups demonstrated different largerstim effects, this might help to explain the causes of the largerstim effect.

The inclusion of this additional response factor also allowed for some investigation into the slight asymmetry of the location effect seen in experiment one. Stimulus pairs in the *middle* location certainly had significantly longer RTs. However, it was also observed for the smaller distances of 0.05, 0.1 and 0.2, that those the near_one location had smaller RTs (but slightly more errors) than those in the near_zero location. This slight asymmetry might have been caused by attention being drawn to the top of the zero-to-one range by the requirement to choose the larger stimulus.

Because of the second modification to the experiment, at least twice as many participants were required to take part in the experiment. It was therefore considered impractical to interview each participant verbally about their strategy.

3.3 Method

3.3.1 Participants

Fifty-eight staff and students at the University of Huddersfield initially participated in the second experiment. However, subsequent analysis showed that several participants performed no better than chance and were therefore rejected from the analysis; as detailed in the results section. Thus a second session of the same experiment was run involving nine further participants, all students at the University of Huddersfield. Equivalent conditions and the same software and equipment were used during both sessions. There were sixty-seven participants in total, all of whom completed the task.

3.3.2 Stimuli

The 120 stimuli used in Experiment 1 were again used in Experiment 2. An additional 24 stimuli were created, 12 each at distances 0.6 and 0.7. The stimuli again consisted of one fraction presented in its simplest terms with denominator of 11, 13, 15, 17 or 19 and one three-digit decimal fraction. The decimal fractions were generated by adding and subtracting the distances 0.6 and 0.7 from each fraction and rounding to three decimal places, ignoring pairs in which the decimal fell outside of the range 0 to 1.

The 88 possible new stimulus pairs were split into 12 groups; two for each of the 6 combinations of distance (0.6 and 0.7) and location (near-zero, middle, and near-one). This allowed for one in which the decimal was the larger and one in which it was the smaller. Two stimulus pairs were randomly chosen from each of the 12 groups so that, as in the first experiment, one could be presented with the larger number on the right and the other with the larger number on the left.

This resulted in a set of 144 stimuli in total.

3.3.3 Procedure

The experimental procedure was almost identical to that of Experiment 1 with the exception that participants were divided into two conditions. Thirty-three of them had to choose the larger of the two numbers presented, as in Experiment 1 (response-larger condition). The other thirty-four participants had to choose the smaller of the two numbers (response-smaller condition).

The experiment was conducted using SuperLab® 4.0 stimulus presentation software in a quiet, well-lit laboratory at the University of Huddersfield. Participants were instructed that within each trial they had to decide which of the two numbers presented was the largest and to press the leftmost button on the response pad if it was the number on the left and the rightmost button on the response pad if it was the number on the right. They were informed that both speed and accuracy of response were important. In addition, it was made clear that some of the tasks were expected to appear very easy and some extremely difficult and the purpose of the experiment was to find out what factors made the task more difficult.

A practice block of four stimuli preceded the experimental blocks. Participants were given feedback on their accuracy on the practice stimuli and were allowed to ask questions if they did not understand the procedure. The 144 experimental stimuli were then presented in random order in three blocks of 48 with no further feedback on accuracy nor opportunity to ask questions.
Participants were given the opportunity to take a break between the blocks. All participants were presented with the same stimuli.

Response times and accuracy were recorded by the program.

3.4 Results

3.4.1 Pre-analysis data processing

One RT of only 1ms and another of 97s, (twice as long as the next longest RT), were considered extreme outliers and replaced with the next smallest RT and next largest RT of the participants concerned respectively.

The number of errors per candidate ranged between 3 and 89, with a mean average of 33.5. Nine subjects performed no better than chance at the 5% level. That is, they made more than 61 errors out of the 144 trials, B(144, 0.5). It can be speculated that these participants may have misunderstood the task or were not able to, or did not attempt to, complete the task accurately. Whatever the reason for their poor performance, as the purpose of the experiment was to uncover mental representations accessed in completion of the comparison task, the nine subjects were excluded from any further analysis. This left thirty participants in the response-smaller condition and twenty-eight in the response-larger condition.

3.4.2 Response times analysis

3.4.2.1 Skew and transform

Using the method of Crawley (2005), it was found that the RT data were highly positively skewed ($\gamma = 4.023$, $p < .001$). So a log transform was applied to the RTs as for experiment 1. This resulted in much less skewed data ($\gamma = 0.192$, $p = .424$) which a Q-Q plot (figure 3.1) showed to be approximately normal across the majority of the data but with some negative skew caused by 17 of the 8352 responses.

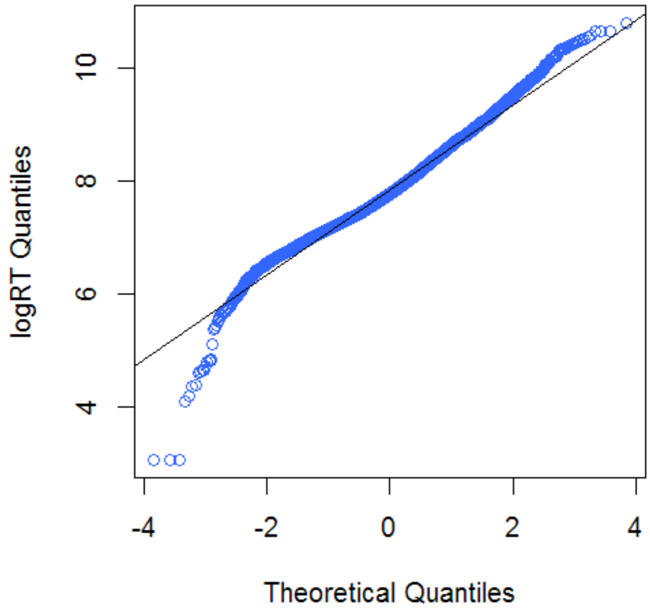

Figure 3.1 Q-Q plot of logRT quantiles against theoretical normal quantiles for experiment two

3.4.2.2 Position factor

Log RTs for each participant were averaged over the seven positions A to G for each of the distances 0.05, 0.1 and 0.2 and an ANOVA was carried out for the factor of position for distances 0.05, 0.1 and 0.2 combined. Again, to minimise the effect of individual differences, the factor of participant was included. This revealed a significant effect of position $F(6,812) = 23.5$, $p < .001$.

Then a pairwise (Bonferroni corrected) t-test between positions A to G was carried out on these data for each of the response types separately . The significance results for response-larger are shown in table 3.1 and for response-smaller in table 3.2.

Certainly the results of these pairwise t-tests for the response-larger condition mirror those of experiment one in that position G nearest to one stands out as having significantly different mean logRTs from the other positions, except, in this case, position A. Also position A, nearest to zero has significantly different mean logRTs from positions B, D and E. Again, the middle positions of B, C, D, E and F do not have significantly different mean logRTs.

larger	A	B	C	D	E	F
B	<.001	-	-	-	-	-
C	.220	1	-	-	-	-
D	<.001	1	1	-	-	-
E	.012	1	1	1	-	-
F	.254	.504	1	.242	1	-
G	.112	<.001	<.001	<.001	<.001	<.001

Table 3.1 Sig. results for pairwise t-test between positions A to G, distances 0.05, 0.1 and 0.2, response-larger condition for experiment two

smaller	A	B	C	D	E	F
B	<.001	-	-	-	-	-
C	<.001	1	-	-	-	-
D	<.001	1	1	-	-	-
E	<.001	.060	.735	1	-	-
F	<.001	.132	.784	1	1	-
G	.004	1	.359	.028	.007	.020

Table 3.2 Sig. results for pairwise t-test between positions A to G, distances 0.05, 0.1 and 0.2, response-smaller condition for experiment two

For the response-smaller condition, the result is similar but with one key difference. Here it is position A, nearest zero that stands out as having significantly different (in fact smaller) mean logRTs from the other positions. Position G, nearest to one, now is significantly different from only A, D , E and F in terms of mean logRT. The middle positions of B, C, D, E and F do not have significantly different mean logRTs except maybe positions B and E which are marginally significantly different from one-another.

As a result of these observations, the factor of position was again re-coded to the new factor of location at the three levels *near _zero*, *middle* and *near_one* as before. Experiment two was initiated before the problem with the recoding for distance 0.3 was discovered for experiment one. So again, there are the same missing combinations of factors for distance 0.3 only. The new distances of 0.6 and 0.7 were not problematic to recode as the definition of the original positions at these distances matched the new location definitions exactly.

3.4.2.3 Distance effect by stimulus

The mean average RT was calculated for each stimulus pair. This was to allow for testing of the strength of a distance effect against another, non-independent effect as was done in the analysis of experiment one.

Specifically, there was concern that any distance effect found might be due to participants comparing the absolute difference between the numerator of the fraction with the first decimal place of the decimal. Comparing the two numbers in this manner would actually have been a very successful tactic. For all of the stimulus pairs of distance 0.3 or more, this comparison was congruent with a holistic magnitude comparison and would therefore have yielded the same correct answer. At the smaller distances of 0.05, 0.1 and 0.2, 59 of the 84 stimulus pairs (70.2%) could also have been correctly compared thus.

Nevertheless, the association of mRT with distance was far stronger (r_s = -.768, p < .001) than it was with the absolute difference between the numerator of the fraction and the first decimal place of the decimal (r_s = -.428, p < .001).

3.4.2.4 Mixed linear modelling for logRT

An initial test of the null model (for logRTs) with no predictors, against a baseline model including random intercepts for participants, showed significant individual differences (L Ratio 2948, p < .001). This indicated that a mixed linear model approach with participant as a random effect was again a suitable method for the analyses of logRT.

3.4.2.4.1 Hierarchy

The experiment design was between groups of participants that were either in the response-smaller condition or the response-larger condition. Hence the experimental design could be thought of as hierarchical in nature with participants nested within response conditions. The addition of this hierarchy to the random effects did not improve the fit of the model by itself (L Ratio < 1, p = .998). This indicates that there was not a significant difference between the average logRTs of the participants in the two response conditions. However, the hierarchy was included in the model for all further analysis of the full data set to allow for detection of any interaction with other factors.

3.4.2.4.2 Individual testing of potential factors

The five possible fixed factors (response, distance, location, largerstim & largelr) were added separately to the random intercepts only (baseline) model. An ANOVA test was then applied to detect improvements of each individual factor on its own to the fit of the model (see table 3.3).

Factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Baseline	4	16259.8	16288.0	-8125.9		
Response	5	16261.8	16297.0	-8125.9	0.027	.870
Distance	5	15691.0	15726.1	-7840.5	570.9	<.001
Location	6	15879.2	15921.4	-7933.6	384.6	<.001
Largerstim	5	16215.6	16250.8	-8102.8	46.25	<.001
Largelr	5	16258.9	16294.1	-8124.5	2.90	.089

Table 3.3 Results of ANOVA comparisons between baseline linear model and linear models including single factors for experiment two

Response on its own made no significant improvement to the model. Neither did whether the larger stimulus was on the right or the left, again implying there is no significant SNARC effect present. The three factors of distance, location and largerstim that were found to be significant improvements to the model for the previous experiment, again individually improve the linear model for experiment two.

3.4.2.4.3 Building the model – single factors

Therefore the model was built up, as before adding the significant single factors. A summary of this process is in table 3.4.

Additional factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Distance	5	15691.0	15726.1	-7840.5		
Location	7	15529.0	15578.2	-7757.5	165.9	<.001
Largerstim	8	15483.3	15539.6	-7733.7	47.71	<.001

Table 3.4 Results of ANOVA comparisons between versions of the linear model as single factors are added for experiment two

This single factor model is summarised in table 3.5. It can be seen that the distance effect was significant (p < .001) with b = -0.617 implying that logRT was reduced on average by 0.0617 for each 0.1 increase in distance (or RT is reduced by 5.98%).

Factor/level	b (95% CI)	SE	df	t-value	p
distance	-0.617 (-0.681, -0.554)	0.032	8290	-19.05	<.001
location: near_zero → middle	0.192 (0.157, 0.226)	0.018	8290	10.87	<.001
location: middle → near_one	0.019 (-0.021, 0.059)	0.021	8290	0.92	.356
largerstim: decimal → fraction	-0.091 (-0.117, -0.065)	0.013	8290	-6.92	<.001

Table 3.5 Summary of the linear model including all significant single factors for experiment two

The shift between location near_zero and middle was significant (p < .001) with an increase in logRT of 0.192 on average (or a 21.2 average increase in RT). However, the shift between the middle and near_one locations was not significant overall (p = .356).

In addition, stimuli in which the larger of the pair was a fraction were significantly (p < .001) faster, on average in log RT by 0.091 (a 8.70% decrease in RT).

3.4.2.4.4 Interaction with response

To find out whether these effects were significantly different for the two response category, as predicted, each of the interactions between response and distance, location and largerstim were added and tested for a significant improvement to the model.

Interaction added	df	AIC	BIC	Log Likelihood	L Ratio	p
No Interaction	8	15483.3	15539.6	-7733.7		
distance/response	9	15464.6	15527.9	-7723.3	20.68	<.001
location/response	11	15411.7	15489.0	-7694.9	56.93	<.001
largerstim/response	12	15394.9	15479.3	-7685.4	18.82	<.001

Table 3.6 Results of ANOVA comparisons between versions of the linear model as interactions with response condition are added for experiment two

It can be seen in table 3.6 that all three of the factors significantly (p < .001) interacted with response. This implies that each of the factors distance, location and largerstim had a different effect on RT for participants in the two response conditions. These differences between response conditions are illustrated in figures 3.2, 3.3 and 3.4 .

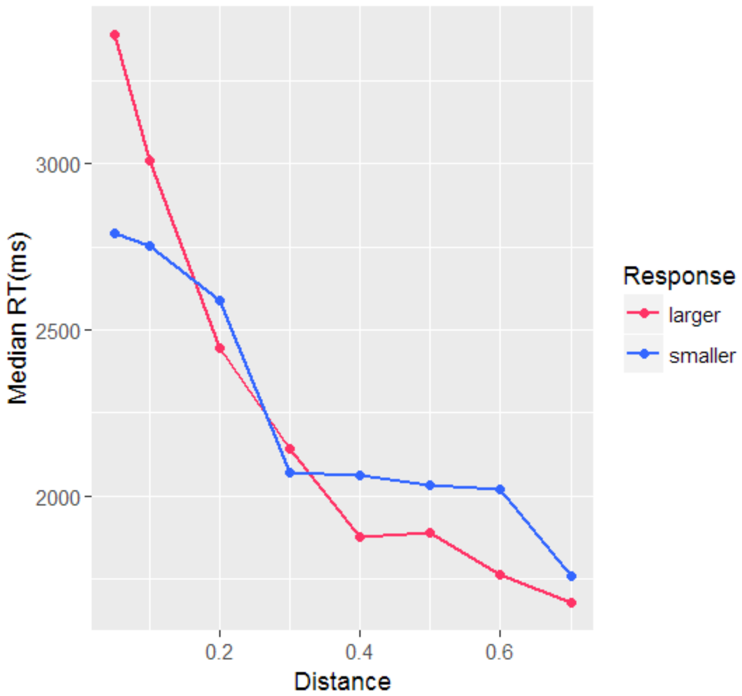

Figure 3.2 Graph of median RT against distance by response condition for experiment two

Figure 3.2 indicates that the distance effect was present for both response conditions. However, it appears to have been stronger for those in the response-larger condition.

Also, the addition of distances 0.6 and 0.7 highlighted that the levelling-off of the distance effect seen in the results of experiment one was not continued in experiment two for either response condition. RTs at distance 0.7 were shorter than at any of the smaller distances.

In fact, for both response conditions, the distance effect appears to have been approximately linear over distances 0.05 to 0.7. Interestingly, virtually the same pattern of median RTs for distances 0.05 to 0.5 is shown here for the response-larger condition as can be seen for experiment 1 (figure 2.3, page 33) even though participants in experiment two responded faster on average at all comparable distances.

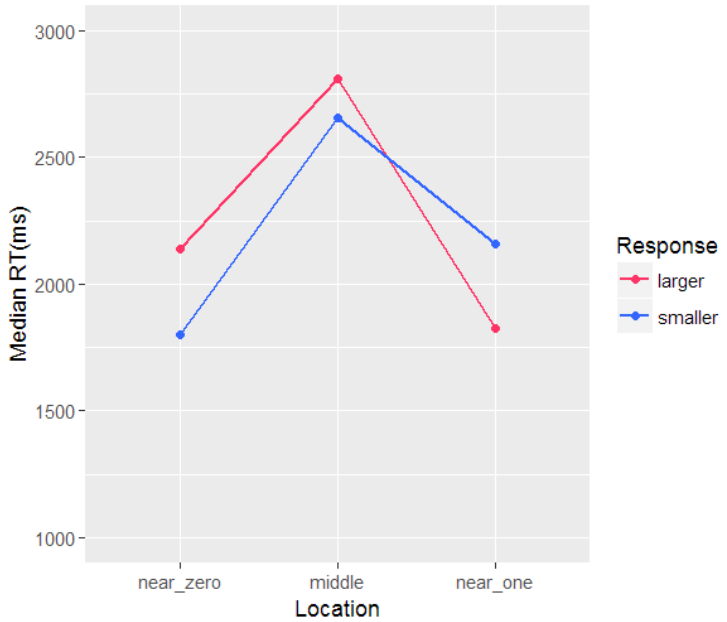

Figure 3.3 Graph of median RT against location by response condition for experiment two

Figure 3.3 demonstrates the difference in average RTs for the location factor between response conditions. The participants that chose the larger number were fastest in the near_one location, next fastest in the near_zero location and slowest in the middle location as in experiment one. However, the participants that chose the smaller number were fastest on average in the near_zero location, followed by the near_one location and then the middle location.

The effect that responses were slower when the decimal was the larger of the stimulus pair was significant in experiment one but does not appear to be so here for the comparable response-larger condition (fig 3.4). It does seem to be markedly present for the response-smaller condition, however.

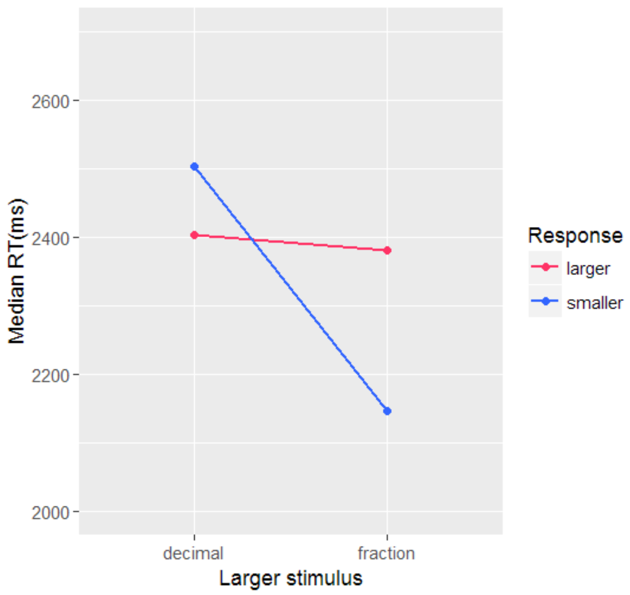

Figure 3.4 Graph of median RT against larger stimulus type by response condition for experiment two

In order to further investigate the differences between the two response conditions, a mixed linear model was built up for each condition separately.

3.4.2.4.5 Response-smaller model (with interactions considered)

The mixed effects linear model was built up by adding the fixed effects of single factors and interactions in the same manner as previous models within this document. The differing baseline intercepts of participants was the random effect specified in the model.

A summary of the final model is shown in table 3.7. The three main fixed effects of distance, location and largerstim made significant (p < .001) improvements to the baseline (random effect only) model. There was no improvement to the model made by any interactions between the factors. This contrasts with the results for experiment one for which there were interactions for the larger distances but not for the smaller distances (0.05, 0.1 & 0.2).

Factor/level	b (95% CI)	SE	df	t-value	p
distance	-0.474 (-0.561, -0.387)	0.044	4286	-10.68	<.001
location: near_zero → middle	0.277 (0.229, 0.324)	0.024	4286	11.45	<.001
location: middle → near_one	0.171 (0.116, 0.227)	0.028	4286	6.08	<.001
largerstim: decimal → fraction	-0.144 (-0.18, -0.109)	0.018	4286	-8.00	<.001

Table 3.7 Summary of the linear model for the response-smaller condition of experiment two

Therefore, for the participants who chose the smaller number of the pair, RTs decreased as distance increased; RTs in the middle location were significantly larger than RTs in the near_zero and near_one locations; RTs were significantly larger when the larger stimulus was a decimal than when it was a fraction.

3.4.2.4.6 Response-larger model (with interactions considered)

Again a similar mixed effects linear model was built up by adding the fixed effects of single factor and interactions; adopting additions to the model only when they made a significant (or near significant) improvement to the model.

The results were somewhat more complex than for those participants in the response-smaller condition. The fixed effects of distance and location made significant ($p < .001$) improvements to the model. However, the factor of largerstim made only a marginal ($p = .075$) improvement to the model as a single factor.

In addition, significant improvements were made to the model by the three interactions between distance and location ($p < .001$); distance and largerstim ($p = .002$); location and largerstim ($p < .001$). The three-factor interaction did not make a significant improvement to the model ($p = .144$).

A summary of the final model in shown in table 3.8.

single factors	b (95% CI)	SE	df	t-value	p
distance	-0.689 (-0.890, -0.489)	0.102	3995	-6.73	<.001
location: near_zero → middle	0.332 (0.233, 0.432)	0.051	3995	6.54	<.001
location: middle → near_one	-0.035 (-0.158, 0.088)	0.063	3995	-0.56	.578
largerstim: decimal → fraction	0.074 (-0.031, 0.179)	0.054	3995	1.37	.169
two-factor interactions					
dist/location near_zero → middle	-0.481 (-0.703, -0.260)	0.113	3995	-4.26	<.001
dist/location middle → near_one	0.123 (-0.129, 0.376)	0.129	3995	0.96	.339
distance/largerstim decimal → fraction	0.241 (0.060, 0.423)	0.093	3995	2.60	.009
largerstim decimal → fraction /loc. near_zero → middle	-0.190 (-0.289, -0.091)	0.051	3995	-3.76	<.001
largerstim decimal → fraction /loc. middle → near_one	-0.300 (-0.416, -0.185)	0.059	3995	-5.09	<.001

Table 3.8 Summary of the linear model including interactions for the response-larger condition of experiment two

Figures 3.5, 3.6 and 3.7 show the nature of the significant interactions found. There was a more pronounced distance effect for stimulus pairs in the middle location than in the other locations (figure 3.5).

There were some small deviations from the general distance effect. For stimulus pairs in which the larger was a decimal, average RTs increased slightly just between distances 0.4 and 0.5. Whereas for stimulus pairs in which the larger was a fraction, average RTs increased slightly just between distances 0.6 and 0.7 (figure 3.6).

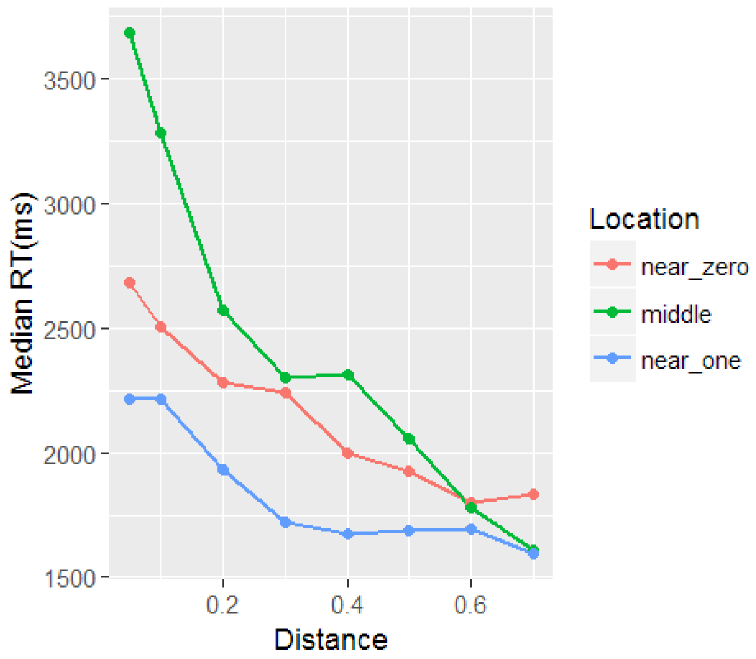

Figure 3.5 Graph of median RT against distance by location for response-larger condition of experiment two

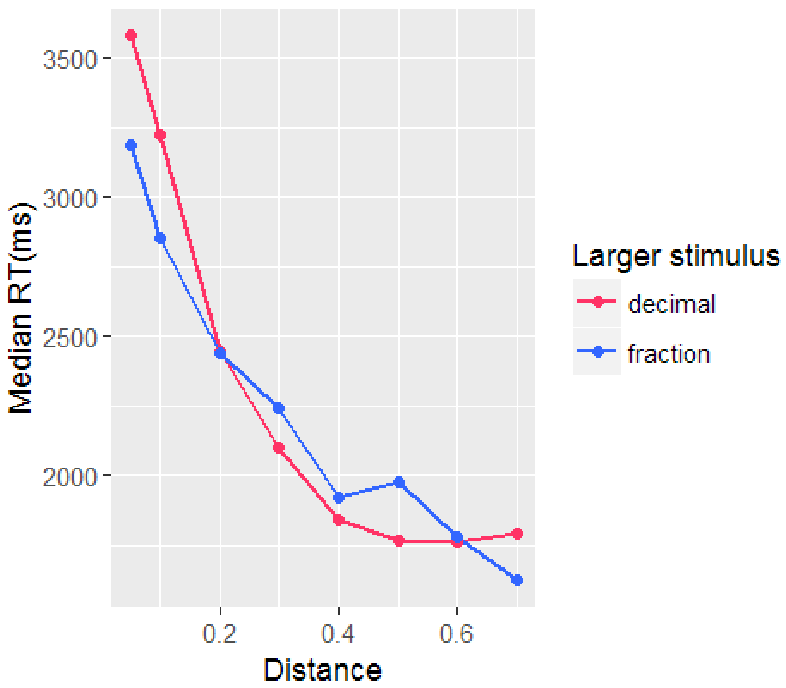

Figure 3.6 Graph of median RT against distance by larger stimulus type for response-larger condition of experiment two

Both when the larger stimulus was a fraction and when it was a decimal, RTs were larger, on average in the middle location than the other two locations. However, only for stimulus pairs in which the larger was a fraction, RTs were smaller in the near_one location than in the near_zero location (figure 3.7).

The interactions are more concerned with the relative strength rather than the general pattern of the effects of distance and the location factor on RT.

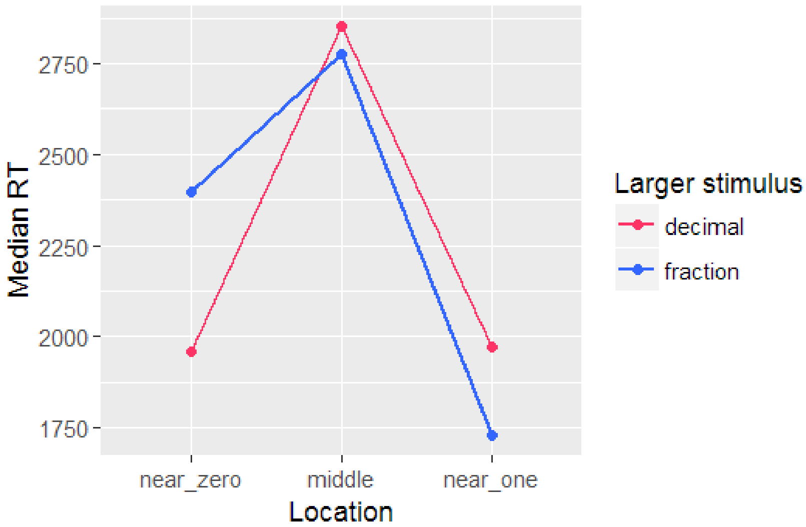

Figure 3.7 Graph of median RT against location by larger stimulus type for response-larger condition of experiment two

3.4.3 Error analysis

Again, as for experiment one, logistic (binomial) regression analysis of errors was carried out. due to the between subjects design, it was not feasible to include response condition as a factor in the analysis. The errors of the participants in the two response conditions were analysed both separately and together. Due to a difference between the results for the two response types, they are reported separately. Those participants in the response-smaller condition made more errors overall. Their error rate was 21.7% as compared to 16.4% for those in the response-smaller condition.

The coefficient and odds ratio results for each response can be seen in table 3.9.

Again, the factor of participant was included in the analysis but results for individuals are not reported in the table. However, individual differences between participants had a significant effect on the number of errors both for those in the response-larger condition ($\chi^2(27) = 223$, $p < .001$) and for those in the response-smaller condition ($\chi^2(29) = 390$, $p < .001$).

Table 3.10 shows the percentages of incorrect responses for each distance, location, larger stimulus type and the position of the larger stimulus. Percentages for the two response conditions are shown separately.

Response larger					
Source	b	SE	odds ratio (95% CI)	z value	p
distance	-3.679	0.297	0.025 (0.014, 0.045)	-12.37	<.001
location: near_zero → middle	0.413	0.133	1.511 (1.165, 1.961)	3.10	.002
location: middle → near_one	-0.245	0.175	0.783 (0.555, 1.103)	-1.41	.160
largerstim: decimal → fraction	-0.441	0.092	0.643 (0.537, 0.771)	-4.77	<.001
largelr: left → right	0.081	0.092	1.084 (0.905, 1.299)	0.89	.375
Response smaller					
Source	b	SE	odds ratio (95% CI)	z value	p
distance	-2.363	0.224	0.094 (0.061, 0.146)	-10.56	<.001
location: near_zero → middle	0.554	0.118	1.740 (1.381, 2.193)	4.71	<.001
location: middle → near_one	0.269	0.141	1.309 (0.993, 1.725)	1.90	.057
largerstim: decimal → fraction	-0.753	0.082	0.471 (0.401, 0.553)	-9.19	<.001
largelr: left → right	0.226	0.081	1.254 (1.070, 1.469)	2.80	.005

Table 3.9 Results of a logistic regression analysis of error data by response condition for experiment two

For both response conditions, there was a significant decrease in errors as distance between stimuli increased $\chi^2(1) = 257$, p <.001 for response-larger and $\chi^2(1) = 170$, p <.001 for response-smaller. Again, the odds ratio is difficult to interpret for distance but its small size implies a sharp drop-off in errors with as distance increases.

In addition, there were significantly more errors made when the larger stimulus was a decimal than when it was a fraction for both conditions. That is, for participants in the response-larger condition, errors were 1.56 times as likely when the larger stimulus was a decimal (or the smaller a fraction) with $\chi^2(1) = 23.0$, p <.001. For the response-smaller condition the figure is 2.12 times as likely ($\chi^2(1) = 86.8$, p <.001).

The location of stimuli within the zero to one range consistently had a significant effect on whether an error was made. For participants in the response-larger condition, $\chi^2(2) = 29.7$, p <.001; for those in the response-smaller condition, $\chi^2(2) = 27.4$, p <.001. For the response-larger condition, in the middle location, errors were 1.51 times as likely than in the near_zero location with no significant difference between the middle and near_one locations. Similarly, for the response-smaller condition, in the middle location, errors were 1.31 times as likely than in the near_zero location but the difference between the middle and near_one locations was not significant.

There was, however, a marked difference between the two response conditions in the effect of whether the larger stimulus was on the left or the right. This factor had no significant effect on errors for the participants that chose the larger stimulus. However, participants that chose the smaller stimulus were significantly more likely to make an error when the larger stimulus was on the right ($\chi^2(1) = 7.85$, p = .005).

Response larger								

distance	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7
% Errors	30.6%	23.9%	14.9%	9.5%	7.7%	6.0%	5.7%	6.3%

location	near_zero	middle	near_one
% Errors	10.7%	21.6%	9.3%

larger stimulus type	decimal	fraction
% Errors	19.0%	13.8%

position of larger stimulus	left	right
% Errors	15.8%	17.0%

Response smaller								

distance	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7
% Errors	33.9%	28.3%	21.0%	14.2%	15.0%	12.5%	12.8%	11.4%

location	near_zero	middle	near_one
% Errors	13.9%	26.5%	17.3%

larger stimulus type	decimal	fraction
% Errors	27.1%	16.2%

position of larger stimulus	left	right
% Errors	20.0%	23.4%

Table 3.10 Tables showing percentage of errors at levels of distance, location and largerstim factors by response condition for experiment two

3.4.4 Summary of results

3.4.4.1 Response times

By far the largest effect on RTs was the classic distance effect. This was true for both response conditions. One of the purposes of experiment two was to find out whether the distance effect for these types of stimuli was limited to smaller distances only (up to distances of approximately 0.3). The results, however, showed RTs did continue to decrease as distance increased up to the maximum distance used of 0.7, though perhaps less sharply for distances above 0.3.

The other main purpose of the experiment was to find out whether the largerstim effect would be affected by changing the response that participants had to give (choosing the smaller rather than the larger number). The results were completely unexpected in that for *both* response conditions, RTs were greater, on average, when the larger of the stimulus pair was a decimal, than when it was a fraction. Moreover, this effect was only significant for participants in the response-smaller condition.

An effect of location persisted for both response conditions in that RTs in the middle position were greatest.

3.4.4.2 Errors

The results for the error analysis of experiment two were largely the same as those for experiment one. For both response conditions, the most significant effect on error rates was the distance between stimuli. Rates of errors decreased as distance increased but again, this effect diminished for the larger distances.

In both response conditions, error rates also demonstrated a significant location effect. They were again, significantly higher for the middle location than for the near_zero and near_one locations. In addition, errors were significantly more likely to occur when the larger of the stimulus pair was a decimal than when it was a fraction.

The only difference between the two response conditions was that participants in the response-smaller condition made significantly more errors when the larger stimulus was (congruently) on the right than when it was (incongruently) on the left. This was the only evidence in either experiment of the SNARC effect.

3.5 Discussion

3.5.1 Response

I did not find a great difference in the nature of any of the main effects between the response-larger group and the response-smaller group. The distance and largerstim effects were significant and in the same direction for both response groups. The effect that RTs were larger in the middle location persisted with the change in response.
This contrasts somewhat with the findings of Arend & Henik (2015). They conducted a SiCE experiment with whole numbers and found that changing the instruction from *choose larger* to *choose smaller* significantly altered the outcome of their experiment. They concluded that the instruction given caused an *attention capture effect* towards the

goal of the task. That is, attention was drawn to the physically larger object presented if asked to choose the larger etc.

Of course, the SiCE effect indicates an unconscious processing of magnitude information and more specifically that the magnitude information is common to numerals and physical size. As discussed before, the magnitude comparison task in experiments one and two required far too much deliberate processing to be noticeably affected by unconscious processes.

The only outcome of experiment two which might be explained by an attention capture effect is the general reverse in the slight asymmetry of the location effect for the two response groups (see figure 3.3). When participants were asked to choose the larger number, they were slightly faster when the stimulus pair was in the near_one location than when it was in the near_zero location. This was reversed for participants asked to choose the smaller number.

These differences were small but might well have been due to attention being captured towards to top of the zero-to-one range by the instruction to choose the larger number and to the bottom of the range by the instruction to choose the smaller number.

3.5.2 Largerstim effect

Indeed, the main purpose of varying the response required of participants was whether the largerstim effect was influenced by the requirement of participants to response to the larger of the stimulus pair in experiment one. That is, was the *larger* part of the largerstim effect partly due to a focus of attention on the larger stimulus? Would it be diminished or reversed by a change in response.

The two stimuli presented in each trial did have physical size differences, as in the experiments of Arend & Henik (2015). Had the largerstim effect been reversed by the change in response, some similar kind of attention capture effect might have proved a plausible explanation for the largerstim effect. However, the largerstim effect was unchanged by the change in response. Even when asked to choose the smaller of two numbers people still responded more slowly and less accurately when the larger number in the stimulus pair was a decimal and the smaller number was a fraction.

Again, the only concrete conclusion that can be made is that there are differences between the processing or representation of fractions and decimals. As mentioned in the previous chapter, it is not surprizing that there would be a difference. Differences in external representations of numbers have been shown to affect responses to mathematical tasks e.g Gonzales & Kolers (1982) comparing Roman and Arabic numerals; Dehaene et al. (1993) comparing verbal with Arabic numbers. Perhaps even more relevantly, understanding of likelihood has been shown to be influenced by the representation of proportion used (Gigerenzer & Edwards 2003).

These differences do not preclude the accessibility of common (possibly approximate) holistic magnitudes. Indeed Dehaene et al. (1993) concluded that though the methods of processing of the magnitudes of the different external representations were distinct, these had to be interconnected to allow for comparison between external representations.

Nevertheless, the differences between the magnitude processing of fractions and decimals, particularly as measured by RTs, still require exploration.

3.5.3 Location effect

The lack of any SNARC effect in experiments one and two reinforces previous findings (Bonato et al., 2007; Kallai & Tzelgov, 2009) that the *holistic* magnitudes of fractional numbers are not processed automatically. However, important facts about the magnitude of fractions *have* been shown to be processed automatically.

Kallai & Tzelgov, (2009) found a SiCE when comparing fraction stimuli to the number one. This implies that the fact that a fraction is less than one is automatically processed. This is magnitude information. Furthermore, this result implies an influence of the number one when considering fraction magnitude. When presented with a fraction, it seems we are automatically aware that it is less than one, whether helpful or not to the task at hand.

The location effect found in my experiments one and two also implies that fraction and maybe decimal magnitude knowledge is tied or *anchored* not only to one but to zero as well. This highlights the need for more investigation into the significance of zero and one in the mental representations of fractional numbers.

In one study, Ganor-Stern (2012) compared the magnitude of unit fractions ($^1/_2$, $^1/_4$, $^1/_6$ & $^1/_9$) with either zero or one. They did not find a distance effect when comparing these unit fractions with one but they did find that this was quicker than comparing them with zero. One interpretation of this result is that there is a fast, automatic response to any given fraction that it is *smaller than one*. Also there is not an equally fast *larger than zero* automatic response.

However, it is problematic that only unit fractions were used. As they were unit fractions, Ganor-Stern's stimuli were all closer to zero than one (except $^{1}/_{2}$) so the fact that comparisons were quicker against one than against zero might be seen as a distance effect of sorts.

It has been shown that different types of fraction stimuli produce different results even when used within the same magnitude comparison task (e.g. Meert et al., 2009; Meert, Grégoire & Noël, 2010; Faulkenberry & Pierce, 2011). Therefore, comparing a wider range of fractions against both zero and one might be throw more illumination on the significance of these two numbers to our mental representation of fractional numbers.

3.5.4 Distance effect

Experiment two successfully replicated the strong distance effect found in experiment one. When comparing a relatively unfamiliar fraction with a three digit decimal, RTs were again dependent on the magnitude distance between the two stimuli with larger distances being associated with shorter RTs. Does this show a comparison between fractions and decimal based upon common holistic magnitudes?

It has been demonstrated that people automatically process the separate components of fractions (Kallai & Tzelgov, 2012a) and each place value component of a (whole) decimal number separately (Kallai & Tzelgov, 2012b). Fraction magnitude comparison tasks have been shown to be carried out using componential magnitude comparison alone (Bonato et al., 2007) However, these findings do not preclude the use of holistic magnitude mental representations of fractions and decimal fractions any more than of multi-digit whole numbers.

Indeed, Zhang & Wang (2005) showed it is possible to remove the distance effect for two-digit whole numbers by manipulating the stimuli presented. Similarly, it has been shown that the use of components only for the comparison of the magnitude of two fractions is dependent on the stimuli presented.

People might have more automatic responses to the single digit components of a number presented to them. However, they can and will use holistic magnitudes to make a comparison between two fraction stimuli if they do not allow for a componential comparison (e.g. Meert et al., 2009, 2010; Schneider & Siegler, 2010; Faulkenberry &

Pierce, 2011). Indeed, people were observed to process components of fractions before forming a holistic magnitude representation in the mouse-tracking study of Faulkenberry, Montgomery & Tennes (2015). Also recent studies from neuroscience imply that the brain encodes whole magnitudes rather than components when presented with fractions (Ischebeck, Schocke & Delazer, 2009; Jacob, Vallentin & Nieder, 2012).

Yet in my experiments people were comparing fractions with decimals so consideration must be made of how people process and represent decimal magnitudes also. When directly comparing the magnitude of two 3-digit decimal fractions with one another, a distance effect is found that mirrors that of 3-digit decimal whole numbers (DeWolf, Grounds, Bassok & Holyoak, 2014). This is an outcome similar to that found by Huber, Klein, Willmes, Nuerk & Moeller (2014) in an eye-tracker study contrasting the comparison of pairs of decimals with the comparison of pairs of whole numbers.

Matthews, Chesney & McNeil (2014) did cross-format magnitude comparison tasks for fractions and diagrams of proportion. They found a distance effect for fractions versus non-symbolic representations implying participants were accessing holistic magnitude representations of fractions that were directly comparable to visual proportions.

But what of comparisons *between* fractions and decimals? Response times for my experiments one and two were relatively long. They clearly were a difficult set of magnitude comparisons to make. However, with a very few exceptions, participants were remarkably successful at the task and a strong, significant distance effect was found in both. The stimuli were designed to make it very difficult indeed to successfully compare them in any way other than via holistic magnitudes. The fact that such a strong distance effect was found implies that there is indeed a common, (if fuzzy), mental number line for fractions and decimals.

The distance effect has often been found to diminish in size with an increase in distance (e.g. DeWolf et al., 2014; Matthews et al., 2014; Schneider & Siegler, 2010). For simpler experimental designs than mine, that used a target-stimulus paradigm, it has been possible to accurately model this shape as logarithmic. However, this is not a universal result and a linear model implying no decrease in t distance effect as distance between fractions increases has also been found (Faulkenberry, 2011).

In my experiment one, the distance effect appeared to completely plateau between distances 0.4 and 0.5. Part of the changes to the design of experiment two was to look for an extension to the distance effect at larger distances. This certainly seems to have been found. In experiment two, the average RT continued to decrease as distance increased up to 0.7 (see figure 3.2). It also appeared to be approximately linear in nature for each response group. However, though the distance effect was fairly consistent across all three location levels, I do have some concern that for the largest distances (particularly 0.7) one of the stimuli would be fairly close to either zero or one for all stimulus pairs.

3.5.5 Next steps

Nevertheless, the questions left open after experiments one and two are not about the distance effect. They are about the emergent results of the location and largerstim effects.

What requires further investigation is whether differences can be found between the magnitude representations of fractions and decimals that might help explain the causes of the largerstim effect. Additionally, are estimations of the size of fractions and decimals anchored to the end points of the zero-to-one range? Is this the cause of the location effect?

Chapter 4 Experiments 3 and 4

4.1 Introduction

This fourth chapter is a commentary on the short third and fourth experiments that were carried out simultaneously for the purpose of this thesis. It starts with a summary of the reasoning which led to the design of these experiments; in particular, how they follow on from the findings of the previous two experiments. The details of the design of both experimental tasks and stimuli follows.

The next part of the chapter covers the methodology and outcomes of experiment three. This includes analysis of the results and any findings thereof. The third section of the chapter similarly covers the experimental procedure, analysis and results of experiment four.

Finally, the last section of the chapter is a consideration of the implications of the results of these two experiments with an assessment of whether they were helpful in clarifying outcomes of the first two experiments.

4.2 Justification

4.2.1 Revisiting the results of experiments one and two

Both of the magnitude comparison tasks of relatively unfamiliar fractions and decimals, that were experiments one and two, produced evidence of three significant influences on both RTs and error rates.

The largest of these was the predicted distance effect. This implied the existence of a magnitude understanding of these numbers that is to some extent, common to both forms of fractional numbers of the kind presented. The extent of commonality and difference of the magnitude representations of fractions and decimals was left in question.

Possible differences were highlighted by a surprising effect on RTs uncovered by experiment one which was again found in experiment two. This was that when the larger stimulus in the magnitude comparison task was a decimal, RTs were significantly greater. Experiment two confirmed that this *largerstim effect* was not due to the phrasing of the question asked. Asking participants to choose the smaller of two numbers rather than the larger led to the same result.

This outcome suggests that there might be some key difference in how the magnitudes of fractions and decimals are represented in the brain. It could be, for example, that comprehension of the magnitude of a decimal gets harder as the decimal gets larger or that comprehension of the magnitude of a fraction gets easier as the fraction gets larger.

Additionally, in both experiments one and two, there was an effect on RTs of the location of stimuli within the zero-to-one range. Responses were significantly faster when one of the stimulus pair was very near to (within 0.1 of) zero or one than when they were in the middle. This effect was not universally found to interact with the largerstim effect. So this would seem to have been an effect that was acting on both fractions and decimals in the same way. Therefore implying that judging the magnitude of *both* fractions and decimals is easier when they are very close to zero or one than when they are not.

Another finding to consider comes from the reporting of strategies in experiment one. Many participants reported that they were using ½ or 0.5 as an anchor point against which to judge which stimulus was larger or smaller. However, the consequent effect was not found that responses were faster when the stimuli were positioned either side of 0.5. This can be seen, particularly in the results of trials at the smaller distances of 0.05, 0.1 and 0.2. For these trials, the central position (D) was the only one with stimuli either side of 0.5 but RTs were not significantly quicker for this position than they were for any of the other middle positions in either experiment.

Indeed, the effect of location, would more imply that the key *anchor points* used against which to make a magnitude judgement are zero and one only; rather than, for example 0.5, as reported by participants.

Within the stimulus pairs presented in experiments one and two, one number was always closer to zero and the other number was closer to one. If the number comparison was, at least partly, being made by judging the closeness of stimuli to zero and one then the location effect could have arisen from an anchoring effect. This has been observed before in the magnitude estimation of whole numbers (Izard & Dehaene, 2008).

In the context of the decimals and fractions presented in experiments one and two, the anchoring effect would mean the judgement of their magnitudes were each anchored at either zero or one and then adjusted away from the anchor. So the further away from the anchor(s) the stimulus was, the longer the estimation took because more adjustment needed to take place. However, it should be noted that there was not any evidence of gradation of the location effect in experiments one and two. Responses were quicker for stimuli positioned within 0.1 of zero or one but similarly slower otherwise.

A simple same/different to zero or one judgement would not be sufficient to make all size comparisons. However, when only one of the numbers passes some threshold which makes it sufficiently the same as zero or one, a comparison could thus be made. If neither number is judged as such, some other methodology would need to be employed. This *could* be the explanation of the location effect found in experiments one and two, particularly at the smaller distances.

There did seem to be some plateauing of the distance effect for the larger distances (above 0.4). For these, even the stimulus pairs in the *middle* position, might be judged at either end to be sufficiently close to zero or one. The judgement that near_zero and near_one are within 0.1 of the ends of the zero-to-one range was somewhat arbitrary and based on the need to classify the positions chosen in the original design of experiment one. So perhaps the stimulus pairs at the larger distances were similarly as fast as each other because for these, one of the stimuli was beyond the threshold at which it could be recognised as almost zero or almost one.

The three questions that experiments three and four were designed to answer were:

1. Could more evidence be found of similarities and differences between the magnitude representations of fractions and decimals; specifically for the types of fractions and decimals used in experiments one and two?
2. Could any differences found help to explain the largerstim effect? That is, could an explanation be found for an increase in the difficulty of the magnitude comparison task when the larger stimulus is a decimal and the smaller a fraction rather than vice-versa?
3. Could more evidence be found of a location effect? That is, are magnitude-type judgements of fractions and decimals particularly easy/fast near to zero and one?

4.2.2 Experiment three design

Experiment three was a magnitude comparison task with the *two* targets of zero and one. It was designed to discover whether the effect of location on RTs and error rates in experiments one and two might be somewhat explained by judging the individual stimuli against *both* zero and one. That is, would far smaller RTs (and error rates) again be found for stimuli very near to zero and one with fairly uniform RTs (and error rates) for those in between?

Experiment three was additionally intended to investigate similarities and differences in the pattern of responses for fraction stimuli and decimal stimuli. It was hoped to thus throw light onto the largerstim effect found in the previous experiments. That is, would RTs and errors be particularly low for fractions very close to one and for decimals very close to zero?

A very simple task was utilised. Participants were shown single fractions or decimals and asked to judge whether they were closer to zero or closer to one. Response times and accuracy were recorded.

4.2.3 Experiment four design

Experiment four was a magnitude estimation task. It also was designed to investigate causes of the largerstim effect and find out whether a different pattern of responses would be found, in RTs to fractions and decimals stimuli. Again, would RTs be particularly low for fractions very close to one and for decimals very close to zero?

In addition, it was intended to discover whether the effect of location on RTs and accuracy in experiments one and two could, in fact, be explained by the challenge associated with estimating the magnitude of numbers of different sizes within the zero-to-one range. That is, would far smaller RTs again be found for stimuli very near to zero and one with fairly uniform RTs for those in between?

The size of errors for an estimation task would be quantifiable. Therefore, rather than looking at error rates, would the size of errors be greater in the middle of the zero-to-one range? Would errors for fractions be particularly small nearest to one; and for decimals be particularly small nearest to zero?

In the experiment four task, participants were shown single fractions or decimals and asked to estimate their size by placing a mark on a line. It was particularly important that the stimuli were presented in random order for this task as it has been shown that successive number estimations can be biased by the first number presented (Sullivan, Juhasz, Slattery & Barth, 2011).

4.3 Experiment three

4.3.1 Method

4.3.1.1 Participants

38 psychology students at the university of Huddersfield, (5 men), participated in the experiment in return for course credit. Their ages ranged from 18.6 to 37.0 years with an average age of 20.9 years and a standard deviation of 3.6 years.

4.3.1.2 Stimuli

Stimuli consisted of one relatively unfamiliar number with a magnitude between zero and one; either a decimal or a numerator/denominator fraction. The fractions were the sixty-four fractions with denominators of 11, 13, 15, 17 or 19 that cannot be simplified. The decimals were the decimal equivalents of these sixty-four fractions, rounded to three decimal places. This made 128 stimuli in total.

4.3.1.3 Procedure

The experiment was conducted using E-Prime® 2.0 stimulus presentation software in a lab at the University of Huddersfield. Participants were instructed that within each trial they had to press the 'Z' key if they decided if the number shown was closer to zero or the 'M' key if it was closer to one. They were informed that both speed and accuracy of response were important.

A practice block of four stimuli preceded the experimental blocks. Participants were given feedback on their accuracy in the practice block stimuli and were allowed to ask questions if they did not understand the procedure. All of the 128 experimental stimuli were then presented in random order with no further feedback on accuracy nor opportunity to ask questions.

Response times and accuracy were recorded by the program.

4.3.2 Results

4.3.2.1 Response time analysis

4.3.2.1.1 Pre-analysis data processing

One participant made 115 errors out of 128 trials (90%). All participants had obtained at least a level two maths qualification, (GCSE or equivalent). Therefore it was not considered possible that this participant was unable to tell if a number was closer to zero or one. It was judged, instead, that they either misunderstood the task or had deliberately not completed the task correctly. In either case, their results were not considered valuable and were not included in any analysis.

The number of errors for the remaining thirty-seven participants ranged from 0 to 48. All performed better than chance at the task at the 5% level. That is, they made fewer than 55 errors out of 128 trials, B(128, 0.5). Therefore, all were included in the analysis. Their ages ranged from 18.6 to 37.0 years with an average age of 20.9 years and a standard deviation of 3.6 years.

Two datasets were formed for analysis. One was the average response times (mRTs), across participants, for each of 64 decimal stimuli. The other was the mRTs, across participants, for the 64 fraction stimuli. Measurements are in ms.

One potential outlier was identified for the decimal mRTs. This was for the decimal stimulus 0.462 which had an outcome 3.71 standard deviations above the mean for the decimal mRTs. Investigation of the individual responses for this stimulus, (decimal 0.462), identified only one value (out of 37) that might have been considered an outlier for the candidate concerned. It was 3.38 standard deviations above the mean for that candidate's responses (p = .0004). The analysis detailed in this section was also run with the outlier replaced by $M + 2SD$ for the decimal mRTs. This had virtually no effect upon the results obtained. Therefore, it was decided that there was no overwhelming reason to omit or adjust any data for the decimal 0.462 stimulus and the analysis reported herein is of the unaltered data.

4.3.2.1.2 Separate analysis of mRTs for decimal and fraction stimuli

Histograms showed both datasets to be approximately normal though with some possible positive skew for the decimal stimuli mRTs (figure 4.1). However, the method of Crawley (2005), found insignificant skew for both decimal stimuli mRT (γ = 1.08, p = .141) and fraction stimuli mRT (γ = -0.057, p = .523).

To investigate the shape of the RTs across the zero-to-one range, scatter diagrams were produced of mRT against the size of the stimulus presented for each of the fraction and decimal datasets (figure 4.2). Locally weighted regression (loess) calculations were performed to find smoothed patterns for the data (span = .75, polynomial degree = 2). The results are shown in green on figure 4.2.

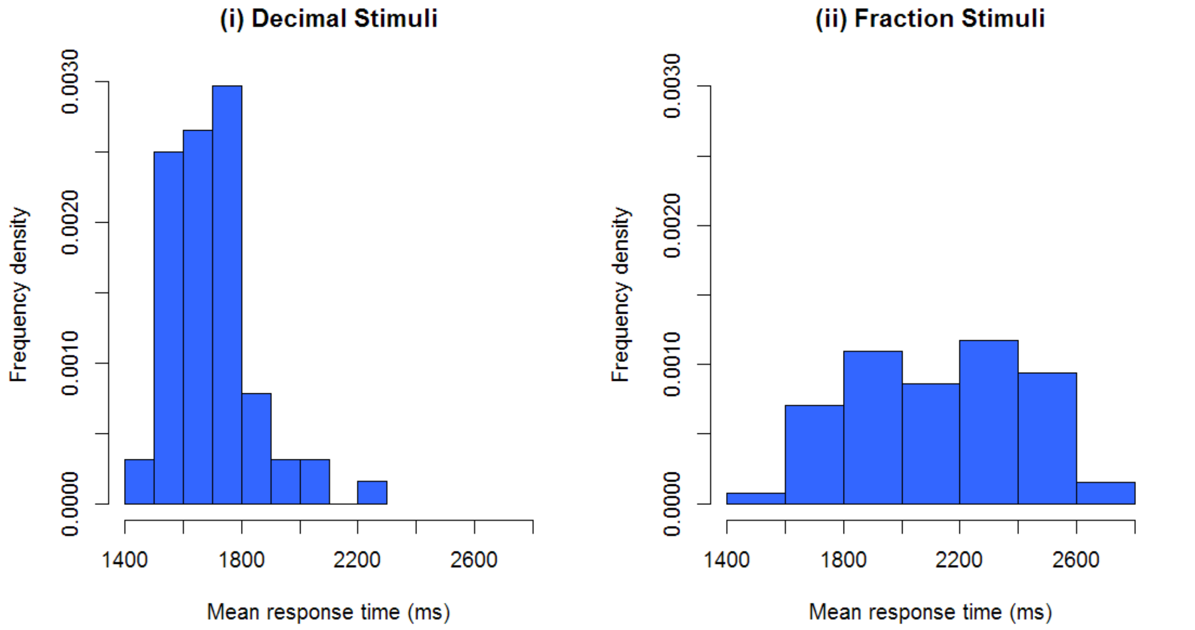

Figure 4.1 Histograms showing distribution of mean RTs for (i) decimal and (ii) fraction stimuli for experiment three

Neither of the distributions of mRTs had the shape of the RTs for the seven positions of the smaller distances in experiments one and two (see figure 2.4, page 34). These were generally flat in the middle with sharp declines at either end. grew had a distinct peak

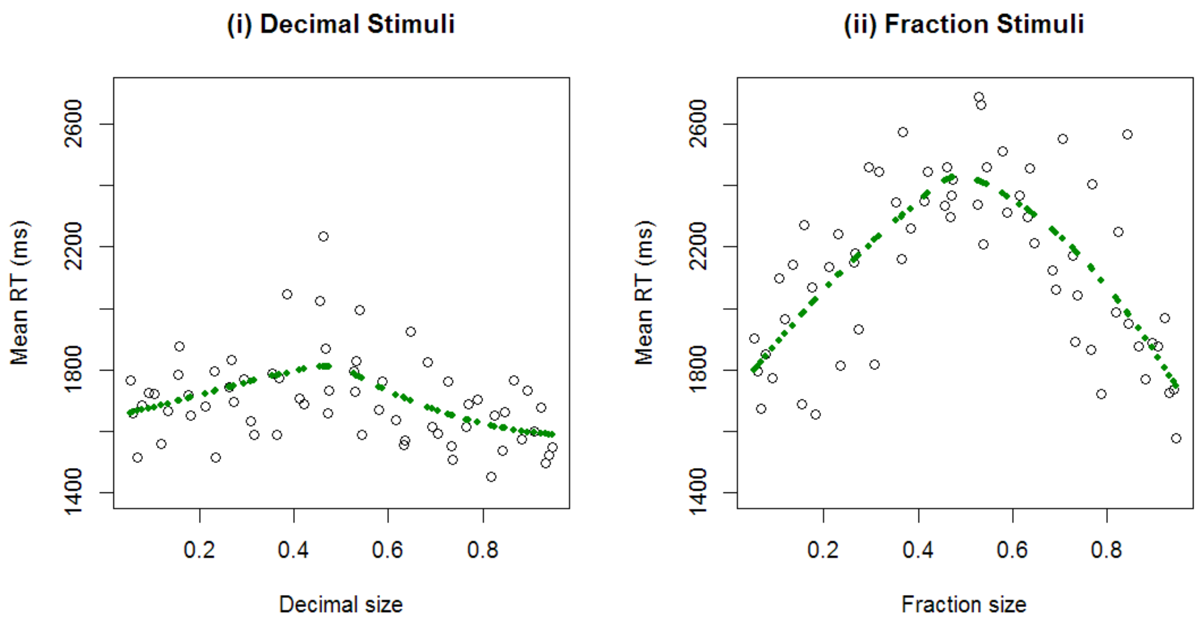

Figure 4.2 Scatter diagrams of mean RT against stimulus size for (i) decimal and (ii) fraction stimuli for experiment three

Particularly different from those was the distribution of the mRTs for the fraction stimuli of experiment three which had an approximately negative parabolic shape. That is, faster mRTs for the smallest and largest stimuli, (nearest to zero and to one), with mRTs increasing for the middle stimuli; reaching a maximum mRT for a stimulus close to 0.5 in size.

For the decimal stimuli the pattern was similar, rising in the middle of the range but not with the distinctive parabolic curvature. There was also a much less pronounced difference between mRTs for stimuli at the centre of the zero-to-one scale and those nearest the ends of the scale. The potential outlier (0.462, 2235) can be clearly identified on the scatter diagram (figure 4.2) as being the largest mRT but it does follow the general pattern of responses increasing toward the middle of the stimulus range.

Due to the approximately parabolic shape of the data, multiple linear regressions were calculated to predict mRT based on stimulus size and (stimulus size)^2 for each dataset. The purpose of these models was descriptive rather than predictive. The fraction stimuli were an exhaustive list within set parameters and not a random nor a representative sample of all fractions. Therefore any models obtained could not legitimately be used to predict anything other than a general pattern of results for other types of fractions.

For decimal stimuli a significant regression equation was found ($F(2,61) = 7.67$, $p = .0011$). Both stimulus size ($t = 2.73$, $p = .0082$) and (stimulus size)^2 ($t = -3.311$, $p = .0016$) were significant predictors of mRT. A significant regression equation was also found for fraction stimuli ($F(2,61) = 36.0$, $p < .001$). Both stimulus size ($t = 8.24$, $p < .001$) and (stimulus size)^2 ($t = -8.49$, $p < .001$) were significant predictors of mRT. The regression models are displayed on figure 4.3 along with r^2 values for the models, adjusted for the additional predictor. Details of the coefficients are shown in table 4.1.

		Decimal stimuli		
Coefficient	Estimate	SE	95% CI (model)	Bootstrapped 95% CI
intercept	1618	55.3	(1510, 1727)	(1528, 1699)
stimulus size	701	257	(198, 1204)	(268, 1254)
(stimulus size)^2	-826	249	(-1315, -337)	(-1390, -405)
		Fraction stimuli		
Coefficient	Estimate	SE	95% CI (model)	Bootstrapped 95% CI
intercept	1580	83.3	(1417, 1744)	(1423, 1707)
stimulus size	3178	386	(2421, 3934)	(2615, 3804)
(stimulus size)^2	-3185	375	(-3920, -2449)	(-3738, -2595)

Table 4.1 Details of the coefficients of the parabolic models for mean RT for decimal and fraction stimuli for experiment three

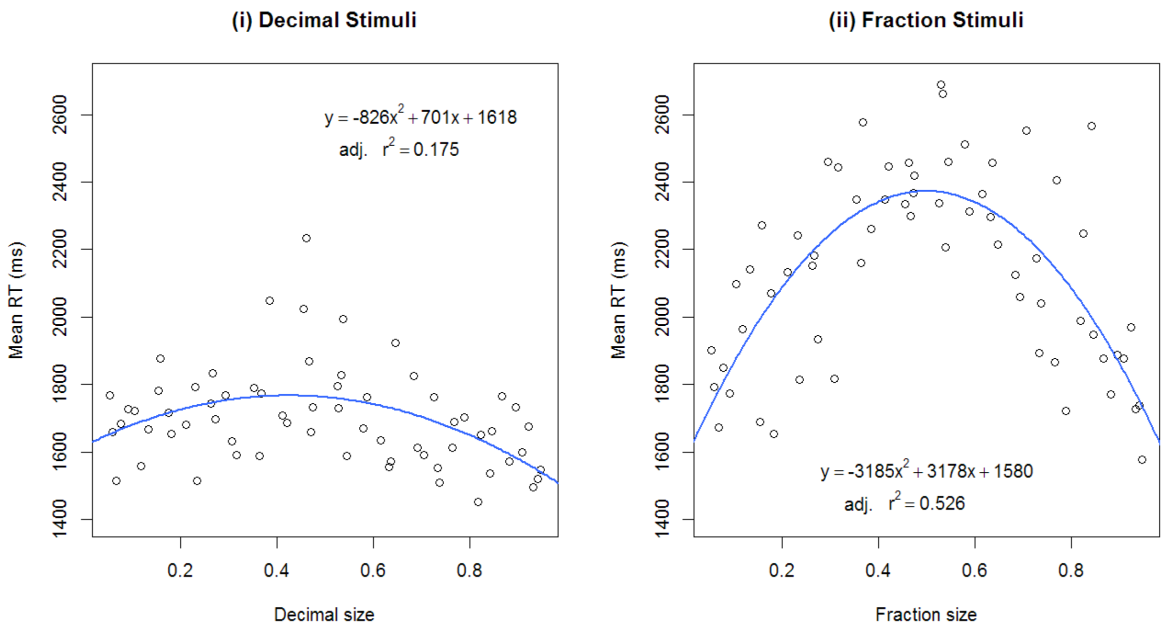

Figure 4.3 Parabolic models of mean RT for (i) decimal and (ii) fraction stimuli for experiment three

There is no particularly systematic pattern to the scatter around the model on either graph. The residual plots for the two models (figure 4.4) do not have obviously non-random patterns. They are centred roughly around zero with one outlying value each (at decimal 0.462 and at fraction $^{10}/_{13}$). This indicates that the parabolic multiple linear regression models using stimulus size and (stimulus size)^2 as predictors of mRT were somewhat appropriate.

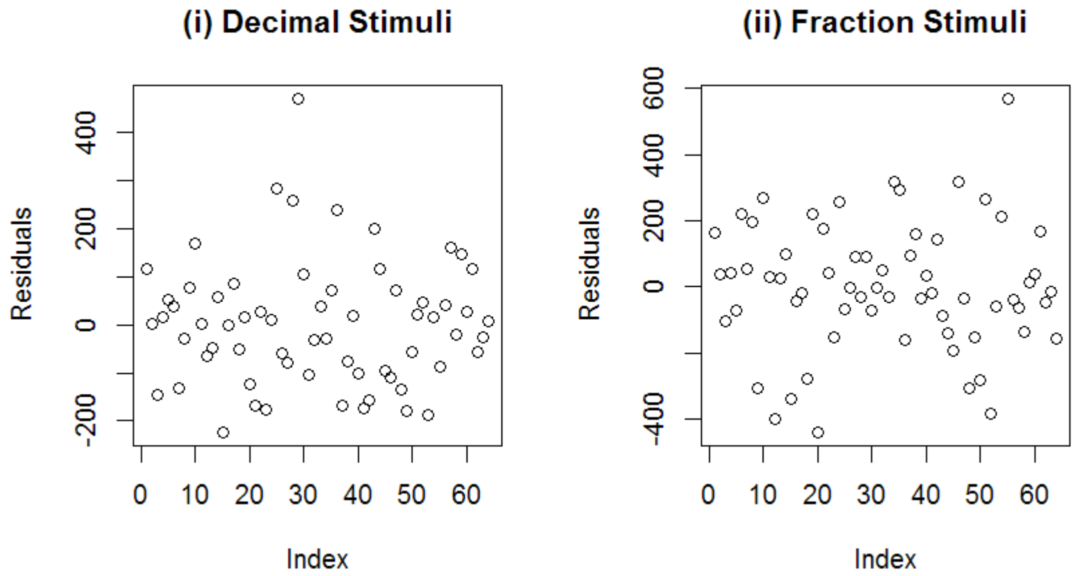

Figure 4.4 Plot of residuals for parabolic models for mRT for (i) decimal and (ii) fraction stimuli for experiment three

Confidence intervals for the model coefficients were also found by bootstrapping with 5000 replications. The results are shown in table 4.1. The results do not differ greatly from those found by regression analysis implying that the datasets are close enough to normal distributions for the parametric regression analysis to be valid.
The adjusted r^2 values indicated that the model for the decimal stimuli accounted for only 17.5% of the variance within mRTs to decimal stimuli whereas the model for the fraction stimuli accounted for 52.6% of the variance within mRTs to fraction stimuli. The fraction model has a considerably steeper curve (gradient function $y = 3178 - 6370x$) than the decimal model (gradient function $y = 701 - 1652x$)The fraction model peaks at the centre of the zero-to-one range (0.499). The decimal model peaks at 0.424, slightly to the left of the centre of the range.

Responses to fraction stimuli were far more affected by the position of the stimulus within the zero-to-one range than were responses to decimal stimuli. Additionally, the nature of the effect was virtually symmetrical about ½ for fractions.

One additional test of mRTs was applied to the results of the fraction stimuli only. This was a Kruskal-Wallis test to check whether there was evidence of any difference in the average mRT results for the five different denominators (11, 13, 15, 17 & 19). If one or some of the denominators made the task particularly easier than the others that could have confounded the results. No significant difference between denominators was found $\chi^2(4) = 2.04$, $p = .728$.

4.3.2.1.3 Analysis of the difference between fraction and decimal mRTs

The scatter diagrams in figure 4.2 and the histograms in figure 4.1 indicated that mRTs for the decimal stimuli were smaller than those to the equivalent fraction stimuli. A paired t-test was conducted on mRT of decimal and fraction stimuli of the same size. Average response times to decimal stimuli (M = 1700ms, SD = 144ms) were significantly lower than those to fraction stimuli (M = 2134ms, SD = 286ms), with very large effect size; $t(63) = 13.7$, $p < .001$, $d = 1.92$. Effect sizes are interpreted using the benchmarks proposed by Cohen (1988) and Rosenthal (1996).

To mirror the largerstim results of experiments one and two, the pattern of the paired RT differences (fraction – decimal stimuli) should have been positive near zero and negative near one. That is because the largerstim effect possibly implies that it is more difficult to judge the size of fractions than decimals near zero (when they are smaller) and more difficult to judge the size of decimals than fractions near one (when they are larger).

To investigate the pattern of differences, a scatter diagram of difference in mRT against stimulus size was plotted with loess calculations performed as above, to find a smoothed pattern for the data (figure 4.5). It showed an approximately parabolic shape with lower differences near to the stimulus size end points of zero and one; with higher differences in the middle of the range. The differences are almost all positive and roughly of the same size at either end of the zero-to-one range.

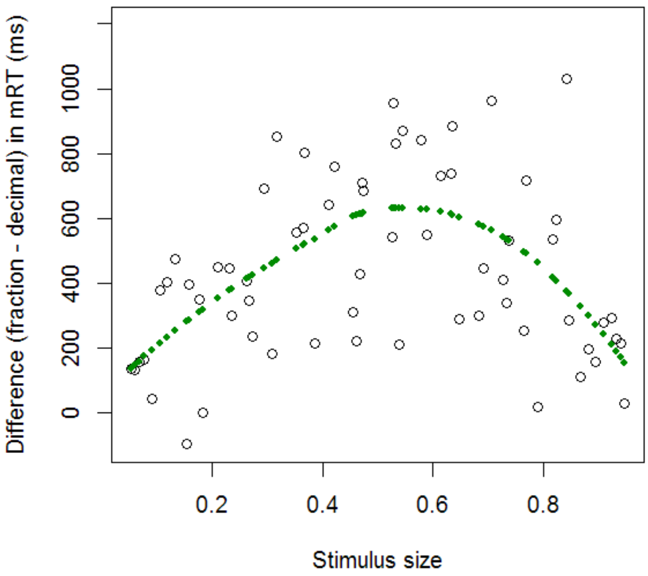

Figure 4.5 Scatter diagram of difference in mean RT (fraction – decimal stimuli) against stimulus size for experiment three

So again a multiple linear regression was fitted using stimulus size and (stimulus size)^2 as predictors of mRT differences. A significant regression equation was found ($F(2,61) = 15.7$, $p < .001$). Both stimulus size ($t = 5.59$, $p < .001$) and (stimulus size)^2 ($t = -5.48$, $p < .001$) were significant predictors of mRT. Details of the coefficients are shown in table 4.2.

Coefficient	Estimate	SE	95% CI (model)	Bootstrapped 95% CI
intercept (not significant)	-37.8	95.6	N/A	(-198, 86)
stimulus size	2477	443	(1609, 3345)	(1768, 3211)
(stimulus size)^2	-2358	431	(-3203, -1514)	(-3037, -1596)

Table 4.2 Details of the coefficients of the parabolic model for differences in mean RT for experiment three

No systematic pattern to the scatter around the model is apparent on the graph and the residual plot for the model (figure 4.6) has a suitably random pattern around zero with just one potential positive outlier (at decimal 0.842, fraction $^{16}/_{19}$). Confidence intervals for the model coefficients were also found by bootstrapping with 5000 replications, shown in table 4.2. The results do not differ greatly from those found by regression analysis.

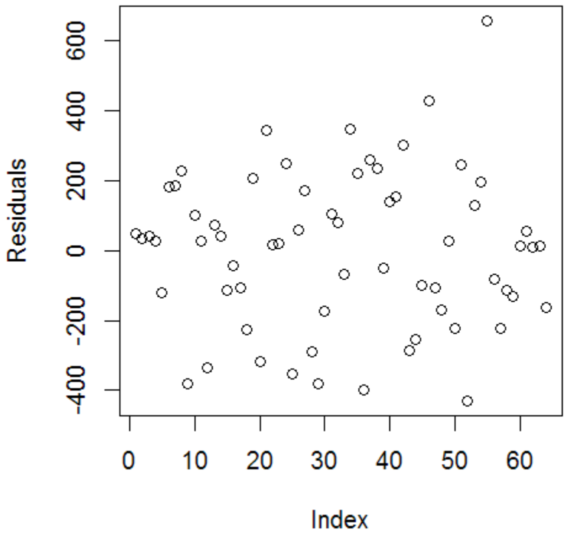

Figure 4.6 Plot of residuals for parabolic model for differences in mean RT for experiment three

These findings imply that the multiple linear regression model using stimulus size and (stimulus size)2 as predictors of mRT difference was appropriate. The adjusted r2 value indicates that the model accounted for 31.7% of the variance within mRT differences.

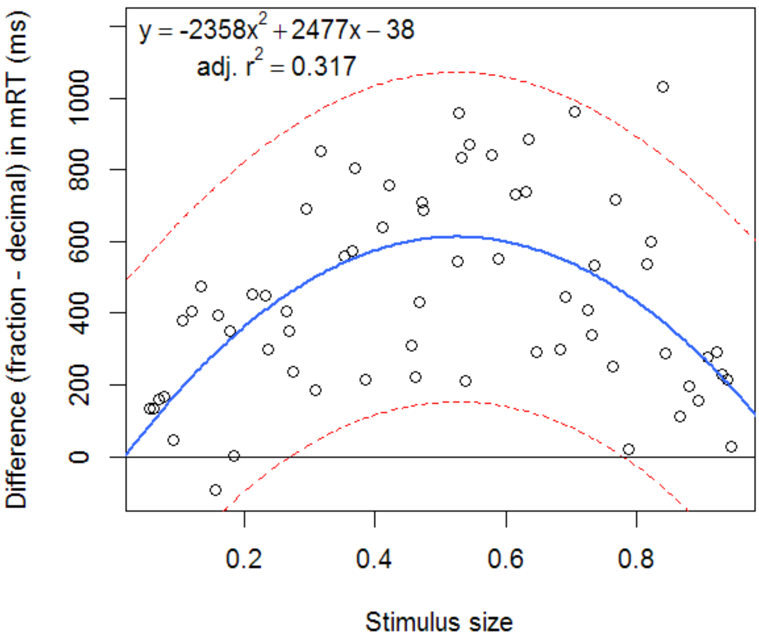

Figure 4.7 Parabolic model of mean RT for differences in mean RT with 95% CI for experiment three

The regression model with the associated adjusted r^2 is shown on figure 4.7, graphically depicted as a solid blue line, with 95% CI for the model shown as red dotted lines. For $0.269 < \text{stimulus size} < 0.780$ the 95% CI is completely above the difference = 0 line. This implies that for stimulus sizes in the interval (0.269, 0.780), responses to fraction stimuli were significantly larger than they were to their equivalent decimal stimuli.

It also can be seen that the model reaches its maximum only slightly to the right of the centre of the zero-to-one range (at 0.525). This implies that within the middle of the range it is significantly harder to identify whether a fraction is closer to zero or one than it is for the equivalent decimal. However, at the far ends of the range it is not significantly harder to make this judgement for a fraction than a decimal.

4.3.2.2 Error analysis

4.3.2.2.1 Pre-analysis data processing

Analysis of the errors made by participants was conducted. These data were far more problematic than the response times data as participants might have given incorrect responses even when they knew the correct response to give. They had no opportunity to correct their responses. All participants had at least a level two qualification (minimum GCSE grade C or equivalent). As such, the task itself should have been relatively easy for all participants. Indeed, several verbally reported frustration at having "pressed the wrong button".

Again, the erroneous responses were processed into two separate datasets. One was the percentage rate of errors, across participants, for each of 64 decimal stimuli. The other was the percentage rate of errors, across participants, for the 64 fraction stimuli.

4.3.2.2.2 Separate analysis of errors for decimal and fraction stimuli

Both datasets were found to be somewhat but not significantly positively skewed using the method of Crawley (2005). The error rate for decimal stimuli produced $\gamma = 1.59$, $p = .058$ and for fraction stimuli, $\gamma = 1.06$, $p = .147$. However, as can be seen in histograms of the datasets (figure 4.8), neither distribution looks approximately normally distributed.

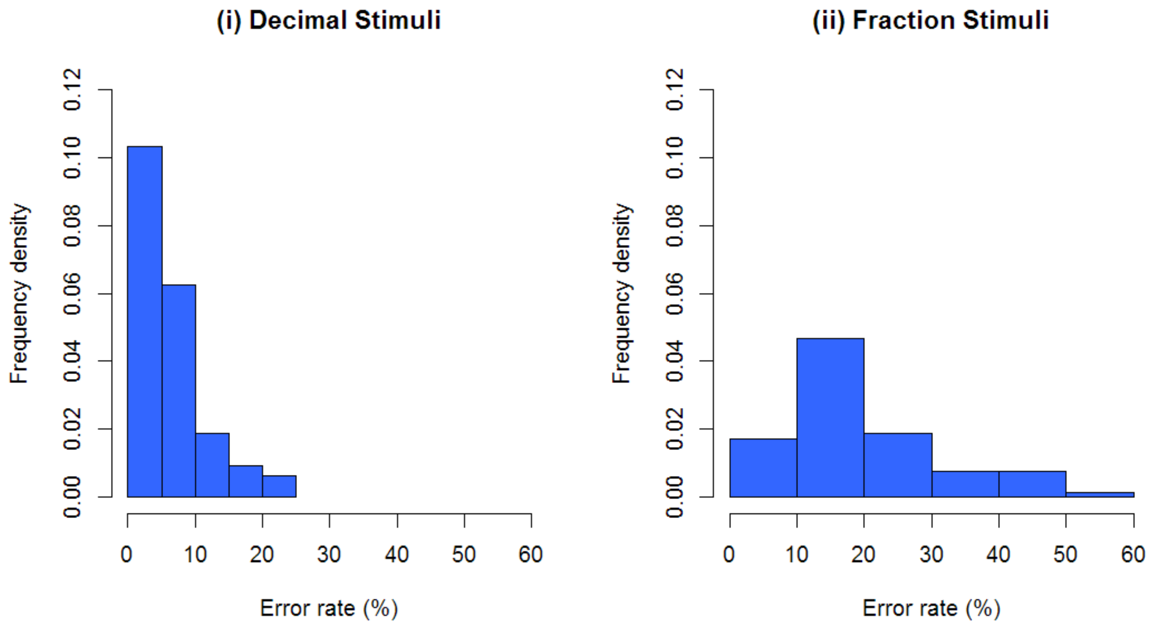

Figure 4.8 Histograms showing distribution of error rates for (i) decimal and (ii) fraction stimuli for experiment three

Scatter diagrams of error rate against stimulus size were plotted for each dataset (figure 4.9). The pattern of errors for decimal stimuli is generally flat but with peaks in error rate around stimulus sizes 0 (zero) and 0.5. It is unlikely participants could not tell whether a decimal was closer to zero or one. So the little peak around 0.5, perhaps is allied to the hesitation shown by the peak in mRTs here. Hesitation that could lead to fluster and pressing the wrong button.

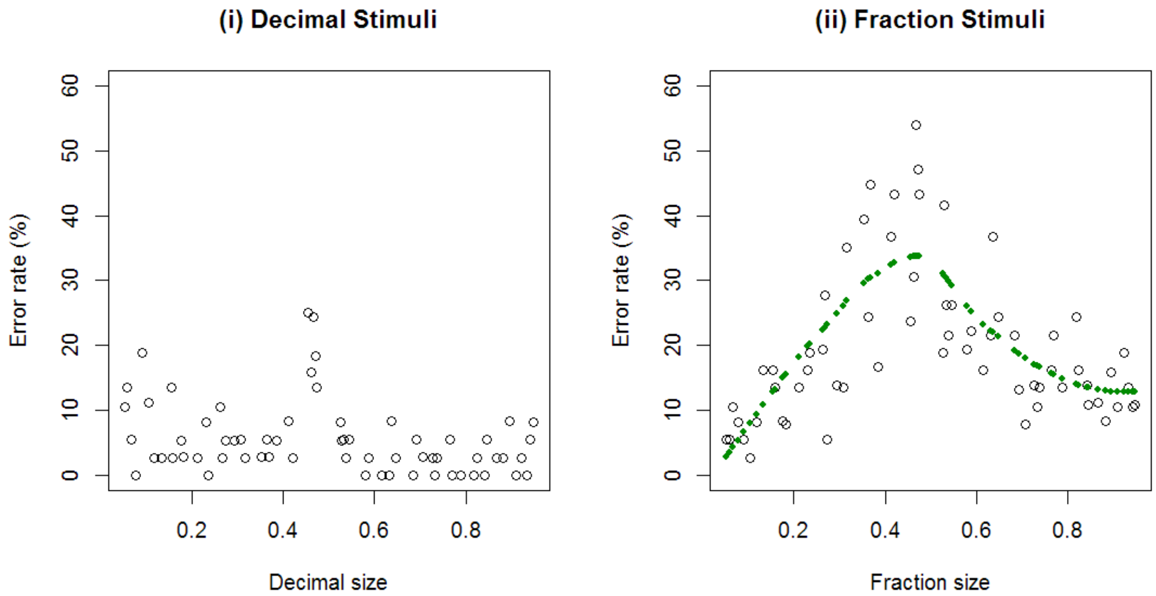

Figure 4.9 Scatter diagrams of error rates against stimulus size for (i) decimal and (ii) fraction stimuli for experiment three

The pattern of errors for fraction stimuli is more useful as participants might have not known the correct answer. It has a distinctive shape with lower error rates around the ends of the zero-to-one range and a steep rise to a peak in errors in the middle of the range, around 0.5. However, this is not an approximately parabolic pattern as demonstrated by the plots of the loess calculations shown in green on the scatter graph of error rates against stimulus size for the fraction stimuli only (figure 4.9(ii)). The lower error rates near to zero and one do mirror the location effect of experiments one and two.

The lack of normal distributions and the complexity of the shapes of the scatter diagrams preclude more in-depth analysis of the errors for fractions and decimals separately.

4.3.2.2.3 Analysis of the difference between fraction and decimal errors

The differences calculated by subtracting the error rate (%) for each decimal stimulus from the error rate (%) for the equivalent fraction stimulus were not significantly skewed, $\gamma = 0.265$, $p = .40$. Additionally, the histogram produced from these differences (figure 4.10) demonstrated an approximately normal distribution.

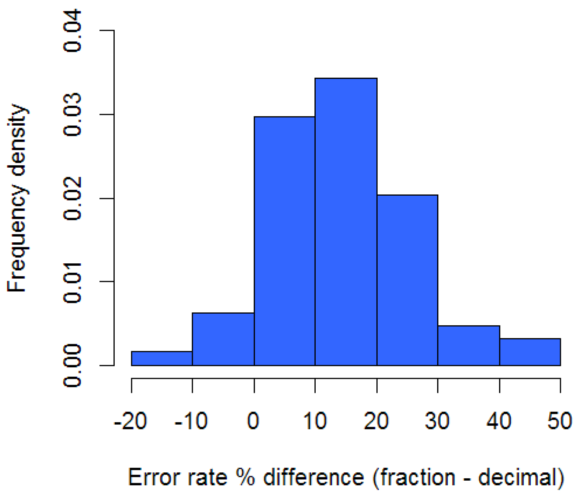

Figure 4.10 Histogram showing distribution of difference in error rates (fraction – decimal stimuli) for experiment three

Therefore a paired t-test was conducted to compare the error rate of decimal and fraction stimuli of the same size. Average % error rates for decimal stimuli (*M* = 5.53, *Mdn* = 2.78, *SD* = 5.71, *range* = 25) were significantly lower than those for fraction stimuli (*M* = 19.5, *Mdn* = 16.2, *SD* = 11.7, *range* = 51.4), with very large effect size; $t(63) = 9.63$, $p < .001$, $d = 1.51$. So significantly more errors were made when the stimulus was a fraction than when it was a decimal.

4.3.2.3 Summary of results

Average RTs to both fraction and decimal stimuli were affected by their size (or position in the zero-to-one range with larger mRTs in the middle of the range than at the ends of the range. However, responses to fraction stimuli were far more affected by the position of the stimulus then those to decimal stimuli. In addition, the pattern of the effect was symmetrical about ½ for fractions but with a peak lower than 0.5 for decimals indicating generally lower RTs for larger decimals.

The pattern of errors also peaked around the middle of the zero-to-one range for both fractions and decimals. For decimal stimuli, other than this peak, there was little pattern to the errors. However the errors for the fraction stimuli demonstrated a graduated rise to the middle peak.

Responses to fraction stimuli were longer and less accurate than responses to decimal stimuli, particularly in the middle of the zero-to-one range.

Taken together, this implies that the task of deciding whether a number is closer to zero or to one is more difficult when that number is a fraction than a decimal. Near to zero and one, there is not a significant difference in the difficulty of the task but as the stimuli move towards 0.5, the task becomes increasingly more difficult for fraction stimuli than for decimal stimuli.

4.4 Experiment four

4.4.1 Method

4.4.1.1 Participants

32 psychology students at the university of Huddersfield, (5 men), participated in the experiment in return for course credit. Their ages ranged from 18.6 to 24.4 years with an average age of 20.5 years and a standard deviation of 1.5 years.

4.4.1.2 Stimuli

Stimuli consisted of relatively unfamiliar fraction or decimal with a magnitude between zero and one. The stimuli were presented, in black on a white screen, one-by-one, in the centre of the screen. Simultaneously, below this, a horizontal line was shown, taking up approximately the central third of the screen. The ends of the horizontal line were marked with two equal-sized, small vertical lines. Beneath the left hand vertical line was a '0'. Beneath the right hand vertical line was a '1'.

The fractions used were the sixty-four fractions with denominators of 11, 13, 15, 17 or 19 that cannot be simplified. The decimals were the decimal equivalents of these sixty-four fractions, rounded to three decimal places. This made 128 stimuli in total.

4.4.1.3 Procedure

The experiment was conducted using E-Prime® 2.0 stimulus presentation software in a lab at the University of Huddersfield. The same computer was used by all participants. Stimuli were preceded by a fixation cross in the centre of the screen. Participants were instructed that within each trial they had to judge the size of the fraction or decimal stimulus and mark it on the number line using a mouse click. The participants were able to see their judgement on the number line as a thin green line that appeared for 500ms after clicking. Once made, participants were not able to change their estimation.

The experimental block was preceded by two preparatory blocks. A practice block of four stimuli first accustomed participants to the procedure and allowed them to pause and ask questions of the experimenter for clarification of the task, if required . Then a calibration block followed in which participants clicked twice on zero and twice on one. The 128 experimental stimuli were then presented in random order with the opportunity to take a break offered half way through.

Response times and the position of mouse clicks in pixels were recorded by the program for the practice, calibration and experimental blocks.

4.4.2 Results

4.4.2.1 Response time analysis

4.4.2.1.1 Pre analysis data processing

As for the analysis of RTs for experiment three, two datasets of average response times were formed. One was the mRTs, across participants, for each of 64 decimal stimuli. The other was the mRTs, across participants, for the 64 fraction stimuli (measurement in ms).

All of the mRTs in each dataset were less than 3 standard deviations away from the mean for that dataset. So none of these data were considered to be outliers.

4.4.2.1.2 Separate analysis of mRTs for decimal and fraction stimuli

Both datasets were found to be somewhat but not significantly positively skewed using the method of Crawley (2005). The RTs for decimal stimuli produced γ = 0.711, p = .240 and for fraction stimuli, γ = 0.451, p = .327. Histograms of the datasets are shown in figure 4.11.

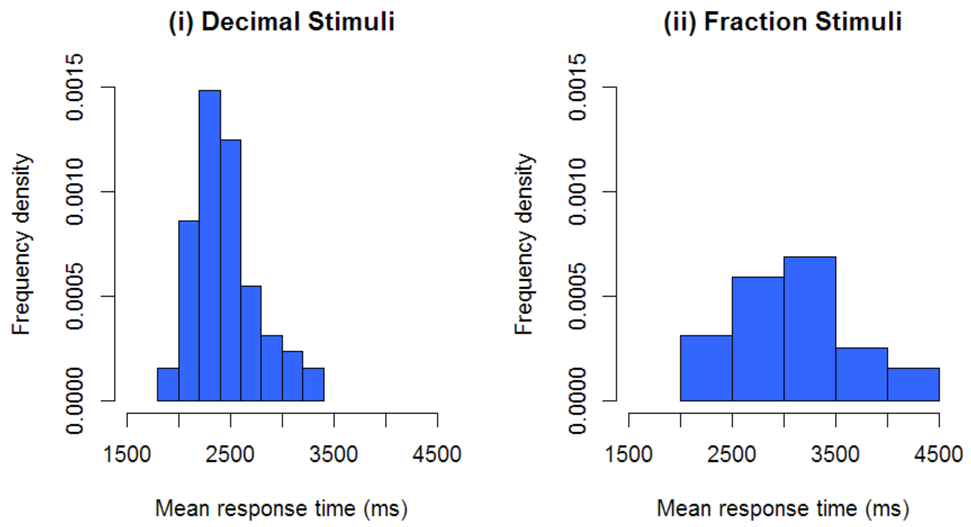

Figure 4.11 Histograms showing distribution of mean RTs for (i) decimal and (ii) fraction stimuli for experiment four

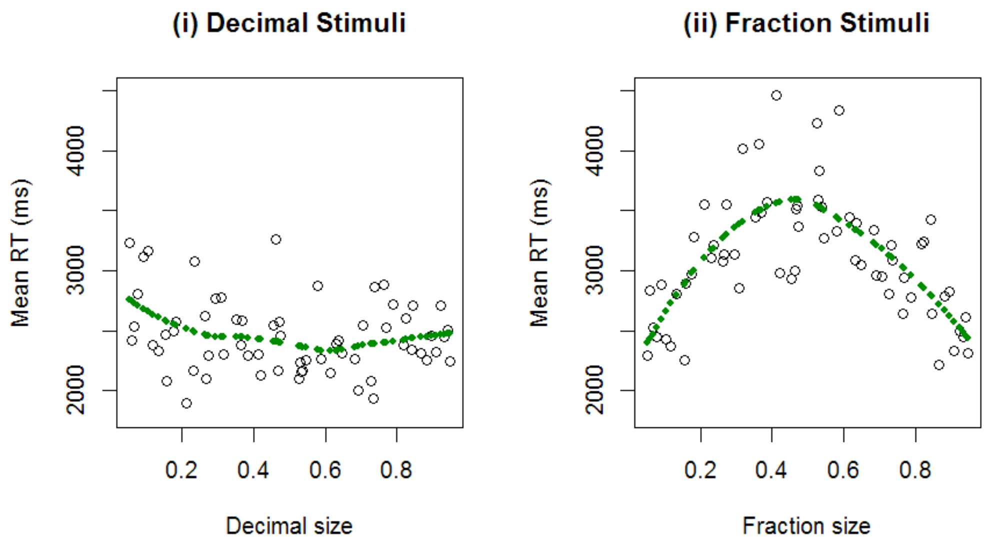

Figure 4.12 Scatter diagrams of error rates against stimulus size for (i) decimal and (ii) fraction stimuli for experiment four

Scatter diagrams were produced of mRT against the size of the stimulus presented for each of the fraction and decimal datasets (figure 4.12). As for experiment three, loess calculations were performed to find smoothed patterns for the data (span = .75, polynomial degree = 2). The results are shown in green on figure 4.12.

For decimal stimuli there was a generally flat shape to the mRTs. They appeared to be fairly consistent across the zero-to-one interval, possibly slightly larger close to zero. This was the opposite of what the largerstim would lead one to expect which is that RTs should be lower for smaller decimals than for larger decimals. Neither a linear (r^2 = .0376, $F(1,62) = 2.42$, $p = .125$), nor a parabolic ($adj.r^2$ = .0732, $F(1,62) = 3.49$, p = .0369) model accounted for any notable proportion of the variance.

For fraction stimuli, the shape of mRT against stimulus size was again not the relatively flat pattern with a sharp drop-off at either end seen for mRT against position in experiment one (figure 2.4). There was instead, the same approximately parabolic shape as seen in experiment three. Mean RTs were greater in the middle of the zero-to-one range than at the ends but the transition between the middle and ends was relatively smooth.

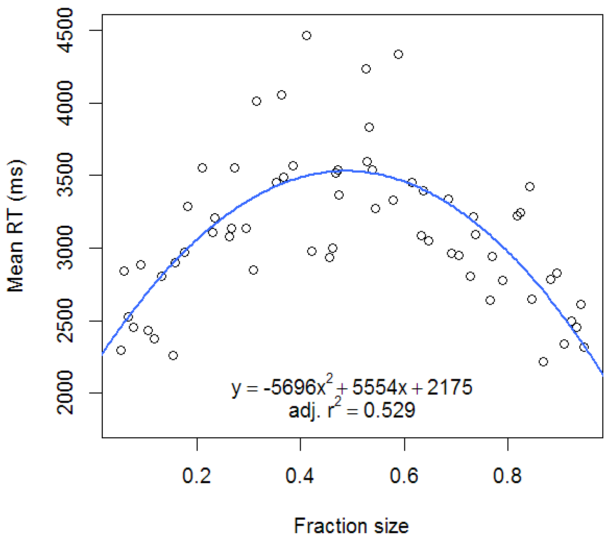

Figure 4.13 Parabolic models of mean RT for fraction stimuli for experiment four

A multiple linear regression was fitted using fraction size and (fraction size)^2 as predictors of mRT. A significant regression equation was found ($F(2,61) = 36.5$, $p <$ .001). Both fraction size ($t = 8.05$, $p < .001$) and (fraction size)^2 ($t = -8.49$, $p < .001$) were significant predictors of mRT. The model as well as the adjusted r^2 value are shown in figure 4.13. Details of the coefficients are in table 4.3 along with confidence intervals for the model coefficients found by bootstrapping with 5000 replications.

The bootstrapped results differ only very slightly from those found by regression analysis. Also, the residual plot for the model (figure 4.14) has a suitably random pattern, approximately around zero. This implies that the multiple linear model with fraction size and $(\text{fraction size})^2$ as predictors of mRT is suitable for these data.

Coefficient	Estimate	SE	95% CI (model)	Bootstrapped 95% CI
intercept	2175	149	(1883, 2466)	(1928, 2409)
fraction size	5554	690	(4202, 6906)	(4388, 6872)
$(\text{fraction size})^2$	-5696	671	(-7011, -4382)	(-7003, -4543)

Table 4.3 Details of the coefficients of the parabolic model for fraction stimuli for experiment four

The adjusted r^2 value indicates that the model accounted for 52.9% of the variance within mRTs. The model reaches its maximum only slightly to the left of the centre of the zero-to-one range (at 0.488).

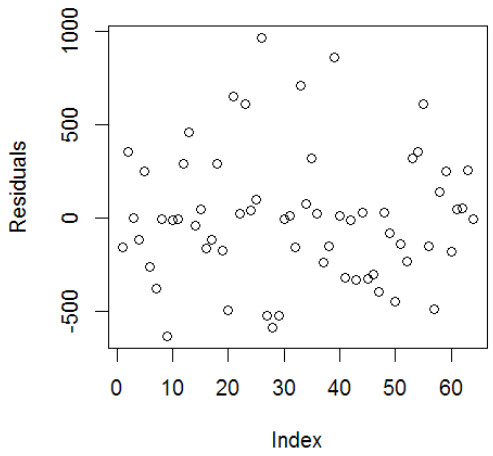

Figure 4.14 Plot of residuals for parabolic model for fraction stimuli for experiment four

Again, a Kruskal-Wallis test was conducted on mRTs for the fraction stimuli to check whether there was evidence of any difference in the average mRT results for the five different denominators (11, 13, 15, 17 & 19). No significant difference was found $\chi^2(4) = 1.92$, $p = .750$.

4.4.2.1.3 Analysis of the difference between fraction and decimal mRTs

Just as for experiment three, mRTs for fraction stimuli appeared to be generally longer than for decimal stimuli. The paired differences calculated by subtracting the mRT for each decimal stimulus from the mRT for the equivalent fraction stimulus were not significantly skewed, $\gamma = 0.148$, $p = .441$. The histogram produced from these differences (figure 4.15) demonstrated an approximately normal distribution.

Therefore, a paired t-test was conducted on these paired differences. Average response times to decimal stimuli (M = 2458ms, SD = 2254ms) were significantly lower than those to fraction stimuli (*M* = 3100ms, *SD* = 2787ms), with very large effect size; $t(63) = 7.50$, $p < .001$, $d = 1.51$.

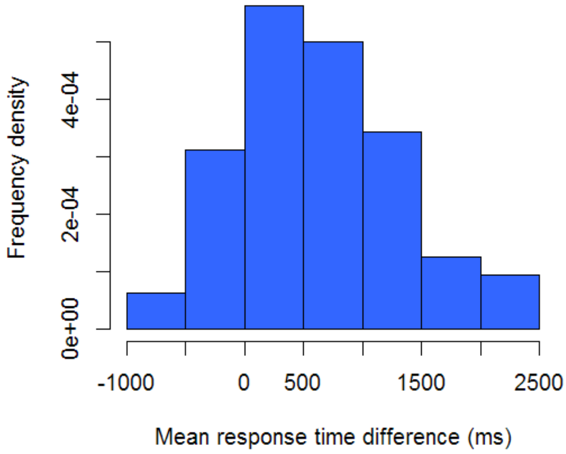

Figure 4.15 Histogram showing distribution of difference in mean RTs (fraction – decimal stimuli) for experiment four

The pairwise (fraction – decimal) differences in mRT were potted against stimulus size (figure 4.16). Loess calculations were performed to find a smoothed pattern for the points which again demonstrated an approximately parabolic shape (shown in green). The largerstim effect led to the expectation that this difference would be greater near to zero than near to one. However, the curve appears to be approximately symmetrical with the turning point in the centre of the zero-to-one range.

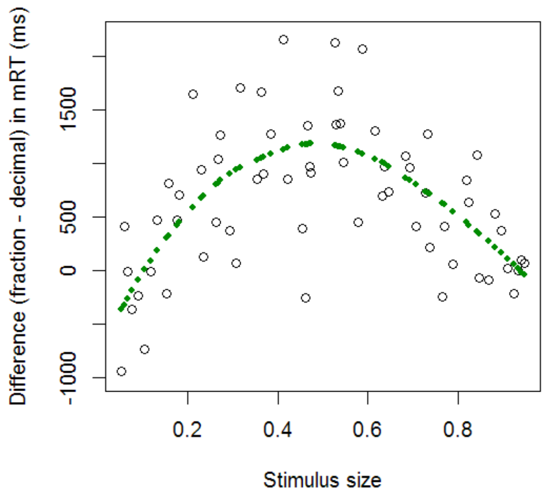

Figure 4.16 Scatter diagram of difference in mean RT (fraction – decimal stimuli) against stimulus size for experiment four

Again, a multiple linear regression was fitted using stimulus size and (stimulus size)^2 as predictors of mRT differences. A significant regression equation was found ($F(2,61)$ =

24.32, p < .001). Both stimulus size (t = 6.85, p < .001) and (stimulus size)^2 (t = -6.97, p < .001) were significant predictors of mRT differences. Details of the coefficients are shown in table 4.4.

Coefficient	Estimate	SE	95% CI (model)	Bootstrapped 95% CI
intercept (not significant)	-603	220	(-1034, -172)	(-1003, -181)
stimulus size	6973	1018	(4978, 8968)	(4980, 8909)
(stimulus size)^2	-6897	990	(-8837, -4957)	(-8747, -5065)

Table 4.4 Details of the coefficients of the parabolic model for differences in mean RT for experiment four

The residual plot for the model (fig 4.17) has a suitably random pattern around zero with just one potential negative outlier (at decimal 0.462, fraction $^6/_{13}$). Confidence intervals for the model coefficients were also found by bootstrapping with 5000 replications, shown in table 4.4. The results differ only very slightly from those found by regression analysis.

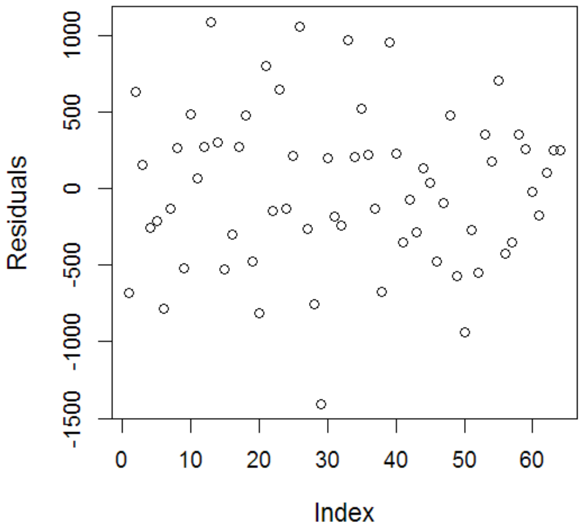

Figure 4.17 Plot of residuals for parabolic model for differences in mean RT for experiment four

These findings imply that the multiple linear regression model using stimulus size and (stimulus size)^2 as predictors of mRT difference was appropriate. The adjusted r^2 value indicates that the model accounted for 42.5% of the variance within mRT differences.

The regression model with the associated adjusted r^2 is shown on figure 4.18, graphically depicted as a solid blue line, with the 95% CI for the model shown as red dotted lines. The 95% CI is completely above the difference = 0 line on the interval (0.383, 0.629), implying that for these stimulus sizes, responses to fraction stimuli were significantly larger than they were to their equivalent decimal stimuli. However, nearer to the end points of zero and one, the model dips below the difference = 0 line implying that here responses to fraction stimuli were smaller than they were to their equivalent decimal stimuli.

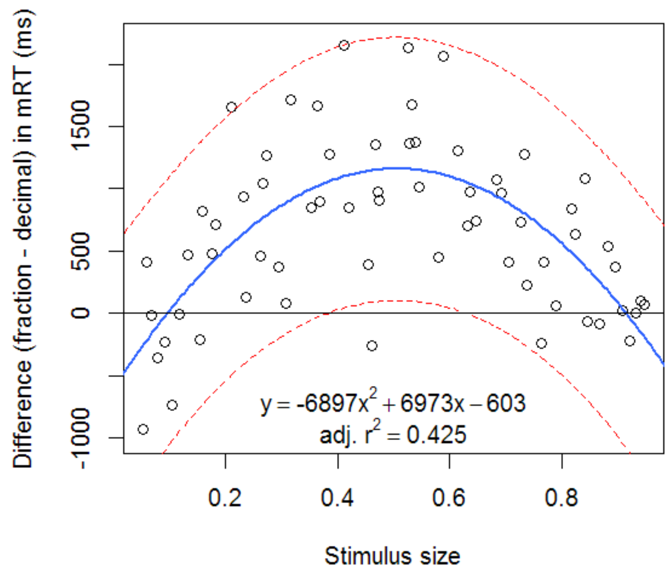

Figure 4.18 Parabolic model of mean RT for differences in mean RT with 95% CI for experiment four

It also can be seen that the model reaches its maximum only slightly to the right of the centre of the zero-to-one range (at 0.506). These outcomes taken together imply that within the middle of the range it is significantly harder to judge the size of a fraction than it is to judge the size of a decimal. At *both* ends of the range it is not significantly harder to make this judgement for a fraction than a decimal, in fact the opposite might be true.

4.4.2.1.4 Comparison between experiments three and four of fraction mRTs

Response times for fraction stimuli in experiment four appeared to be generally longer than for experiment three. Therefore, a paired t-test was conducted on the paired differences of fraction mRTs (experiment four – experiment three). Average response times to fraction stimuli in experiment three (M = 2134ms, SD = 286ms) were significantly lower than those to the same stimuli in experiment four (M = 3100ms, SD = 2787ms), with very large effect size; $t(63) = 18.9$, $p < .001$, $d = 2.33$).

The differences between mRTs for fraction stimuli in experiments four and three ranged between 331ms and 2115ms (M = 966, SD = 408).

There did not appear to be a particular pattern to the differences across the zero-to-one range. Loess calculations showed a slight rise and then fall in differences so a parabolic model was fitted. A significant fit was found, ($F(2,61) = 6.43$, $p = .003$), with both stimulus size (t = 3.22, p = .002) and (stimulus size)^2 (t = -3.50, p < .001) significant predictors of mRT differences. However, the adjusted r^2 value was only .147 implying that only 14.7% of the variance in mRT differences could be accounted for by the model. This implies that the effect of fraction size (or position) upon the difference in mRTs for fraction stimuli in experiments four and three was small in size.

4.4.2.2 Accuracy analysis

4.4.2.2.1 Calibration and pre-analysis data processing

Participants' calibration results were collated and used to locate the zero and one ends of the estimation number line. For both the zero and one calibration clicks, the mode, median and mean pixel locations were the same; with zero placed at 159 pixels and one at 471 pixels. These were very consistent between participants and therefore taken to be reliable locators of the end points of the estimation line. It followed that each pixel between these end points was representing a distance of approximately 0.003205 on the number line.

This was not ideal as the decimal numbers presented were given to three decimal places but participants had not been given the facility to respond with that degree of accuracy. Nevertheless, it was considered a reasonable degree of accuracy for the purposes of detecting the general shape of responses and any differences in the shape of responses between decimal and fraction stimuli.

Participants' responses were therefore converted from pixels into decimal values. Then for the purpose of analysis of the accuracy of estimation, the average response size was calculated, across participants, for each of 64 decimal stimuli and then again for each of the 64 fraction stimuli. They were not significantly skewed for either the fraction stimuli (γ = -0.364, p = .641) or the decimal stimuli (γ = -0.330, p = .629). Histograms for each of these datasets (not included herein) demonstrated reasonably normal distributions.

4.4.2.2.2 Separate analysis of accuracy for decimal and fraction stimuli

Average estimations (mean response size) were plotted against stimulus size for both the decimal and fraction stimuli and linear regression lines calculated for each. These are shown in figure 4.19 (i) and (ii). The target line y = x is shown for comparison.

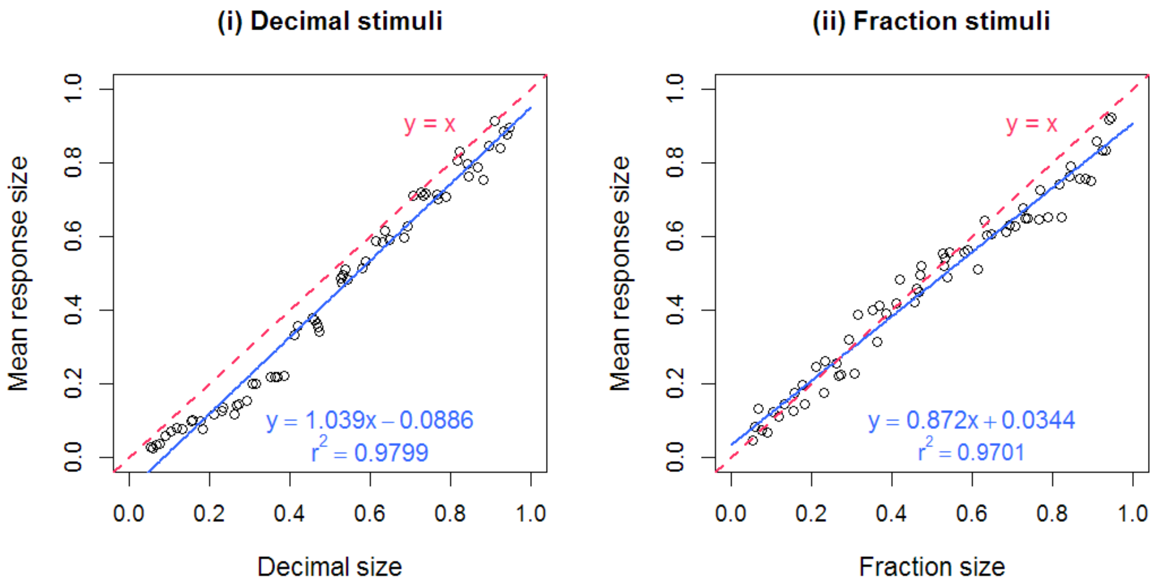

Figure 4.19 Linear models for mean response size for (i) decimal and (ii) fraction stimuli for experiment four

It can be seen that the linear regression lines were a very good fit to both the decimal and the fraction data. Indeed, the linear correlation between stimulus size and mean response size was very large for both decimals (r = .990, $t(62)$ = 55.5, p < .001) and fractions (r = .985, $t(62)$ = 45.2, p < .001).

4.4.2.2.3 Comparisons of accuracy between fraction and decimal stimuli

A paired t-test showed that the responses to decimal stimuli were generally underestimates of the true value ($t(63)$ = -13.1, p < .001, d = 0.244) by approximately 0.07 on average. It should be noted that these are under- estimates of approximately 0.02 even based on solely their leading decimal. A t-test of mean response against leading decimal yielded $t(63)$ = -3.96, p < .001, d = 0.067).

The responses to fraction stimuli appear to be approximately correct or slightly over estimated between zero and roughly 0.6 and then slightly underestimated between 0.6 and one. A generally accurate/over- and then under-estimate would lead to the decrease in the gradient of the regression line from the expected value of 1 to 0.872.

An over- and then under-estimation (inverse ogival) pattern of proportion estimations was seen by Varey et al. (1990). So paired t-tests were carried out. First, for stimuli below 0.6 in size, no significant error in the size of estimation was found ($t(38)$ = 0.238, p = .813). Stimuli above 0.6 in size, however, were found to be significantly underestimated ($t(24)$ = -9.08, p < .001, d = 0.733) by approximately 0.08 on average.

A paired t-test comparison between fractions and decimals of the absolute difference between estimate and target number showed that estimates of decimal stimuli were significantly less accurate than of their equivalent fraction stimuli ($t(63) = 2.79$, $p = .003$) and that this effect was of medium size ($d = 0.527$).

4.4.2.3 Summary of results

Average RTs to fraction stimuli were affected by their size (or position in the zero-to-one range) with larger mRTs in the middle of the range than at the ends of the range. The pattern of the effect was symmetrical about approximately ½ for fractions.

The size/position of decimal stimuli within the range had a far less clear-cut effect on RTs. However, it was clear that responses to fraction stimuli were longer than responses to decimal stimuli, particularly in the middle of the zero-to-one range.

For both fractions and decimals, the size of estimation had a very strong linear association with the size of the stimulus presented. However, there was a systematic underestimation of decimal stimuli and estimations of fraction stimuli were therefore more accurate generally.

Taken together, this implies that estimation of the size of fractions is a more lengthy and deliberate procedure than is the estimation of the size of a decimal, particularly as the stimuli move away from zero and one towards 0.5. Yet this lengthier procedure yields a more accurate result (at least for numbers below 0.6).

The factor of denominator did not confound the results for either experiment.
Average RTs to fraction stimuli were significantly much greater in experiment four than they were in experiment three. The difference was not greatly affected by the size of the fraction.

4.5 Discussion

4.5.1 Contrasting the results for fractions and decimals

Decimal magnitudes were accessed more quickly than fraction magnitudes. Their proximity to the end points of zero and one was more accurately and quickly accessed too. These differences were significant in the middle of the zero-to-one range only.

Similarly, DeWolf et al. (2014) also found RTs for fractions were significantly greater than for 3-digit decimals. There's were magnitude comparison tasks in which they presented very similar stimuli to mine. However, unlike in my experiments one and two, the fractions and three-digit decimals were separately, and not directly, compared. They concluded that the increased RTs for fractions implied that representations of fraction magnitudes are more fuzzy in nature and less easy to access than those of decimals.

This is a conclusion supported by the longer RTs and greater error rates for fractions than for decimals in experiments three and longer RTs for experiment four. However, I also found that this effect was mediated by the location of the stimuli within the zero-to-one range. Close to zero and close to one there were no differences in RT for fractions and decimals. Conversely, I found that fraction *estimates*, though they took longer, were more accurate. The underestimation of decimals was systematic across the whole of the zero-to-one range.

Iuculano & Butterworth (2011) did not analyse the pattern of RTs in terms of the location of the stimulus in the zero-to-one interval, (their experiments were more concerned with the differences between adult's and children's estimations of different types of stimuli). However, they too had generally longer RTs for fraction than decimal stimuli in their NP task for their adult participants.

Both fraction and decimal estimations in my experiment four had an approximately linear relationship to their target number. In their NP task, Iuculano & Butterworth also presented their participants with fractions and decimals to estimate and mark on a number line. They used a smaller set of fractions and decimals which included familiar number, such as $^{1}/_{4}$ and more unfamiliar numbers like $^{7}/_{9}$. Despite the different numbers used, they got very much the same kind of results for accuracy of estimation. That is, linear models were a very good fit to the pattern of mean estimate against stimulus size for *both* fractions and decimals. However, they did not discuss whether fractions or decimals had any systematic under- or over- estimation.

4.5.2 Patterns of estimation in experiment four

In the estimation task of experiment four, different patterns of size estimations were observed for fractions and decimals. Both results differed from the cyclical over- and then under-estimation, inverse ogival pattern for judgements of spatial proportion found by e.g. Varey et al. (1990), Hollands & Dyre (2000) or of numerical interval estimations found by e.g. Karolis, Iuculano & Butterworth (2011). If this had been observed it would have implied that judgements had been biased away from the end points of zero and one. Conversely an under-and then over- estimation pattern would have implied judgements biased towards the end points.

However for decimals there was an markedly linear pattern of systematic slight under-estimation that is consistent with the dominance of the leading decimal. For fractions there was some evidence of underestimation for the larger fractions but generally estimates were very accurate. Neither of these patterns demonstrates any influence of the theoretical anchor points of zero and one upon estimation accuracy. So the patterns in the accuracy of estimation do not support my theory of the use of zero and one as anchor points in size judgements about fractional numbers.

The systematic underestimation of the decimal stimuli implies that participants were overly influenced by the leading decimal digit, neglecting the rest of the number, rather than rounding to one decimal place. However, the estimates of decimals were also an underestimation of their leading decimal digit. This combined with the greater accuracy of estimation for fraction magnitude estimations seems at odds with the findings of DeWolf et al. (2014).

They found faster and more accurate magnitude comparisons between pairs of decimal stimuli than pairs of fraction stimuli. They concluded that mental magnitude representations of fractions were more *fuzzy* and less easily accessed than those of decimals. It could be argued that I have found the opposite in this estimation task.

Systematic underestimation is consistently found when people are asked to bisect a line segment (see Jewell & McCourt, 2000 for a review and meta-analysis). This is the same task as marking the decimal 0.5 or familiar fraction ½ on a line. Longo & Lourenco (2007) saw the same type of underestimation in the bisection of pairs of whole numbers. The conclusion was of the leftward spatial bias being mirrored as a similar numerical bias. In other words, a numerical bias that supports the mental number line theory for whole numbers because it matches a known spatial bias.

By giving participants a spatial task to do – mark numbers on a line, spatial biases could have been elicited. They might have had a stronger affect on the decimal stimuli precisely because they are *less* fuzzily mapped onto a spatial mental number line. This could have been the reason for the systematic underestimation of decimals even against their leading digit.

4.5.3 Location effect and anchor points

It is important to note that the anchor points of zero and one were explicit in both experiments. In experiment three they were asked to judge against these two numbers. In experiment four, both zero and one were visible on the screen at all times and participants were asked to click on each three times at the start of the experiment for calibration.

Results of both experiments three and four indicate faster and more accurate magnitude judgements near zero and near one than in the middle of the range. However these differences were more gradual than the location effect seen in experiments one and two. Nevertheless, the pattern of results for RTs and errors still imply that size judgements, particularly for fractions, are more difficult in the centre of the zero-to-one range.

This bias in the RTs towards zero and one supports the theory that these numbers are used as anchors for judging the magnitude of fractions, if not decimals (Tversky & Kahneman, 1974). RTs increasing as increasing adjustments are made away from the anchor points.

However, the pattern of RTs might also be interpreted as a bias away from the centre of the range. This could mean that judgements of the magnitudes of fractions are automatically made against ½. Effectively a distance effect would then produce a decrease in RTs as stimuli move away from ½. It would be interesting to know if the same pattern of RTs would have emerged if participants had been primed to make that judgement. That is, if they had been asked in experiment three "Is this number greater or smaller than ½?" or had been presented with a number line with only ½ marked in experiment four.

The shape of the location effect of experiments one and two, particularly at the smallest distances of 0.05, 0.1 and 0.2,argues against this interpretation. For those distances, there were no significant differences in RTs to the positions B to F (see Chapter two, pages 33 to 36). This is why they were amalgamated into a single *middle* location. It was only for the extreme end positions that any differences in RTs were found.

The fact that there was far more influence of position in the zero-to-one interval on RTs on fraction stimuli than there was on decimal stimuli implies that the location effect on the magnitude comparison tasks of experiments one and two was due more to the fraction in the stimulus pair than the decimal. Yet the smooth parabolic shape of RTs seen for fraction stimuli in both experiments three and four is not the same as the shape of the location effect in experiments one and two. Of course, in a magnitude comparison task there is the interference between the two numbers being compared. That is evident in the distance effect.

However, RTs to fraction stimuli were the same size or even perhaps shorter than to decimal stimuli very close to zero and one. The differences between fractions and decimals were only evident in the middle of the range.

4.5.4 Largerstim effect

The systematic underestimation of the size of decimals in experiment four might throw some light on the largerstim effect seen in experiments one and two.

It has been shown that decimal fractions are processed componentially i.e. decimal place by decimal place (Kallai & Tzelgov, 2014). This, along with the systematic decimal underestimation implies that the estimates of the size of decimal stimuli were based mainly on their leading digit alone. Though they were still generally an underestimation of this too.

In the magnitude comparison task, if the decimal is the smaller of the stimulus pair, then an underestimation of its size (based on its leading digit) would not be problematic. However, if the decimal is the larger number, its underestimation could lead to errors being made, (particularly at smaller distances). Part of the largerstim effect was indeed a significantly greater number of errors when the larger stimulus was a decimal than when it was a fraction.

The other part of the largerstim effect was a similar increase in RTs when the larger stimulus was a decimal. The RT analysis for experiments three and four above really did not throw any light on this effect. The RTs for fractions were symmetrical about ½ in both experiments. The RTs for decimals were basically consistent across the zero-to-one range for experiment four and if anything, slightly shorter for higher decimals for experiment three. This is exactly the opposite of what one might expect to see which would be longer RTs when making size judgements about increasing larger decimals.

However, the underestimation of the size of decimals and the implication that their size was being judged predominantly on their leading digit does provide a possible explanation of the largerstim effect on RTs. In experiments one and two, participants might have tried first to make a comparison based on the leading decimal alone. If that comparison was not clear-cut enough for them they would have to *take more time* to look at further decimal places before coming to a judgement on which was the smaller or larger.

This need to look at more decimal places would occur more often when the decimal was the larger number because of the inherent rounding-down effect of looking only at the leading decimal. This *could* be an explanation for the longer RTs when the decimal was the larger of the stimulus pair.

4.5.5 Reported strategies from experiment one

There is one final problematic outcome of experiment one that can be addressed by the results, specifically of experiment three.

There was an apparent discrepancy in experiment one between participants' reported methodology and the results obtained. The majority of participants in experiment one claimed to have made their judgements by testing which side of 0.5 the two stimuli were. Only 26.7% of the stimulus pairs were either side of 0.5. There was no evidence that it was easier to judge the larger in these than in other pairs. Indeed, these stimulus pairs were mostly in the slower middle location. Thus, I judged that they were not aware of their own strategy.

The judgement of whether a number is one side or the other of 0.5 is effectively the same as judging whether it is nearer to zero or to one. This was the task in experiment three. Therefore the results of experiment three imply that judging whether either a fraction *or a decimal* is above or below 0.5 takes the longest in the middle of the zero-to-one range. Indeed, for fractions the RT reaches it maximum pretty much at 0.5.

The participants in experiment one were not reporting strategies on a trial-by-trial basis. They were asked to recall methods they had used after the experiment was over. As such, the methods that might have been the most salient were the ones they had deliberated over the longest. That is the most difficult judgements, close to 0.5.

4.5.6 Next steps

The next chapter of this thesis is not concerned with the differences between fractions and decimals. It revisits the results of experiments one and two and considers individual differences in more depth. Specifically in the three main effects of distance, location and largerstim. The links between these effects in individuals and their success at the number comparison task are investigated.

Chapter 5 Individual differences in experiments one and two

5.1 Introduction

In this very short chapter, experiments one and two are revisited. First is an argument for reconsidering the outcomes of these experiments in terms of individual differences. Then the statistical methodology used for investigating and quantifying individual differences is detailed.

The next part of the chapter details the analysis made. In particular, analysis of the associations between patterns in participants' response times with their accuracy at the task.

The chapter finishes with a short discussion of the implications of the analysis of individual differences.

5.2 Individual differences in experiments one and two

During the analysis of the results of experiment one and experiment two, significant individual differences for participants in average response time were found (pages 36 and 59). This resulted in the use of mixed linear models for the analysis of RTs, with random intercepts for participants. It was also demonstrated that there were significant differences between participants in the number of errors made (pages 43 and 66). Individual participants gave as few as 3 to as many as 45 incorrect responses out of the 120 trials in experiment one. In experiment two, the number of errors ranged from 3 to 61 out of 144 trials.

So individuals significantly differed in their baseline mRTs and error rate. General effects on mRT were found for the factors of distance, location and largerstim. The distance effect was sought and expected. Additionally, the possibility that zero and one would act as anchor points for magnitude estimation meant that some effect of location was not unexpected.

However, the effect that RTs would be longer and error rates greater when the larger of the stimulus pair was a decimal and the smaller a fraction was not expected. Experiments three and four highlighted some variability in the pattern of RTs and errors across the zero-to-one range, particularly for fractions and decimals. These differences did offer some grounds for speculating on the cause of the largerstim effect. However, there was no conclusive reason to explain the result.

Taking all of the above into consideration, I was prompted to look more deeply at individual differences in participants' results. More particularly, to look at whether there were large variations in their distance, location and largerstim effects and whether the size of these effects bore any relation to participants' ability at the number comparison task.

These distance, location and largerstim effects might be considered *cognitive indicators* for individuals. That is, they might allow some insight into the cognitive representations of small numbers within individual participants' minds. After all, the distance effect is classically used as an indication of number magnitudes being internally represented on a mental number line. If there are links between the strength of these cognitive indicators in individuals and their success at the number comparison task, this could throw light on what kinds of mental representations of fractional numbers are most effective.

5.3 Method

5.3.1 Verifying that individuals differed significantly

To investigate whether it might be the case that individuals were also significantly varied in the nature of their distance, location and largerstim effects, the final mixed linear models of logRT for experiments one and two were revisited.

Variable slopes were added to these models for distance by participant, location by participant and largerstim by participant. Each factor was added separately from the others. These models would have been contained far too many variables to be considered useful. Therefore this procedure was only carried out to demonstrate that individual differences in the results of experiment one and two were worth exploring.

5.3.2 Quantifying individual differences

Next, for each participant of the two experiments an ANOVA tests was run on logRT for the three factors of distance, location and largerstim. Interactions were not included for simplicity. These analyses yielded both the significance and the partial η^2 effect size of the distance, location and largerstim effects for each individual.

The partial η^2 effect size measures the percentage of the variance within responses that might be explained by a particular factor. As such, they were used as a measure of the three cognitive indicators for each individual. The greatest sum of these three partial η^2 values for any individual in either experiment was 61.2% and the least was 0.8% (*M* = 19.0, *SD* = 12.8).

The partial η^2 values were recorded as positive if the effect for that participant was in the expected direction and negative otherwise. The expected direction for the distance effect was decreasing RT with increasing distance (tested via Spearman's r for each participant); for the largerstim effect, was larger median RT for trials in which the decimal was the larger stimulus; for the location effect, was median RT highest in the middle location larger. These measures were then used to make comparisons between the strength of these three cognitive indicators and participants' success in the experimental task.

5.4 Results

5.4.1 Experiment one results

Variable slopes for distance by participant were added to the final logRT mixed linear model and were shown to make a significant improvement to the fit of the model (L Ratio 86.39, $p < .001$). The same result was found for the location (L Ratio 56.11, $p < .001$) and largerstim (L Ratio 14.61, $p < .001$) factors.

These results demonstrated that there were significant differences between the strength of individuals' distance, location and largerstim effects. Hence ANOVA tests were run on logRT for each participant for the three single factors of distance, location and largerstim. The results of these analyses can be found in appendix 1 (page 157).

A significant distance effect on logRT was found for 26 of the 30 participants (87%). Three of the four who did not demonstrate a significant distance effect were, by far, the least successful participants at the task; making 45 (38%), 43 (36%) and 38 (32%) errors each (as compared to a range of 3 to 27 errors for the remaining participants). This observation already implied that the distance effect, might have been associated with improved success at this comparison task. All participants except one (number 7) demonstrated a negative association between distance and RT. Participant 7 had an insignificant but positive distance effect and made 38 errors.

Importantly, no speed-accuracy trade-off was found for this task. Those who were faster on average, were not also significantly less accurate (Spearman's *r* = -.174., *p* = .178 for median RT against total errors). This indicates that accuracy and RTs were independent. The distance effect is an effect upon response times, specifically that response times decrease as distance increases. The independence between accuracy and RTs was necessary, therefore, for the investigation to take place into whether the strength of the distance effect was associated with the number of errors.

A large, significant negative association was found between the size of the distance effect, (participants' partial η^2) and the number of errors made by individuals (Spearman's *r* = -.659, *p* < .001). This association is illustrated in figure 5.1(i). The inference to be drawn from this outcome is that participants with a stronger distance effect and so, presumably a stronger magnitude element to their mental representations of fractional numbers, were more successful at the task.

The effect of location on logRT was significant, (p < .05), for only 18 participants of the 30 (60%). However, no participants at all were faster in the middle location than the other locations. A marginally significant, small, negative association was found between the size of the location effect and the number of errors for individuals (Spearman's *r* = -.289, *p* = .060), illustrated in figure 5.1(ii).

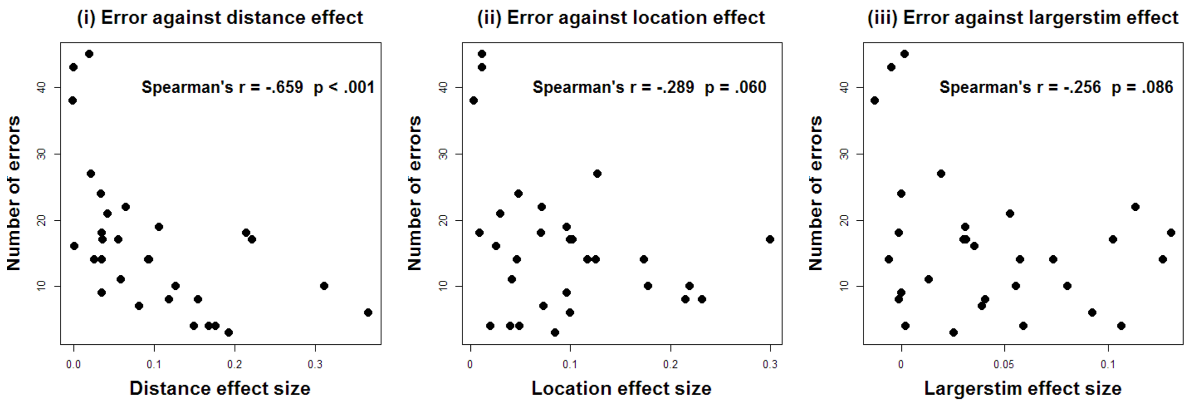

Figure 5.1 Scatter diagrams of number of errors against distance, location and largerstim effect sizes for experiment one

The effect of largerstim was significant (p < .05) for only 15 participants of the 30 (50%). Additionally, six of the participants had the reverse largerstim effect. These were participants {7, 10, 15, 16, 17 , 18}; none of whom had a significant location effect. A marginally significant, small negative association was found between the size of the largerstim effect and the number of errors for individuals (Spearman's *r* = -.256, *p* = .086), illustrated in figure 5.1(iii).

The last two results imply an intriguing possibility that the location effect and largerstim effects are also manifestations of features of the mental representation of small numbers that are useful for the task of comparing magnitudes. However, the results were marginal and lacked consistency. All participants performed better than chance at the task and so could be considered to have sufficiently functional mental representations of fractions and decimals to complete the task successfully. The lack of consistency across participants of the largerstim and location effects implies that these effects might not be highlighting functionally necessary features of the mental representation of small numbers.

However, one further detail to note is that only three participants in experiment one did not demonstrate a significant distance, location or largerstim effect. These were the three participants that produced the most errors (see appendix 1, page 157).

5.4.2 Experiment two results

There were separate final mixed linear models for the two response groups of experiment two. So these two groups were initially analysed separately for individual differences.

For the response-smaller group, variable slopes first for distance by participant were added to the final logRT mixed linear model and were shown to make a significant improvement to the fit of the model (L Ratio 48.7, $p < .001$). The same result was found for the location (L Ratio 44.1, $p < .001$) and largerstim (L Ratio 15.0, $p < .001$) factors.

For the response-larger group, the same significant result was found for the addition of variable slopes first for distance by participant (L Ratio 59.7, $p < .001$). However, due to the complexity of the model, it would not converge when variable slopes for location by participant were added. The addition of variable slopes for largerstim by participant did not significantly improve the model (L Ratio 4.17, $p = .125$).

Though not completely consistent, these results were sufficient evidence to investigate the effects of individual differences on responses for experiment two. Therefore, as for experiment one, ANOVA tests were run on logRT for each participant for the three single factors of distance, location and largerstim. The results of this analysis can be found in appendices 2a and 2b, pages 159-160.

A significant, ($p < .5$) distance effect on logRT was found for 22 of the 28 (79%) participants in the response-smaller group and for only 20 of the 30 (67%) participants in the response-larger group. Again, only one participant (number 48) demonstrated a reverse distance effect but it was very much not significant.

A large significant negative association was found between the size of the distance effect and the number of errors made by individuals (Spearman's *r* = -.770, *p* < .001). This association is illustrated in figure 5.2(i). Again, to support the validity of this analysis, no speed-accuracy trade-off was found for experiment two. That is, there was no significant association between median RT and the number of errors made (Spearman's *r* = -.002, *p* = .495).

The effect of location on logRT was significant (*p* < 0.05) for only 13 participants of the 28 (46%) in the response-larger group, though a further 7 were marginally significant (*p* < .10). Only 16 out of the 30 (53%) participants in the response-smaller group demonstrated a significant effect of location on logRT (with 2 marginally significant). Six participants were not slowest in the middle location. These were participants {7, 14, 34, 43, 50, 57}; none had a significant location effect, though participant 43's was marginally significant.

For experiment two, a significant, medium sized, negative association was found between the size of the location effect and the number of errors participants made (Spearman's *r* = -.362, *p* = .003, see figure 5.2(ii) for illustration). However, there was one apparent outlier (participant 32) with, by far, the largest location effect but also a large number of errors as can be seen on figure 5.2(ii).

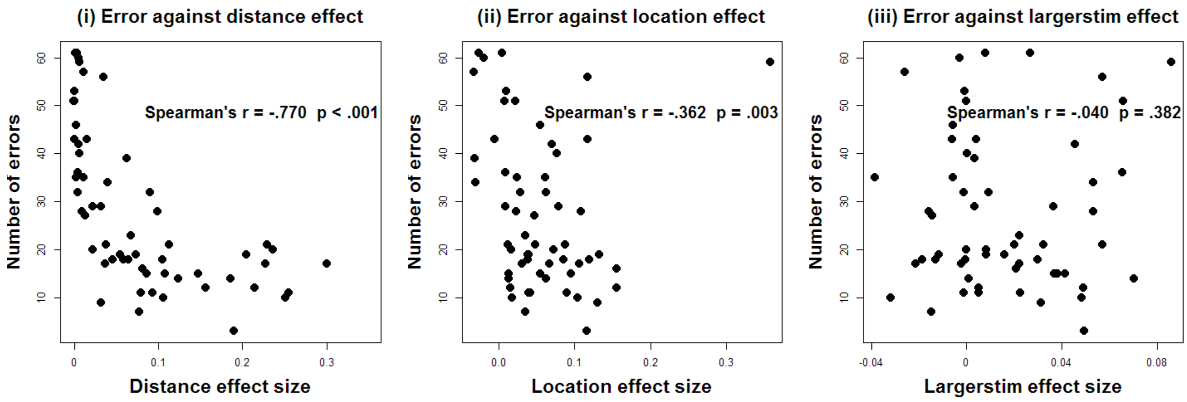

Figure 5.2 Scatter diagrams of number of errors against distance, location and largerstim effect sizes for experiment two

Nevertheless, in general, the greater the difference between RTs in the middle locations and those in the other locations, the fewer errors were made. So this is stronger evidence that the location effect might be indicative of some understanding of fractional numbers that was helpful in this comparison task.

The effect of largerstim on logRT was significant for only 6 (21%) of the response-larger group, (2 further marginally significant), but 15 (50%) of the response-smaller group, (3 further marginally significant). Indeed, 20 of the 58 participants did not have a larger median RT when the larger stimulus was a decimal than when it was a fraction. These were participants {3, 5, 6, 11, 13, 15, 16, 20, 21, 23, 24, 26, 27} of the response-larger group and {29, 31, 43, 50, 53, 57, 58} of the response-smaller group. Of these, four had a significant largerstim effect in the opposite direction to that expected 15, 26, 27 & 43.

No association was found between the size of the largerstim effect and the number of errors participants made (Spearman's r = -.040, p = .382, see figure 5.2(iii) for illustration).

5.4.3 Summary

Most participants demonstrated a significant distance effect in the expected direction. A little over half had a significant location affect, with longest RTs in the middle location. Under half of participants had a significant largerstim effect. For several of these it was in the opposite direction to that expected.

In both experiments, the distance effect was very strongly associated with success at the magnitude comparison task. The location effect was significantly associated with success at the magnitude comparison task in experiment two only. However, this association was marginally significant in experiment one. There was no significant association between the largerstim effect and success at the magnitude comparison task.

5.5 Discussion

5.5.1 Individual differences

Like Schneider & Siegler (2010), I found a significant distance effect for the majority of my participants. Unlike them, I had presented a task that afforded very little choice of strategy other than holistic magnitude comparison.

It might have been expected that participants who completed the magnitude comparison tasks more quickly would have made more errors. If so, this could have invalidated my attempts to establish an association between the distance effect as mentioned in the analysis. However, no speed-accuracy trade-off was found. This a result which has been found to be a feature of several simple mathematical tasks (Ratcliff, Thompson & McKoon, 2015).

The location and largerstim effects were much less consistently observed in individuals than the distance effect. However, though only 47 of the 88 participants in the two experiments (53%) demonstrated a *significant* location effect, almost all participants were slowest when responding to stimulus pairs in the middle location. This implies that when prompted to think about the size of fractions (and possibly decimals), their location near the anchor point of zero or one is a salient feature for many people. Moreover, it one around which they build a strategy.

It also appears to be part of a successful strategy or suite of strategies. There was a significant negative association between the strength of participants' location effect and the number of errors they made. Yet only three participants reported using a strategy that involved comparison with either zero or one in experiment one. This was also not one of the strategies reported by participants in Faulkenberry (2011).

This implies that the magnitude relationship with zero and one might be an unconscious response to fractional numbers. The swift comparison of fractions to the number one found by Kallai & Tzelgov (2009) does partly support this conjecture.

In the mixed effects modelling of logRTs for experiment two, RTs were significantly longer when the larger of the stimulus pair was a decimal, for both response groups. Looking at the individual differences, it can be seen that, out of 58 participants, only 38 showed this effect at all with only 17 of these demonstrating a significant effect. These were countered by 4 who significantly demonstrated the opposite effect. The fact that the largerstim effect nevertheless appeared to be significant in both response groups is an example of a minority strategy having an undue effect upon the general conclusions of analysis (see Siegler (1987) for another example).

The largerstim effect is still of interest as it might indicate some strategy that individuals are using to make magnitude comparisons that is either effective or problematic. It may be linked to useful strategies for fraction and decimal tasks other than magnitude comparison.

5.5.2 Next steps

The results from this further analysis of experiments one and two imply that not only the distance effect, but also the location effect might be indicative of useful features of the mental representation of fractional numbers.

The following chapter details the last experiment carried out for the purposes of this thesis. The intent of which was to find out whether individuals' success at general mathematical tasks involving fractions and decimals is associated with the strength of their distance, location and largerstim effects.

Chapter 6 Experiment five

6.1 Introduction

This chapter is a commentary on the two-stage final experiment carried out for the purpose of this thesis. It starts with a summary of the justification of the experiment and how it was intended to bring together the findings of the previous experiments. Details of the design of the experimental task and stimuli follow with reference to how they were devised to addresses the intention of the experiment.

The second part of the chapter covers the experimental procedure. Then the third section contains the results of the experiment along with the methodology of the analysis.

Finally, the last section of the chapter is a consideration of the implications of the results of the experiment.

6.2 Justification and design

The analysis into individual differences in performance in experiments one and two detailed in the last chapter highlighted possible associations between

an interesting and significant result outcome referred to as the expert-distance effect. This was the outcome that participants with a greater distance effect also made fewer mistakes in their size comparisons. Experiment five was designed to find out whether this expert-distance effect could also be seen for fraction and decimal skills other than size comparison. In other words, to discern whether there is a link between the strength of an adult's distance effect in number comparison tasks and their general knowledge of fractions and decimals.

As the intention of this experiment was to find a link between participants' outcomes for experiment one and their understanding of fractions and decimals, it was necessary to find an effective way to assess their understanding of fractions and decimals. Professor Margaret Brown of King's College London kindly offered use of the fraction and decimal portions of the CSMS (Chelsea Diagnostics Mathematics Tests) for this purpose.

The CSMS tests were formulated in the 1970s by experts in mathematical education and due to their effectiveness in discerning children's level of knowledge and misconceptions, they are still in use today. Extensive studies of schoolchildren's understanding of mathematics using these tests by professor Brown and others are published in Hart et al.

(1981). and Dickson et al. (1984). The CSMS tests and associated documentation are freely available for researchers to use via the website http://iccams-maths.org/CSMS/.

Built into the CSMS tests, there is an assessment mechanism by which participants can be assigned a level for each mathematical area of understanding. The levels are intended to be hierarchical in that level 1 indicates greater understanding than level 0 and so on. For fraction understanding, the levels range from 0 "unable to make a coherent attempt" to 4 ability to complete "questions where more than one operation is needed". For decimal understanding, the levels range from 0 "little grasp of place-value", through increasing and more sophisticated understanding of place-value, to level 6 "decimals as the result of a division; infinite number of decimals". A score of 60% or more is required on a designated set of question parts in order to attain a level.

Opportunity to study how the effects found are linked to individual differences in understanding and ability. Look at failures/rejected candidates on E2 and what they said about their method.

Despite the slight effect of response on RTs for the locations near_zero and near_one noted in experiment two, it was decided, for simplicity, to not vary the response required and ask all participants to choose the larger stimulus. The three main effects had the same general direction for both response groups and as was demonstrated in the last chapter, differences in the size of the location and largerstim effects between the groups might well have been down to the influence of individual differences between participants.

6.3 Method

6.3.1 Participants

Fifty-four healthy adults, (5 men) aged between 18 and 48 years (M=23.5, SD=7.65) participated in the study. All were psychology students at the University of Huddersfield who volunteered for the study in return for course credit.

6.3.2 Stimuli and materials

The stimuli for the first part of the experiment were identical to those used in experiment 1. For the second part of the experiment, slightly modified versions of two of the CSMS tests were used. Detailed assessment of both fraction and decimal knowledge was necessary for the purposes of experiment 5. Thus both the *Fractions 2* and *Place-value* *and Decimals* CSMS tests were employed. However, there were concerns that participants might abandon the test due to fatigue or fail to maintain concentration if the questions were overly repetitious or the test too long. So questions 21 to 26 were omitted from the *Fractions 2* test and questions 1 to 4 and 7 and 8 were omitted from the *Place-value and Decimals* test. It was judged that the skills tested in these questions were covered in other questions. Also, as adults rather than children were to be tested, the questions remaining emphasised real-life contextual use of number skills over classroom-based and abstract use of number.

Thus in the modified test, questions 1 to 23 covered fraction knowledge. These comprised 39 question parts. Questions 24 to 35 covered decimal knowledge. These comprised 59 question parts. The greater number of question parts for decimal knowledge reflecting the greater variety in the body of knowledge for decimals rather than fractions that the test designers discerned.

This omission of some questions caused the CSMS method for assessing levels of understanding in fractions and decimals to be slightly affected. In order to maintain a method of assessment that was consistent with the original CSMS method, a level was deemed to be reached if 60% or over of the remaining relevant question parts were completed correctly.

6.3.3 Procedure

The experiment was composed of two parts. Part one was identical to Experiment 1 with the exception that participants completed the task alone in small, quiet, well-lit, laboratory booths at the University of Huddersfield.

Exactly as for experiment one, the stimuli were presented on SuperLab® 4.0 stimulus presentation software. Participants were instructed that within each trial they had to decide which of the two numbers presented was the largest and to press the leftmost button on the response pad if it was the number on the left and the rightmost button on the response pad if it was the number on the right. They were informed that both speed and accuracy of response were important. In addition, it was made clear that some of the tasks were expected to appear very easy and some extremely difficult and the purpose of the experiment was to find out what factors made the task more difficult.

A practice block of four stimuli preceded the experimental blocks. Participants were given feedback on their accuracy on the practice stimuli and were allowed to ask questions if they did not understand the procedure. The 120 experimental stimuli were then

presented in random order in three blocks of 40 with no further feedback on accuracy nor opportunity to ask questions.
Participants were given the opportunity to take a break between the blocks. All participants were presented with the same stimuli. Response times and accuracy were recorded by the SuperLab® program.

After completing part one, participants were allowed to take a break of up to 5 minutes before they started part two.

For part two of the experiment, participants remained in their booths and completed the modified CSMS tests on paper. They were instructed to work alone, use no tools other than pen and paper and to respond to every question. For any questions they were completely unable to answer mathematically they were asked to write either "I don't know" or "?". Once they had embarked upon the experimental task, participants were not allowed to ask questions of the experimenter.

There was no time limit on completion of the task and the time taken was not recorded. Every participant completed the two tasks in this order.

6.4 Results

6.4.1 Introduction

The results section of this chapter is split into four parts. Part one is a summary of the mixed linear modelling analysis of logRTs for the first part of the experiment. That is, a replication of the analysis of response times for experiment one as detailed in chapter two of this thesis. General logistic regression error analysis was not carried out as error rates were investigated in terms of individual differences.

Part two is the analysis of individual differences in terms of the associations between the size of individuals' cognitive indicators (distance, location and largerstim effects) and their error rates on the task in part one of the experiment. This mirrors the analysis of experiments one and two that was reported in the last chapter.

Part three is a short summary of the general results of the full cohort of participants on the CSMS tests that formed part two of this experiment.

The last part of the analysis is an investigation into associations between the participants' various measures of performance on both parts of the experiment. That is the size of individual participants cognitive indicators and measures of their performance on the CSMS fractions and decimals paper-based tests.

6.4.2 Response time analysis of the first task

6.4.2.1 Pre-analysis data processing

Of the 54 participants, 8 were excluded from this part of the analysis for performing no better than chance at the 5% level (more than 50 errors out of 120 trials). Six of these made errors on the majority of trials indicating either a misunderstanding of the task (possibly choosing the smaller number rather than the larger) or very faulty reasoning.

Two response times were identified as extreme outliers, one of 58ms and another of 148ms. These were respectively 5.65 and 5.22 standard deviations below the mean logRT for the candidates concerned. These were identified as skewing the results of the mixed linear models and so were replaced with the candidates' next lowest RTs (1160ms and 1120ms respectively).

Only three candidates demonstrated a significant SNARC effect. This was indicated by the factor largelr being identified by the ANOVA test as having a significant result on mean logRT. They showed no other specific detectable differences from the rest of the cohort so no further analysis of factor largelr was carried out in any part of the analysis of this experiment.

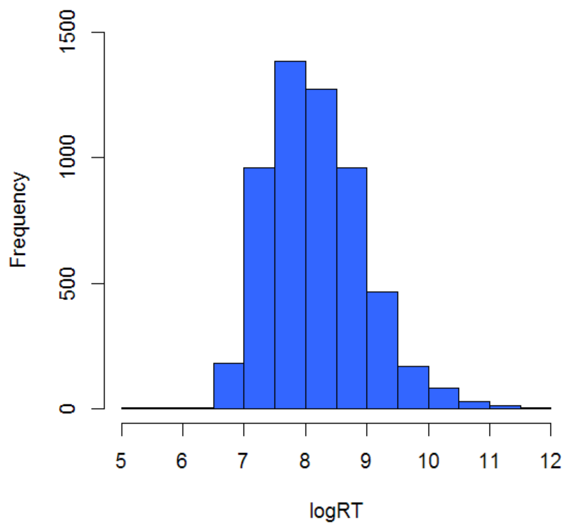

Figure 6.1 Histogram showing the distribution of logRT for experiment five

The natural log transform of RTs resulted in an approximately normal distribution (see figure 6.1) that was not significantly skewed ($\gamma = .606$, $p = .272$).

6.4.2.2 Linear modelling for logRT

The same method of formulating a mixed linear model was used as for the results of experiment one (see chapter two pages 36-43) and experiment two (see chapter three pages 59-66). As for both of these previous experiments the position factor was recoded into the location factor.

6.4.2.2.1 Individual testing of potential factors

First, the three possible fixed factors were added separately to the random intercepts only (baseline) model. An ANOVA test was then applied to detect improvements of each individual factor on its own to the fit of the model (see table 6.1). As in previous versions of the experiment, the factors of distance, location and largerstim each made a significant improvement to the baseline model.

Factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Baseline	3	10914	10934	-5454.1		
Distance	4	10367	10394	-5179.6	548.9	<.001
Location	5	10274	10307	-5132.1	644.1	<.001
Largerstim	4	10902	10929	-5447.1	13.87	<.001

Table 6.1 Results of ANOVA comparisons between baseline linear model and linear models including single factors

6.4.2.2.2 Building the model – single factors

The mixed linear model was then built up in stages by adding the fixed factors in turn. An ANOVA test was applied to test for an improvement in the model (see table 6.2). The addition of location and largerstim to the distance only model significantly improved the fit of the model.

Additional factor	df	AIC	BIC	Log Likelihood	L Ratio	p
Distance	4	10367	10394	-5179.6		
Location	6	9946	9986	-4967.0	425.3	<.001
Largerstim	7	9934	9980	-4960.1	13.85	<.001

Table 6.2 Results of ANOVA comparisons between versions of the linear model as single factors are added for experiment five

A summary of the mixed effects model for the single factors can be seen in table 6.3.

Factor/level	b (95% CI)	SE	df	t-value	p
distance	-1.036(-1.146, -0.927)	0.056	5470	-18.50	<.001
location: near_zero → middle	0.162(0.120, 0.205)	0.022	5470	7.56	<.001
location: middle → near_one	-0.274(-0.324, -0.224)	0.025	5470	-10.76	<.001
largerstim: decimal → fraction	-0.059(-0.089, -0.028)	0.016	5470	-3.72	<.001

Table 6.3 Summary of the linear model including all significant single factors for experiment five

Within this linear model without added interactions, the b-value for distance implies that for every 0.1 increase in distance, there was a reduction of 0.1036 in logRT (or a 9.8% reduction in RT). This distance effect was significant ($p < .001$).

There were significant ($p < .001$) changes in average logRT with the shift in stimulus location between zero and one. The b-value (0.162) between locations near_zero and middle was smaller in magnitude than that between middle and near_one (-0.274). In general, location near_one response times were fastest and middle response times were slowest. These were, on average, around 17.6% longer than those for the near_zero location and 24.0% longer than those for the near_one location.

The significant ($p < .001$) effect found for largerstim had a b-value of only -0.059. This implies that RTs for stimulus pairs in which the decimal was the larger number were, on average, around 5.7% longer than for those in which the fraction was the larger number.

6.4.2.2.3 Building the model – interactions

Interactions were added to the model to see if they would improve the fit to the logRT data. The addition of an interaction between distance and location significantly improved the fit of the model; as did the further addition of an interaction between location and largerstim. The addition of no other interactions significantly improved the fit of the model. A summary of the significant ANOVA comparisons can be seen in table 6.4.

Interaction added	df	AIC	BIC	Log Likelihood	L.Ratio	p
No Interaction	7	9934	9980	-4960.1		
distance/location	9	9898	9958	-4940.1	39.9	<.001
location/largerstim	11	9875	9948	-4926.5	27.2	<.001

Table 6.4 Results of ANOVA comparisons between versions of the linear model as interactions are added for experiment five

A summary of the final mixed effects model including interactions is in table 6.5. It can be seen that the greatest effect on logRT was that of the distance factor. The location factor remained significant when interactions were taken into consideration but the largerstim was only significant in its interaction with location.

The nature of the significant interaction between distance and location can be seen in figure 6.2. The expected distance effect can be seen for all three locations with median RTs decreasing as distances increase. For the middle locations it is a steeper effect with median RTs decreasing more rapidly generally and consistently across the range of distances.

In addition, at the distance between stimuli of 0.3 there is a hiatus in the expected distance effect particularly for the near_zero location, as was seen in experiments one and two. Again, this is probably an artefact of the unfortunate choice of stimuli at this distance.

single factors	b (95% CI)	SE	df	t-value	p
distance	-1.002(-1.218, -0.787)	0.110	5466	-9.10	<.001
location: near_zero → middle	0.278(0.193, 0.363)	0.043	5466	6.42	<.001
location: middle → near_one	-0.272(-0.378, -0.167)	0.054	5466	-5.05	<.001
largerstim: decimal → fraction	0.053(-0.019, 0.125)	0.037	5466	1.45	.148
two-factor interaction					
dist/location near_zero → middle	-0.344(-0.61, -0.077)	0.136	5466	-2.53	.012
dist/location middle → near_one	0.513(0.209, 0.818)	0.156	5466	3.30	.001
L.stim fraction/loc. near_zero → middle	-0.099(-0.181, -0.018)	0.042	5466	-2.38	.017
L.stim fraction/loc. middle → near_one	-0.258(-0.358, -0.159)	0.051	5466	-5.10	<.001

Table 6.5 Summary of the linear model including all interactions for experiment five

Figure 6.3 demonstrates the interaction between the factors of largerstim and location. There is a greater increase in mean logRT between the locations near_zero and middle for stimuli pairs in which the decimal is the larger of the stimuli. Between the location middle and near_one, there is a more significantly larger decrease in mean logRT for trials in which the fraction is the larger of the stimuli.

Figure 6.2 Graph of median RT against distance by location for experiment 5

It can be seen that for each location, the distance effect is approximately linear in nature. The inclusion, therefore, of the distance-location interaction in the model allays the concern that linear modelling is being used to model a non-linear effect.

A very similar interaction between distance and location was also found in experiment two (see pages 64-65).

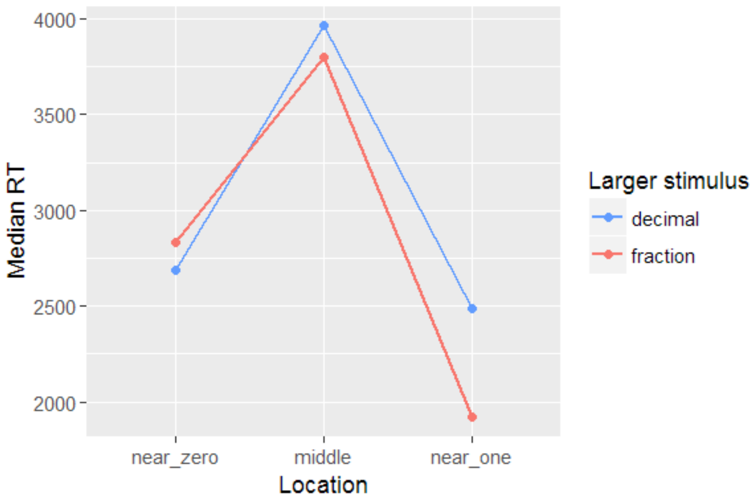

Figure 6.3 Graph of median RT against location by largerstim for experiment 5

A visual inspection of the qqplot of the residuals of the final experiment five linear mixed model for logRT (including interactions) indicated no obvious deviation from the assumption of normality (see figure 6.4). However, the slight wedge shape of the plot of the residuals did indicate some degree of heteroscedasticity, (see figure 6.5), which would imply the model did not account for all significant variables.

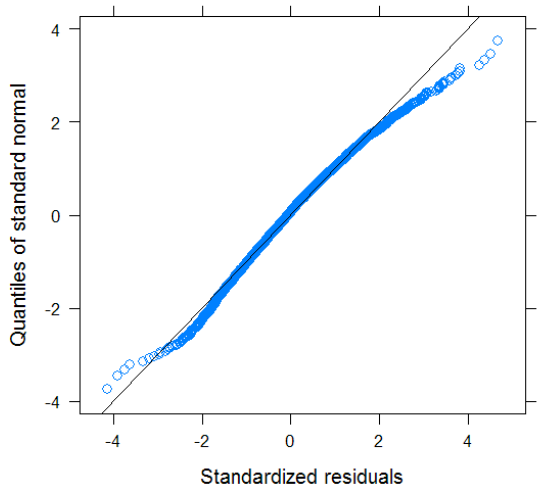

Figure 6.4 Qqplot of the residuals of the final mixed linear model for experiment five

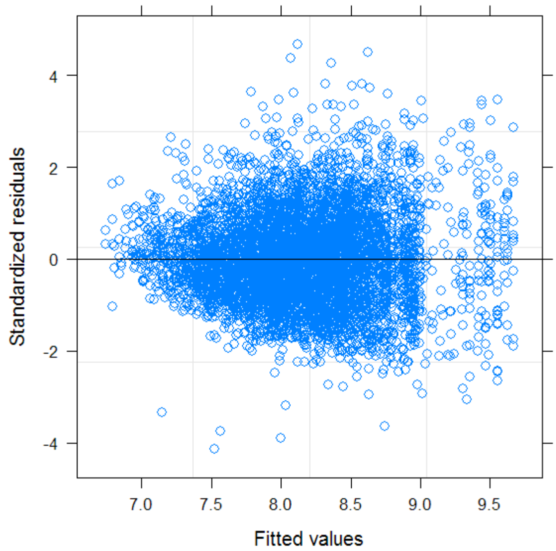

Figure 6.5 Plot of standardised residual against fitted values of the final mixed linear model for experiment five

6.4.2.2.3 Building the model – considering additional factors

Seven additional factors within the recorded data were added to the model in an attempt to account for the heteroscedasticity in the final model for logRT. Six made no notable improvement to the model. These were the decimal (p = .150), fraction (p = .803), denominator (p = .777) and numerator (p = .488) sizes, the absolute difference between the numerator and denominator of the fraction (p = .382) and the absolute difference between the numerator of the fraction and the first decimal place of the decimal (p = .186).

These six factors are not independent of the distance between stimuli and/or the location of the stimuli. If they had significantly improved the model this would have thrown the nature of the distance and/or location effects into question as they might, in fact, have been artefacts of one of these other effects.

Adding the factor of stimulus presentation order to the model did make a significant (p < .001) improvement. This is unsurprising as it is simply indicative of a learning effect taking place with response times decreasing as participants have more practice at the task in hand.

The addition of the factor trial order to the model made virtually no difference to the output of the model for the other factors. Indeed order had a b value of only -0.0016 implying an average decrease of 0.0016 in logRT for each successive stimulus which would amount to a 17.4% decrease in RT over all 120 trials. The addition of the order factor did not greatly affect the residuals plot (see figure 6.6).

As the purpose of this analysis was to identify the nature and comparative sizes of any influences of the experimental factors on response times, I decided that this degree of unaccounted for variance did not invalidate such conclusions being drawn from the mixed linear model.

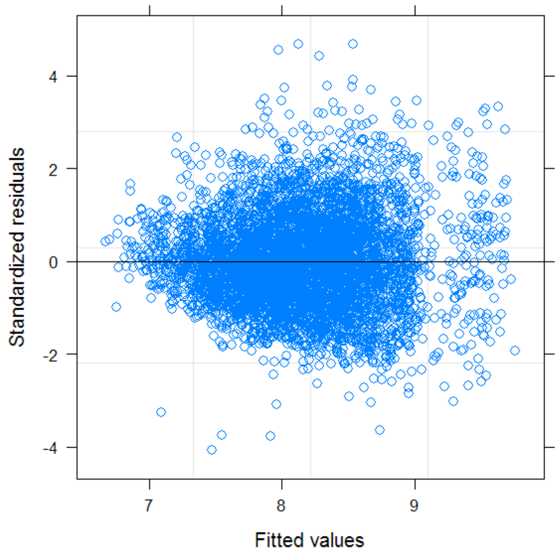

Figure 6.6 Plot of standardised residual against fitted values of the final mixed linear model with additional factor of trial order for experiment five

6.4.3 Individual difference analysis of the first task

6.4.3.1 Pre-analysis data processing

The analysis made of individual differences for this part of chapter six was conducted in much the same way as the analysis in the last chapter. As this experiment was intended as investigation of individual differences, no candidate was completely excluded because of poor performance. However, analysis was conducted first excluding the 8 participants who performed worse than chance termed *unsuccessful* (the others are termed *successful*). The results including the unsuccessful participants are given for completeness. Nevertheless, there are good reasons for excluding these participants as the reasons for their lack of success are unknown and might have been due to some misunderstanding of the task or lack of concentration.

For each participant an ANOVA tests was run for logRT on the three single factors of distance, location and largerstim. Again, interactions were not included for simplicity. These analyses yielded both the significance and the partial η^2 effect size of the distance, location and largerstim effects for each individual. *P*-values < .05 were considered significant.

The partial η^2 effect sizes measure the percentage of the variance within responses that might be explained by a particular factor. As such, they were used as measures of the effect of each factor on the responses of each individual. The greatest sum of these three partial η^2 values for any individual participant in experiment five was 50.4% and the least was 1.5% (*M* = 19.8, *SD* = 11.4). The results of this analysis can be found in appendix 3, page 161.

As in the last chapter, the partial η^2 values were recorded as positive if the effect for that participant was in the expected direction and negative otherwise. The expected direction for the distance factor was decreasing RT with increasing distance (tested via Spearman's r for each participant); for the largerstim factor was larger median RT for trials in which the decimal was the larger stimulus; for the location factor was median RT in middle position larger than in the other positions. Comparisons were then made between the strength of the three factors and participants' success in the first experimental task.

6.4.3.2 Analysis of associations between effect sizes and error rates

As for the analysis of experiments one and two in the last chapter, it was important to establish that there was no speed accuracy trade-off to enable the analysis of an association between distance effect and error rate. In fact, an insignificant association was found in the opposite direction to that expected, i.e. larger median RTs associated with greater error rates (Spearman's $r = .098$, $p = .742$). This indicates that accuracy and response time are independent for the successful participants.

However, when the unsuccessful participants were included, the association was unexpectedly increased to Spearman's $r = .257$, $p = .970$, now a significant reverse result. This can be explained by one outlying participant who took a particularly long time to complete the task and yet performed no better than chance.

For the successful participants, a significant distance effect on logRT was found for 37 of the 46 (80%). Only one, number 40, had an insignificant reverse distance effect; that is, increasing RTs associated with increasing distance between stimuli. A very large, significant negative association was found between the size of the distance effect, (participants' partial η^2) and the number of errors on the number comparison task for these individuals (Spearman's $r = -.794$, $p < .001$). Again, this was strong evidence that participants with a stronger distance effect were more successful at the task.

Three of the 8 unsuccessful participants demonstrated a significant distance effect. Another 4 of them showed an insignificant reverse distance effect {24, 27, 33, 46}. The association between distance effect size and error was little changed when the unsuccessful participants were included, (Spearman's $r = -.736$, $p < .001$). However, the above significant positive association between speed and inaccuracy should be born in mind when considering *this* result.

The effect of location on logRT was significant, for 38 of the 46 successful participants (83%). In addition, none of the successful participants were not slowest in the middle location. A medium-sized, significant, negative association was found between the size of the location effect and the number of errors for these individuals (Spearman's $r = -.288$, $p = .026$). This implies, again that there is some link between faster comparison of the magnitude of numbers located close to one than numbers in the middle of the zero-to-one range and general success at the number comparison task.

Two of the 8 unsuccessful participants had a significant location effect. For one of these, participant 24, it was a reverse effect, faster in the middle location than in the near_one location. Additionally, participant 46 had an insignificant but reversed location effect. So with the unsuccessful participants included, the magnitude of the association between the location effect and error rates was greatly increased (Spearman's $r = -.503$, $p < .001$).

Only 10 (22%) of the successful participants demonstrated a significant largerstim effect. For participants {3, 4, 6, 12, 15, 17, 25, 28, 29, 30, 31, 34, 36, 37, 38, 39, 40, 44} their median RT was greater when the larger stimulus was the fraction; 29 and 34 significantly so. Thus only 8 successful participants demonstrated a significant largerstim effect in the expected direction. There was an insignificant negative association between the largerstim effect and number of errors (Spearman's $r = -.145$, $p = .168$). So altogether this shows that the significantly larger logRTs when the larger of the stimulus pair is a decimal, that were seen in the mixed linear model analysis, was not a very common individual effect. It also was not linked to success at the comparison task.

Of the 8 unsuccessful participants, 3 of them showed a significant largerstim effect and one other, participant 16, demonstrated an insignificant reverse effect. With these participants included, Spearman's $r = .023$, $p = .564$, highly insignificant and implying, for this dataset, an increase in errors associated with a stronger largerstim effect.

6.4.4 Analysis of general performance on the second task

Unfortunately participant 27, did not complete the CSMS tests. This participant was one of the unsuccessful participants who performed worse than chance on part one of the experiment. The lack of completion was due to an oversight rather than any unwillingness so to do. Nevertheless, this participant had to be excluded entirely from any analysis involving the results of part two of the experiment. All of the other 53 participants were included.

6.4.4.1 Conflicts in the knowledge hierarchy

In only 5 (9.3%) of the decimals assessments and 3 (5.6%) of the fractions assessments were there conflicts within the hierarchy of assessment. That is when a participants scored < 60% at one level but ≥ 60% in a higher level. In only one case did more than one lower level score < 60%. This conflict was considered to be at a sufficiently low occurrence rate that it might be considered to not greatly affect the use of the assessment levels as a hierarchical measure of fraction and decimal understanding.

For the 8 assessments affected by the conflict, participants were awarded the higher level if the lower levels were all ≥ 50% correct and their overall percentage, at the relevant levels, was ≥ 60%. This accounted for 7 of the 8 affected assessments. For the other, the next lowest level which scored ≥ 60% was awarded.

6.4.4.2 Summary of correct responses

The minimum number of correct answers on the tests (out of 98) was 44 (44.9%). One person answered every single question correctly. The median result was 74 (75.5%) correct answers.

For the 39 fraction questions, the minimum number of correct answers was 13 (33.3%) and the median was 30 (76.9%). For the 59 decimal questions, the minimum number of correct answers was 29 (49.2%) and the median was 44(74.6%).

6.4.4.3 Summary of fraction and decimal levels obtained

Table 6.6 shows a summary of the fraction levels obtained by the 53 participants who completed the tests. The median level obtained was 3 (out of 4). Table 6.7 shows a summary of the decimal levels obtained by the 53 participants. The median level obtained was 4 (out of 6).

It is difficult to directly compare the fraction and decimal *levels* as they are on different scales. It is interesting, however, that one participant failed to gain any level at all on the fractions test but nobody scored less than two on the decimals test. In fact, it was the same participant who scored the lowest level on both tests. That was a participant successful at part one of the experiment.

Fraction level	Number of participants	Percentage
0	1	1.9
1	12	22.6
2	8	15.1
3	28	52.8
4	4	7.5

Table 6.6 Breakdown of the fraction levels obtained in the CSMS test

Decimal level	Number of participants	Percentage
0	0	0
1	0	0
2	1	1.9
3	11	20.8
4	20	37.7
5	8	15.1
6	13	24.5

Table 6.7 Breakdown of the decimal levels obtained in the CSMS test

6.4.5 Comparing individuals' performance on the tasks.

6.4.5.1 Pre-analysis considerations

Of the 53 participants that completed part two of experiment five, 7 were judged unsuccessful at part one. When comparing the performance of individuals in the two tasks, it seemed somewhat problematic to include participants who had possibly misunderstood the task in part one of the experiment.

In this part of the analysis, therefore, the results for the 46 *successful* candidates were considered first. However, the same analysis was repeated including the 7 unsuccessful candidates who did complete the CSMS tests. Both sets of results are reported herein.

The partial η^2 effect sizes from the first part of the experiment were again, used as measures of the size of individuals' distance, location and largerstim effects. As in part two of the analysis, these were recorded as negative if they were in the reverse direction to that expected. The fact that data from both parts of the experiment were analysed together was the reason for my reservations about automatically including the candidates unsuccessful at the first part.

6.4.5.2 Investigation of associations between features of performance

This time, associations were sought between the strength of the three factors in the first part of the experiment and participants' performance in the CSMS tests that constituted the second part of the experiment. The participants' distance, location and largerstim effect results were assessed against their overall success (% error), decimal reasoning success (% error and score) and fraction reasoning success (% error and score). The associations with the three error rates were expected to be negative, i.e. increasing effect implying decreasing error. The associations with the decimal and fraction levels were expected to be positive, i.e. increasing effect implying a higher level of skill.

As would be expected, there was a large, significant positive association between the number of errors made on the two parts of the experiment (Spearman's r = .601, p < .001 for the 46 successful participants and Spearman's r = .551, p < .001 for the 53 who completed the CSMS tests). The lack of perfect correlation between the results of the two test reflects the opportunity the CSMS test gave participants to demonstrate fraction and decimal skills other than number magnitude comparison.

Even more marked was the positive association between participants' error rates on the fraction and decimal sections of the CSMS tests (Spearman's r = .802, p < .001 for the 46 successful participants and Spearman's r = .794, p < .001 for the 53 who completed the CSMS tests). The association between participants' fraction and decimal levels was also large but less strikingly so (Spearman's r = .558, p < .001 for the 46 successful participants and Spearman's r = .567, p < .001 for the 53 who completed the CSMS tests). This is significant evidence that skills in fractions and decimals are very strongly linked and as skill in one increases so does skill in the other.

Distance effect	successful only		complete cohort	
	Spearman's *r*	*p*	Spearman's *r*	*p*
Total CSMS error	-.620	<.001	-.546	<.001
Decimal error	-.597	<.001	-.514	<.001
Decimal level	.574	<.001	.555	<.001
Fraction error	-.601	<.001	-.538	<.001
Fraction level	.463	<.001	.405	.001

Table 6.8 Spearman's rank associations between the distance effect and the five measures of success on the CSMS test of experiment five

First were considered the associations of the factor of distance with total error on the CSMS tests, decimal test error, decimal test level, fraction test error and fraction. The results are summarised in table 6.8.

As can be seen in table 6.8, there were significant associations between size of the distance effect and performance in both fraction and decimal tasks. These were large associations except perhaps for the fraction level measurement. Participants with a larger distance effect made fewer errors and achieved higher skill levels than those with a smaller distance effect.

Table 6.9 shows the results of the association tests between the size of participants' location effects and the five measures of their performance on the CSMS tests.

Location effect	successful only		complete cohort	
	Spearman's r	p	Spearman's r	p
Total CSMS error	-.276	.032	-.303	.014
Decimal error	-.275	.032	-.277	.022
Decimal level	.162	.141	.211	.065
Fraction error	-.253	.045	-.298	.015
Fraction level	.201	.091	.236	.044

Table 6.9 Spearman's rank associations between the location effect and the five measures of success on the CSMS test of experiment five

There are significant associations between the size of participants' location effects and all three of the error rates but not consistently so for their fraction and decimal level scores. Taken together these results imply that the location effect is a manifestation of some aspect of mathematical cognition which is helpful for skill with both fractions and decimals. However, it is not helpful at all *levels* of fraction and decimal skill.

The values of Spearman's r for the location effect are far lower than for the distance effect. So the distance effect is clearly far more strongly associated with success at fraction and decimal tasks than is the location effect.

Finally, table 6.10 contains the results of the Spearman's rank association tests on the size of participants' largerstim effect and their success on and levels for the CSMS test. All of these associations are in the expected direction. However, amongst the results there is only a marginally significant negative association between the strength of the largerstim effect and the decimal error rate.

Largerstim effect	successful		complete cohort	
	Spearman's r	p	Spearman's r	p
Total CSMS error	-.171	.128	-.089	.264
Decimal error	-.223	.069	-.131	.174
Decimal level	.057	.353	.010	.471
Fraction error	-.124	.205	-.056	.346
Fraction level	.070	.323	.028	.421

Table 6.10 Spearman's rank associations between the largerstim effect and the five measures of success on the CSMS test of experiment five

6.4.6 Summary of results

6.4.6.1 Mixed linear modelling

Much the same general results were found for the mixed linear modelling of logRTs as in experiments one and two. They confirm the influence on responses of the location of stimulus pair in the zero-to-one range.

They also confirm that there is some effect on RTs of whether the larger stimulus is a fraction or a decimal. However, this did appear to be restricted to a very small interaction with the location factor in the results of experiment five.

The most important confirmation was that the strongest effect found on logRTs was the magnitude of the distance between stimuli.

6.4.6.2 Associations between effect sizes and error rates in the number comparison task

Most participants demonstrated a significant distance effect in the expected direction. Most also had a significant location affect, with longest RTs in the middle location. A small minority of participants had a significant largerstim effect. For two of these it was in the opposite direction to that expected.

The distance effect was very strongly associated with success at the magnitude comparison task. The location effect was significantly associated with success at the magnitude comparison task but the association was small in size. There was no significant association between the largerstim effect and success at the magnitude comparison task.

6.4.6.3 Associations between effect sizes and the five measures of success in the CSMS tests.

There was a very strong positive association between success at the fraction and decimal tests. The strength of a participant's distance effect was significantly associated with success in both tests and with the levels of understanding of both fractions and decimals.

The location effect had a far smaller but still significant association with success on the two tests but not on fraction or decimal level attained. There were no significant associations between participants' largerstim effects and the measures of success in the CSMS tests.

6.5 Discussion

6.5.1 Largerstim effect

There was no significant association found between the strength of the largerstim effect and any measure of ability for the fraction and decimal tasks. There was a marginally significant negative association between the strength of the largerstim effect and the decimal error rate. This might imply that whatever produces the largerstim effect is a facet of small number knowledge that is helpful for decimal tasks at the lowest levels only.

The largerstim effect appears to be a cognitive indicator of some sort of bias regarding the mental representations or processing of fractions and/or decimals. It also appears to affect only a minority of people. This bias seems to be neither useful nor relevant to the effective processing of fractions and decimals. That could be because it is related to the underestimation of the magnitude of decimal as discussed in chapter four. However, the results of experiment five do not imply that the largerstim effect is a hindrance to fraction and decimal knowledge and ability.

Lastly, the lack of association between the largerstim effect and ability implies that it is not indicative of a wider array of appropriate strategies being used by the participants who demonstrate the largerstim effect. If it did we would expect to see a positive association between the largerstim effect and success. This is because the use of an increased variety of appropriate strategies for magnitude comparison of fractions has been linked to improved performance on fraction tasks (Fazio, DeWolf & Siegler, 2016).

6.5.2 Distance effect

The distance effect was again the strongest effect found in the mixed linear modelling analysis across logRTs. The majority of individual participants also had a significant distance effect such that their RTs decreased as the distance between the fraction and decimal pair increased.

A very significant association was found between participants' distance effects and their success in the magnitude comparison task *and* both tests of fraction and decimal skill. In addition, participants with stronger distance effects also achieved higher skill level scores in both the fraction and the decimal tests. The strength of the associations for fractions and decimals were approximately the same but maybe a little higher for the decimal skills level than the fraction skills level.

It should be noted that the distance effect I have found is specifically between the magnitude comparison of fractions and decimals. So when calculated for an individual, it is a measure of how strongly magnitude representations of fractions and decimals share commonalities for that participant. Therefore, my results indicate that mental representations which make stronger links between the magnitude of fractions and decimals are a feature of individuals with higher levels of skill with both fractions and decimals.

Links between the distance effect and number skill have been found by others. The number skill of small children and adults has been shown to be associated with the strength of their individual distance effects (Booth & Siegler, 2008; Fazio, Bailey, Thompson & Siegler, 2014; Holloway & Ansari, 2009).

Also, De Smedt, Verschaffel & Ghesquière (2009) carried out a longitudinal study testing children initially at around age 6 and then a year later. They found that the distance effect was a better predictor of progress in mathematics than age, intellectual ability and speed of response. Their findings also reflect the consistent lack of association between speed of response and accuracy that I have found.

There have also been studies that have made connections between understanding of specifically *fraction* magnitude and mathematical understanding. For example, Torbeyns, Schneider, Xin & Siegler (2015) found that, across three different countries, children's understanding of fraction magnitude was positively associated with their mathematical achievement in general. Also, Booth & Newton (2012) demonstrated that an individual's level of understanding of numerical magnitude in fractions was a strong predictor of their future progress in higher maths.

Furthermore, Faulkenberry (2011) found that students with stronger distance effects in fraction magnitude comparison tasks were more confident and relied less upon the use of calculators. Better understanding of proportions has even been linked to improved rationality in decision making (Alonso & Fernández-Berrocal, 2003).

Importantly, it is deliberate rather than automatic processing of number magnitudes that have been linked to success with mathematical tasks. Automatic responses like the SNARC effect and SiCE/SCE are not found for fractions unless people are trained to associate specific examples of fractions with abstract non-componential symbols (Kallai & Tzelgov, 2012b). This is probably because holistic fraction magnitude processing is deliberate and takes too long to be affected by any automatic response to irrelevant size or location information.

Furthermore, automatic responses to irrelevant information are not necessarily a feature of effective number magnitude representation. For example, Hoffmann, Mussolin, Martin & Schiltz (2014) found that the strength of the SNARC effect is inversely related to ability with whole numbers.

6.5.3 Location effect

The great majority of participants also demonstrated a significant location effect. That is, magnitude comparisons were significantly slowest in the middle location. In addition, the strength of this effect was significantly associated with success in both the fraction and decimal tests. However, it was not associated with the levels obtained in either the fraction or the decimal tasks.

These facts would imply that the location effect might be a cognitive indicator that highlights a feature of numerical magnitude understanding that is helpful for lower level fraction and decimal tasks only. The decimal skills for the lowest levels encompassed the understanding of whole number, tenth and hundredth decimal place values. The lowest level of the fraction skills hierarchy was the part-whole understanding of fractions.

I have suggested that the location effect is an outcome of the use of zero and one as anchor points against which to estimate the size of fractions in particular. However, the association between the strength of the location effect and success at decimal tasks implies that anchor points might also be used for decimal magnitude estimation.

6.5.4 Number lines

Both of the distance and location effects have their basis in features of the number line. The distance effect reflects relative magnitude. The location effect reflects the section of the number line which is the frame of reference for fractional numbers. The fact that these two effects have been linked to greater success at both fraction and decimal tasks might imply that the teaching of fractional numbers should emphasise the use of number lines.

Indeed, the use of number lines in the teaching of fractions and decimals has been shown to improve learning. For example, Fuchs, et al. (2013) found that an intervention that focussed on the *measurement interpretation* rather than part-whole understanding of fractions was most effective in helping low achieving students develop better performance at fraction tasks. The measure interpretation of fractions emphasises the combined meaning of fraction components, the equivalence of different fractions and the place of fractions within the number line.

Mayer, Lewis & Hegarty (1992) found that students were generally more successful at solving proportion questions when they constructed a number line on paper to help them (see pages 140-141 of their book). Furthermore, Jordan et al. (2013) found that number line estimation proficiency was the largest predictor of progress in the learning of fraction knowledge.

Countering this, the study of Bright, Behr, Post & Wachsmuth (1988) suggested that simply using number lines in the teaching of fractions does not necessarily improve fraction skills beyond tasks that specifically make use of number lines. They did not find that the children they studied necessarily transferred their number line knowledge to other tasks; implying they had learned procedures rather than relative magnitude concepts.

So using number lines when teaching the topics of fractions and decimals may well facilitate learning. Encouraging children to explicitly use number lines for tasks involving proportions may also be helpful. This may be because it strengthens the magnitude understanding which leads to the distance effect; maybe even the location effect also. However, number lines must be used to instruct for conceptual understanding not just discrete procedures.

The results of experiment five imply that it would be especially important to ensure that number lines are used to reinforce the commonalities between the magnitudes of fractions and decimals.

6.5.5 Next steps

The next and final chapter of the thesis sums up the analysis and findings of experiments one to five. It responds specifically to the questions asked by the research. What more has been discovered about the commonalities and differences between fractions and decimals and how do these discoveries inform the teaching of proportional knowledge?

Chapter 7 Summary and implications

7.1 Introduction

This final chapter summarises the findings of the research reported within the thesis. It starts with an assessment of the research and analysis techniques used.

The rest of the chapter consists of three short sections responding to the three themes of this thesis. These are, the commonalities between mental representations of fractions and decimals; the differences between mental representations of fractions and decimals; the implications for teaching and learning.

7.2 Techniques of research and analysis

Unlike most researchers that seek the distance effect, I did not use a target-stimulus paradigm in experiments one and two. Using a constant target against which to compare a stimulus certainly leads to a much simpler experimental design.

However there were no appropriate targets to use for the purposes of my investigation. The intention of which was to find out if there are processing routes by which fractions and decimals can be mapped onto a common magnitude mental representation – a mental number line. In other words, whether a distance effect could be found when the magnitude of a fraction is being compared to the magnitude of a decimals. Relatively unfamiliar fractions and decimals needed to be used to find out if these processes exist for fractions and decimals in general rather than just specific, familiar examples.

Not using a target against which to compare meant that it was necessary to control the distance between pairs. Doing so necessitated that the possible confounding factor of the position of the stimulus pairs within the zero-to-one range be accounted for. Indeed, not only did the location factor have a significant effect upon responses but there were significant interactions found between distance and location in both experiments two and five. So to not include some control for location could have resulted in misleading as well as impoverished results.

Not accounting for all possible relevant factors and interactions in the design of an experiment involving mental representations of number can lead to conflicting results. Jiang et al. (2016) demonstrated this in their SiCE task for single digit whole numbers. Like me, they did not use a target-stimulus paradigm which allowed them to control for additional factors which they used to show that responses to a SiCE task are more complex than previously thought.

Still, the design of the number comparison task might be criticised for being too complex and including too many factors. Neither the largerstim factor or the largerlr factor were expected to have any effect. They were mainly included to allow for balanced replication as well as the elimination of the SNARC effect. Nevertheless, the complexity of the experimental set-up did not preclude analysis and it did allow for the effects of several, possibly confounding factors to be eliminated from the investigation of the distance effect.

The decision to use a mixed linear modelling approach to the analysis of the number comparison tasks had drawbacks. In particular, it meant that only comparisons between the relative sizes of the significant effects could be made. Other techniques might have allowed for standardised effect sizes to be calculated. However, the considerable influence of individual differences upon the three main factors invalidated any approach which analysed averages across participants. Using the mixed linear modelling allowed valid conclusions to be drawn despite the inherently noisy data and the unbalanced design.

For the much simpler experiments three and four, participants responded to each stimulus only once so the only replications were between subjects. Particularly for experiment three, there was actually relatively little variance between participants. So the decision to analyse average responses for each stimuli was taken which allowed for comparison with other studies (especially Iuculano & Butterworth, 2011).

7.3 Commonalities – the distance effect

The key finding of this research is the one that it set out to find. That is, access to a common magnitude understanding of fractions and decimals. In three tasks in which the magnitude of a decimal was compared to that of a fraction, not only was the distance effect found but it was consistently by far the strongest effect.

There is some still question of whether the distance effect is indicative of number comparisons based upon holistic magnitudes rather than components. So I conducted a final meta-analysis on the combined results of experiment one, the response-larger data of experiment two (with the responses for distances 0.6 and 0.7 removed), and the magnitude comparison task of experiment five for the 46 successful participants. This combined the results of 110 participants.

The distance effect remained the strongest effect. The location effect remained significant as did the largerstim effect by way of interactions with distance and the interaction of all three factors. Other magnitude features of the stimuli including comparisons between stimulus components were tested. These were the fraction size, the decimal size, the numerator and denominator of the fraction, the absolute difference between the numerator and denominator and the absolute difference between the numerator and the first digit of the decimal.

These factors are not independent of the distance between stimuli and/or the location of the stimulus pair so added individually they did appear to improve the model of logRT. However, the combination of distance, location and largerstim effects superseded the effect of these other factors.

Three digit decimals have been seen to have commonalities in their mental representations as three digit whole numbers (DeWolf et al., 2014). In their comparison tasks for large whole numbers, Barth, Kanwisher & Spelke (2003) found that the format of numbers (visual or auditory) did not affect performance, nor did the size of the numbers. They found that the *ratio* of the sets being compared was what predicted performance best. That is, they found a distance effect scaled for set size. I didn't find this as it would have presented as a monotonic, increasing location effect for each distance.

The persistent and strong distance effect that I have found for small numbers and among people with functional mathematical skills strongly implies that people try to fit small numbers into their existing cognitive structures for whole numbers; that is the mental number line.

7.4 Differences

Though theoretically representing the same amount, fractions and decimal are used differently. For example, relative proportions are more often represented by and better understood using fractions than decimals (DeWolf, Bassok & Holyoak, 2015). Magnitude responses to decimals are faster than to fractions (e.g. DeWolf et al., 2014; Iuculano & Butterworth, 2011). This last result was also found in my experiments three and four at least for stimuli in the middle of the zero-to-one range. In all five experiments responses were affected by the location of stimuli in relation to zero and one.

The specific effect of location upon responses in the magnitude comparison tasks and the magnitude estimation task was one of the key novel findings of this research. It was also found to be a significant individual effect for approximately half of the participants in experiments one, two and five.

Rational numbers – fractions and decimals in this case, reside within a restricted domain. The fact that they are influenced by the end-points of that domain is not remarkable. It seems to be indicative of some sort of processing route to magnitude representation that depends on the use of anchor points. So the magnitude of a fraction is generally accessed as a process of adjustment away from either zero or one.

The results of experiments three and four imply that this is indeed a process applied to fractions alone rather than to decimals and proportions in general. In fact, it seems that only if prompted specifically to compare with zero and one does a decimal's distance from these two points affect the response to it and then only slightly. This is both in terms of response times and errors.

The evidence is that three-digit decimals are processed much like three-digit whole numbers (see DeWolf et al., 2014). This would appear in some ways to be an appropriate approach considering that, especially when used in a scientific or measurement context, the magnitude of a decimal number is dependent on the units used and a two number differing in place value can mean the same thing. For example, 452mm = 0.452m. However, 452 ≠ 0.452 in other contexts such as when representing a proportion. In the task of comparing a decimal's magnitude to that of a fraction, place value is relevant.

Evidently, there is a different processing route to magnitude representation for decimals than fractions. Decimal magnitudes appear to be accessed more quickly with a strong componential bias towards the first decimal place. Fractions appear to be accessed more slowly with the components considered but combined into a holistic magnitude, if necessary, which is more accurate than that of a multi-digit decimal. These observations are generalisations. There were large variations between individuals found in experiments one, two and five.

Differences are significant between individuals' mental representations of fractions and decimals and their routes to processing them. RT results for experiments one, two and five were very noisy and much of the noise was down to the very great differences between individuals. Individuals not only had significantly different baseline RTs but also significantly different sizes of distance and location effects. Some few people also had a significant largerstim effect, not all of them in the same direction.

The results of these number comparison experiments, combined with the findings of others on the effects of stimulus and strategy choice on responses to fractions (e.g. Faulkenberry & Pierce, 2011; Meert et al. 2009, 2010; Schneider & Siegler, 2010; Shaki & Fischer, 2013), build up a picture of a very complicated cognitive structure for processing the magnitudes of fractional numbers. Structures that vary depending on the individual and the task that they are performing.

The apparent visual simplicity of the fraction and decimal stimuli presented in the five experiments carried out for this thesis belie the complex nature of their meaning, processing and representation. Future research into the cognitive structures that support our use and understanding of fractional numbers must take into account all of this complexity within and between individuals and tasks.

Most of the participants of these experiments performed remarkably well at a challenging task. So a holistic magnitude distance effect combined with an effect of location for many individuals and a largerstim effect for a small minority of individuals seems the best account of the response time results of the magnitude comparison task.

7.5 Implications for teaching and learning

Gérard (1998) makes an argument that the notion of understanding representation is important in mathematics education research because we do not directly experience number we use symbolic representations. So understanding the nature of our internal representations and how they work is key to understanding how to teach numerical concepts better.

When we teach, we facilitate the formation of mental number representations in the minds of each of our students. The pedagogical techniques we employ affect the development of these mental representations. An example of how different approaches to teaching result in the different development of mental representations can be seen in the cross-cultural study of Resnick et al. (1989) discussed in chapter one.

As discussed in the previous chapter, the findings of my research add to previous research which has demonstrated that holistic magnitude understanding of fractions is key to achievement in mathematics, not only for fraction tasks. However, I have also shown that people who effectively translate fractional numbers into magnitude representations common to both fractions and decimals are also more successful at solving fraction and decimal problems.

I have also highlighted the possibility that the use of the anchor points of zero and one in making judgements of the magnitude of fractions, (and maybe decimals), is linked to improved basic skills with fractions and decimals. The magnitude representations and anchor points are part of a mental number line. So I have suggested that the use of physical number lines might be used to improve learners' understanding of fractions and decimals and their commonalities.

I have not found a causal relationship between either holistic, common magnitude representations of fractions and decimals, or the use of anchoring for estimation and improvement in fraction and decimal skills. I would suggest, however, that it makes sense for teacher to aim to build up knowledge of fractions and decimals in their students in ways that make use of these available structures that have been demonstrated to be potentially effective.

References

Alonso, D., & Fernández-Berrocal, P. (2003). Irrational decisions: attending to numbers rather than ratios. *Personality and Individual Differences, 35*(7), 1537-1547.

Arend, I., & Henik, A. (2015). Choosing the larger versus choosing the smaller: Asymmetries in the size congruity effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 41(6)*, 1821-1830.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. Cognition, 86(3), 201-221.

Behr, M., Lesh, R., Post, T., & Silver, E. (1983). Rational Number Concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of Mathematics Concepts and Processes* (pp. 91-125). New York: Academic Press.

Bonato, M., Fabbri, S., Umiltà, C., & Zorzi, M. (2007). The mental representation of numerical fractions: real or integer? *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1410-1419.

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology, 37*(4), 247-253.

Booth, J. L., & Siegler, R. S. (2008). Numerical Magnitude Representations Influence Arithmetic Learning. *Child Development, 79*(4), 1016-1031.

Bright, G. W., Behr, M. J., Post, T. R., & Wachsmuth, I. (1988). Identifying Fractions on Number Lines. *Journal for Research in Mathematics Education, 19*(3), 215-232.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes, 79*(2), 115-153.

Chiou, R. Y., Wu, D. H., Tzeng, O. J., Hung, D. L., & Chang, E. C. (2012). Relative size of numerical magnitude induces a size-contrast effect on the grip scaling of reach-to-grasp movements. *Cortex, 48*(8), 1043-1051.

Cohen, D. J. (2010). Evidence for direct retrieval of relative quantity information in a quantity judgment task: decimals, integers, and the role of physical similarity. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*(6), 1389-1398.

Cohen, D. J., Ferrell, J. M., & Johnson, N. (2002). What very small numbers mean. *Journal of Experimental Psychology: General, 131*(3), 424-442.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (Vol. 2nd;2;). GB: Routledge Ltd.

Crawley, M. J. (2005). *Statistics: An introduction using R*. Chichester: Wiley.

de Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition, 110*(2), 198-207.

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., Dupoux, E., & Mehler, J. (1990). Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 626-641.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of experimental child psychology, 103*(4), 469-479.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: relational reasoning with fractions and decimals. *Journal of Experimental Psychology. General, 144*(1), 127-150.

DeWolf, M., Grounds, M. A., Bassok, M., & Holyoak, K. J. (2014). Magnitude comparison with different types of rational numbers. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 71-82.

Dickson, L., Brown, M., & Gibson, O. (1984). *Children learning mathematics: A teacher's guide to recent research*. Oxford: Cassell Education

Faulkenberry, T. J. (2011). *Individual differences in mental representations of fraction magnitude*. Paper presented at the 39th Annual Meeting of the Research Council on Mathematics Learning, Charlotte, NC: RCML.

Faulkenberry, T. J., Montgomery, S. A., & Tennes, S. A. (2015). Response trajectories reveal the temporal dynamics of fraction representations. *Acta Psychologica (Amst), 159*, 100-107.

Faulkenberry, T. J., & Pierce, B. H. (2011). Mental Representations in Fraction Comparison: Holistic Versus Component-Based Strategies. *Experimental Psychology, 58*(6), 480-489.

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.

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 42(1), 1-16.

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, London: SAGE.

Fitousi, D., & Algom, D. (2006). Size congruity effects with two-digit numbers: Expanding the number line? *Memory & Cognition, 34*(2), 445.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R. & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*(3), 683-700.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*(1), 35-42.

Gabriel, F., Coche, F., Szucs, D., Carette, V., Rey, B., & Content, A. (2013). A componential view of children's difficulties in learning fractions. *Frontiers in Psychology, 4*, 715.

Gabriel, F., Szucs, D., & Content, A. (2013). The mental representations of fractions: adults' same-different judgments. *Frontiers in Psychology, 4*, 385.

Gabriel, F. C., Szucs, D., & Content, A. (2013). The development of the mental representations of the magnitude of fractions. *PLoS One, 8*(11), e80016.

Ganor-Stern, D. (2012). Fractions but not negative numbers are represented on the mental number line. *Acta Psychologica (Amst), 139*(2), 350-357.

Ganor-Stern, D., Tzelgov, J., & Ellenbogen, R. (2007). Automaticity and two-digit numbers. *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 483-496.

Ganor-Stern, D., & Tzelgov, J. (2008). Negative numbers are generated in the mind. *Experimental psychology, 55*(3), 157-163.

Gérard, V. (1998). A comprehensive theory of representation for mathematics education. *Journal of Mathematical Behavior, 17*(2), 167-181.

Gigerenzer, G., & Edwards, A. (2003). Simple Tools For Understanding Risks: From Innumeracy To Insight. *BMJ: British Medical Journal, 327*(7417), 741-744.

Goldfarb, L., Henik, A., Rubinsten, O., Bloch-David, Y., & Gertner, L. (2011). The numerical distance effect is task dependent. *Memory & cognition, 39*(8), 1508.

Gonzalez, E. G., & Kolers, P. A. (1982). Mental manipulation of arithmetic symbols. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(4), 308.

Hart, K., Brown, M.L., Küchemann, D.E., Kerslake, D., Ruddock, G. & McCartney, M. (1981) *Children's understanding of mathematics: 11-16*. London: John Murray.

Hasemann, K. (1981). On Difficulties with Fractions. *Educational Studies in Mathematics, 12*(1), 71-87.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & cognition, 10*(4), 389-395.

Hoffmann, D., Mussolin, C., Martin, R., & Schiltz, C. (2014). The impact of mathematical proficiency on the number-space association. *PLoS One, 9*(1), e85048.

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.

Huber, S., Klein, E., Willmes, K., Nuerk, H. C., & Moeller, K. (2014). Decimal fraction representations are not distinct from natural number representations - evidence from a combined eye-tracking and computational modeling approach. *Frontiers in Human Neuroscience, 8*, 172.

Ischebeck, A., Schocke, M., & Delazer, M. (2009). The processing and representation of fractions within the brain: an fMRI investigation. *Neuroimage, 47*(1), 403-413.

Iuculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *The Quarterly Journal of Experimental Psychology, 64*(11), 2088-2098.

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*(3), 1221-1247.

Jacob, S. N., & Nieder, A. (2009). Notation-independent representation of fractions in the human parietal cortex. *Journal of Neuroscience, 29*(14), 4652-4657.

Jacob, S. N., Vallentin, D., & Nieder, A. (2012). Relating magnitudes: the brain's code for proportions. *Trends in Cognitive Science, 16*(3), 157-166.

Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia, 38*(1), 93-110.

Jiang, T., Zhang, W., Wen, W., Zhu, H., Du, H., Zhu, X., Gao, X., Zhang, H., Dong, Q. & Chen, C. (2016). Reevaluating the two-representation model of numerical magnitude processing. *Memory & cognition, 44*(1), 162-170.

Johnson, D. M. E. (1939). *Confidence and Speed in the Two-category Judgement*: Columbia university.

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*(1), 45-58.

Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: an entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1845-1864.

Kallai, A. Y., & Tzelgov, J. (2012). The place-value of a digit in multi-digit numbers is processed automatically. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(5), 1221-1233.

Kallai, A. Y., & Tzelgov, J. (2012). When meaningful components interrupt the processing of the whole: the case of fractions. *Acta Psychologica (Amst), 139*(2), 358-369.

Kallai, A. Y., & Tzelgov, J. (2014). Decimals are not processed automatically, not even as being smaller than one. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*(4), 962-975.

Karolis, V., Iuculano, T., & Butterworth, B. (2011). Mapping numerical magnitudes along the right lines: differentiating between scale and bias. *Journal of Experimental Psychology: General, 140*(4), 693-706.

Kerslake, D. (1986). *Fractions: Children's Strategies and Errors. A Report of the Strategies and Errors in Secondary Mathematics Project*: ERIC.

Koch, L. C., & Li, X. (1996). Students' understanding of computation-related rational number skills. *The Journal of mathematical behavior, 15*(2), 193-205.

Lindemann, O., Abolafia, J. M., Girardi, G., & Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1400-1409.

Longo, M. R., & Lourenco, S. F. (2007). Spatial attention and the mental number line: Evidence for characteristic biases and compression. *Neuropsychologia, 45*(7), 1400-1407.

Matthews, P. G., Chesney, D. L., & McNeil, N. M. (2014). *Are Fractions Natural Numbers, Too?* Paper presented at the Proceedings of the 36th annual conference of the cognitive science society, Cognitive Science Society, Austin TX.

Mayer, R. E., Lewis, A. B. & Hegarty, M. (1992). Mathematical misunderstandings. In Campbell, J. (Ed) (1992) *The nature and origins of mathematical skills* (pp. 137–153). Amsterdam: North-Holland

Meert, G., Grégoire, J., & Noël, M. P. (2009). Rational numbers: componential versus holistic representation of fractions in a magnitude comparison task. *The Quarterly Journal of Experimental Psychology, 62*(8), 1598-1616.

Meert, G., Grégoire, J., & Noël, M. P. (2010). Comparing 5/7 and 2/9: Adults can do it by accessing the magnitude of the whole fractions. *Acta Psychologica (Amst), 135*(3), 284-292.

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature, 215*, 1519-1520.

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition, 82*(1), B25-B33.

Obersteiner, A., Van Dooren, W., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction, 28*, 64-72.

Ratcliff, R., Thompson, C. A., & McKoon, G. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition, 137*, 115-136.

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual Bases of Arithmetic Errors: The Case of Decimal Fractions. *Journal for Research in Mathematics Education, 20*(1), 8-27.

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of social service Research, 21*(4), 37-59.

Schliemann, A., & Carraher, D. (1992). Proportional reasoning in and out of school. In P. Light & G. Butterworth (Eds.) *Context and cognition: Ways of learning and knowing* (pp. 47-73). Hemel Hempstead: Harvester Wheatsheaf.

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1227-1238.

Shaki, S., & Fischer, M. H. (2013). Your neighbors define your value: a study of spatial bias in number comparison. *Acta Psychologica, (Amst), 142*(3), 308-313.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*(3), 250-264.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273-296.

Sprute, L., & Temple, E. (2011). Representations of Fractions: Evidence for Accessing the Whole Magnitude in Adults. *Mind, Brain, and Education, 5*(1), 42-47.

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.

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction, 37*, 5-13.

Tversky, A., & Kahneman, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Varey, C. A., Mellers, B. A., & Birnbaum, M. H. (1990). Judgments of Proportions. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 613-625.

Zhang, J., & Wang, H. (2005). The effect of external representations on numeric tasks. *The Quarterly Journal of Experimental Psychology Section A, 58*(5), 817-838.

Zhou, X., Chen, C., Chen, L., & Dong, Q. (2008). Holistic or compositional representation of two-digit numbers? Evidence from the distance, magnitude, and SNARC effects in a number-matching task. *Cognition, 106*(3), 1525-1536.

Appendix 1 Individual differences for participants in experiment one.

Participant	no. of errors	distance			location			largerstim		
		$F(1,115)$	p	partial η^2	$F(2,115)$	p	partial η^2	$F(1,115)$	p	partial η^2
2	8	31.11	<.001	.119	17.52	<.001	.232	4.87	.029	.041
3	21	6.23	.014	.042	1.71	.186	.029	6.41	.013	.053
4	16	0.89	.348	.002	1.60	.207	.026	4.22	.042	.035
5	9	8.93	.003	.036	6.12	.003	.096	0.00	.985	.000
6	45	2.37	.127	.020	0.68	.510	.011	0.21	.652	.002
7	38	0.09	.765	.001	0.17	.846	.003	1.46	.229	.013
8	19	19.80	<.001	.107	6.15	.003	.097	3.67	.058	.031
9	7	16.54	<.001	.082	4.70	.011	.073	4.65	.033	.039
10	8	35.89	<.001	.155	15.79	<.001	.216	0.11	.745	.001
11	17	42.82	<.001	.222	6.55	.002	.102	13.12	<.001	.102
12	14	15.82	<.001	.095	2.79	.066	.046	9.10	.003	.073
13	17	20.59	<.001	.056	24.82	<.001	.300	3.74	.056	.031
14	14	11.10	.001	.035	12.53	<.001	.174	16.66	<.001	.127
15	18	5.69	.019	.035	0.52	.598	.009	0.12	.728	.001
16	24	6.90	.010	.034	2.90	.059	.048	0.00	.949	.000
17	14	8.41	.004	.027	7.70	.001	.117	0.69	.408	.006
18	43	0.11	.739	.000	0.64	.532	.011	0.55	.461	.005
19	4	25.26	<.001	.150	2.40	.095	.040	0.24	.629	.002
20	4	33.84	<.001	.177	3.05	.051	.049	7.21	.008	.059
21	11	10.02	.002	.059	2.48	.089	.041	1.55	.216	.013
22	10	30.27	<.001	.127	12.61	<.001	.178	6.74	.011	.055
23	14	18.61	<.001	.093	8.27	<.001	.126	7.04	.009	.058
24	3	39.16	<.001	.194	5.35	.006	.084	2.98	.087	.025
25	17	9.78	.002	.036	6.48	.002	.100	3.57	.061	.030
26	27	5.37	.022	.022	8.38	<.001	.127	2.26	.135	.019
27	6	88.36	<.001	.367	6.53	.002	.100	11.68	.001	.092
28	10	65.80	<.001	.312	16.04	<.001	.219	10.05	.002	.080
29	18	40.86	<.001	.215	4.41	.014	.071	17.26	<.001	.131
30	4	27.02	<.001	.168	1.14	.324	.020	13.68	<.001	.106
31	22	12.45	.001	.066	4.47	.013	.072	14.68	<.001	.113

key: [highlighted] p < .05

Appendix 2a Individual differences for participants in experiment two response larger group.

Participant	no. of errors	distance			location			largerstim		
		F(1,139)	p	partial η^2	F(2,139)	p	partial η^2	F(1,139)	p	partial η^2
1	17	78.96	<.001	.300	2.26	.109	.031	3.15	.078	.022
2	23	11.45	.001	.068	2.51	.085	.035	3.17	.077	.022
3	32	21.21	<.001	.090	2.10	.127	.029	0.18	.674	.001
4	21	49.12	<.001	.230	0.82	.441	.012	8.42	.004	.057
5	53	0.11	.743	.000	0.68	.507	.010	0.10	.755	.001
6	18	28.50	<.001	.105	6.50	.002	.086	2.65	.106	.019
7	61	1.49	.225	.003	1.91	.152	.026	1.08	.301	.008
8	19	45.67	<.001	.205	2.89	.059	.040	1.16	.283	.008
9	11	61.28	<.001	.255	3.06	.050	.042	0.70	.405	.005
10	11	21.92	<.001	.093	2.87	.060	.039	3.22	.075	.023
11	17	66.19	<.001	.227	8.26	<.001	.106	0.33	.570	.002
12	32	2.58	.111	.005	4.65	.011	.063	1.28	.261	.009
13	27	2.95	.088	.013	3.47	.034	.047	2.05	.155	.014
14	39	6.15	.014	.062	2.31	.103	.032	0.47	.495	.003
15	10	22.20	<.001	.106	1.20	.304	.017	4.57	.034	.032
16	46	2.24	.137	.002	4.02	.020	.055	0.80	.373	.006
17	15	37.58	<.001	.148	7.32	.001	.095	5.31	.023	.037
18	12	47.71	<.001	.215	1.11	.332	.016	0.70	.405	.005
19	12	50.56	<.001	.156	13.01	<.001	.156	7.16	.008	.049
20	35	4.25	.041	.011	4.53	.012	.061	0.80	.372	.006
21	19	13.05	<.001	.054	2.79	.065	.039	1.63	.204	.012
22	10	68.75	<.001	.251	8.09	<.001	.104	7.04	.009	.048
23	18	17.18	<.001	.046	9.49	<.001	.119	1.86	.175	.013
24	18	14.81	<.001	.058	2.75	.067	.038	0.08	.772	.001
25	20	5.73	.018	.022	5.45	.005	.073	1.14	.288	.008
26	17	9.08	.003	.037	5.00	.008	.067	3.03	.084	.021
27	35	1.17	.282	.002	1.69	.189	.024	5.54	.020	.038
28	14	48.74	<.001	.186	4.63	.011	.063	0.11	.747	.001

Key:
- (white) $p < .05$
- (cyan) $.05 \le p < .10$
- (yellow) $p > .10$

Appendix 2b Individual differences for participants in experiment two response smaller group.

Participant	no. of errors	distance			location			largerstim		
		F(1,139)	p	partial η^2	F(2,139)	p	partial η^2	F(1,139)	p	partial η^2
29	11	23.29	<.001	.079	6.88	.001	.090	0.19	.663	.001
30	28	19.08	<.001	.099	1.74	.180	.023	7.78	.006	.053
31	7	16.80	<.001	.078	2.49	.087	.035	2.09	.150	.015
32	59	8.35	.004	.006	39.31	<.001	.358	13.05	<.001	.086
33	42	3.79	.053	.005	5.43	.005	.071	6.65	.011	.046
34	34	7.16	.008	.040	2.31	.103	.031	7.78	.006	.053
35	43	1.10	.295	.000	9.29	<.001	.117	0.56	.454	.004
36	19	23.92	<.001	.074	10.71	<.001	.132	2.25	.136	.016
37	51	0.19	.668	.000	0.54	.586	.008	0.00	.973	.000
38	29	4.65	.033	.022	0.59	.558	.008	0.46	.497	.003
39	16	28.72	<.001	.081	12.97	<.001	.156	2.94	.089	.021
40	14	26.21	<.001	.124	1.05	.352	.014	10.50	.001	.070
41	40	0.19	.662	.006	5.83	.004	.077	0.04	.845	.000
42	29	12.07	.001	.032	6.06	.003	.079	5.23	.024	.036
43	57	0.60	.439	.011	2.37	.097	.033	3.72	.056	.026
44	15	23.13	<.001	.108	0.95	.389	.013	5.50	.020	.038
45	21	26.09	<.001	.112	3.63	.029	.048	4.61	.034	.032
46	36	0.86	.356	.004	0.63	.534	.009	9.73	.002	.065
47	15	22.29	<.001	.086	4.20	.017	.055	5.96	.016	.041
48	51	0.02	.877	-.000	1.56	.215	.022	9.76	.002	.066
49	21	12.94	<.001	.038	6.81	.002	.088	2.87	.093	.020
50	43	1.75	.188	.015	0.41	.661	.006	0.86	.354	.006
51	56	10.60	.001	.035	9.48	<.001	.117	8.40	.004	.057
52	9	14.81	<.001	.032	10.55	<.001	.130	4.47	.036	.031
53	28	4.99	.027	.009	8.55	<.001	.108	2.22	.139	.016
54	61	0.16	.687	.001	0.30	.742	.004	3.79	.054	.027
55	3	56.66	<.001	.190	9.34	<.001	.117	7.21	.008	.049
56	18	19.45	<.001	.065	6.62	.002	.085	4.29	.040	.030
57	60	1.19	.277	.005	1.38	.254	.019	0.39	.535	.003
58	20	54.63	<.001	.236	1.14	.323	.016	0.04	.836	.000

Key

white	$p < .05$
cyan	$.05 \le p < .10$
yellow	$p > .10$

Appendix 3 Individual differences for participants in experiment five.

Participant	no. of errors	distance $F(1,115)$	distance p	distance partial η^2	location $F(2,115)$	location p	location partial η^2	largerstim $F(1,115)$	largerstim p	largerstim partial η^2
1	27	3.14	.079	.009	4.92	.009	.078	3.07	.082	.026
2	31	.597	.441	<.001	5.19	.007	.080	1.95	.165	.017
3	22	15.8	<.001	.075	3.21	.044	.054	.607	.438	.005
4	14	29.1	<.001	.130	12.7	<.001	.180	.096	.757	.001
5	12	27.4	<.001	.148	1.82	.166	.028	7.21	.008	.059
6	26	7.35	.008	.025	4.70	.011	.077	0.93	.337	.008
7	12	14.4	<.001	.043	13.1	<.001	.182	10.1	.002	.081
8	9	61.0	<.001	.258	8.94	<.001	.130	13.8	<.001	.107
9	2	51.1	<.001	.253	2.73	.069	.045	1.39	.241	.012
10	11	31.2	<.001	.122	10.8	<.001	.156	.867	.354	.007
*11	95	21.1	<.001	.107	3.18	.045	.049	2.89	.092	.025
12	5	32.2	<.001	.170	3.62	.030	.059	2.23	.138	.019
13	13	24.8	<.001	.130	4.17	.018	.067	5.22	.024	.043
14	20	14.2	<.001	.074	6.42	.002	0.10	.285	.595	.002
15	18	8.99	.003	.042	6.86	.002	.107	1.36	.246	.012
*16	77	4.83	.030	.040	.967	.383	.016	1.18	.279	.010
17	19	4.78	.031	.021	3.33	.039	.055	.224	.637	.002
18	26	2.65	.107	.016	5.35	.006	.086	.336	.563	.003
19	14	18.8	<.001	.062	15.0	<.001	.205	4.62	.034	.039
20	19	6.12	.015	.033	1.81	.168	.030	1.67	.198	.014
21	20	21.4	<.001	.164	4.88	.009	.080	1.43	.234	.012
*22	95	.008	.929	.002	2.11	.126	.035	.090	.764	.001
23	13	16.6	<.001	.062	9.21	<.001	.136	2.81	.097	.024
*24	51	.051	.822	.007	7.25	.001	.109	1.22	.272	.010
25	15	17.6	<.001	.073	8.12	.001	.123	.115	.735	.001
26	15	26.5	<.001	.093	15.1	<.001	.206	2.55	.113	.022
*27	56	.110	.741	.004	1.62	.202	.027	8.77	.004	.071
28	11	59.3	<.001	.217	19.2	<.001	.250	.063	.803	.001
29	7	57.7	<.001	.263	8.64	<.001	.131	5.36	.022	.045
30	16	10.0	.002	.046	3.24	.043	.056	1.92	.168	.016
31	28	1.22	.271	.002	5.38	.006	.086	.237	.627	.002
32	10	29.1	<.001	.139	6.50	.002	.099	13.7	<.001	.107
*33	76	1.65	.201	.022	2.65	.075	.041	3.98	.048	.033
34	27	1.78	.185	.004	7.25	.001	.111	4.83	.030	.040
*35	100	16.1	<.001	.097	1.74	.180	.029	12.4	.001	.097
36	11	51.6	<.001	.210	12.5	<.001	.179	.531	.468	.005
37	12	12.4	<.001	.039	8.64	<.001	.132	1.11	.295	.010
38	10	22.9	<.001	.128	1.30	.276	.023	1.15	.287	.010
39	22	4.43	.037	.008	12.0	<.001	.173	1.85	.176	.016
40	50	.134	.715	.002	.459	.633	.009	1.31	.255	.011
41	32	3.74	.055	.003	9.77	<.001	.144	.345	.558	.003
42	14	11.0	.001	.036	9.53	<.001	.142	.327	.569	.003
43	12	11.3	.001	.025	14.2	<.001	.197	.277	0.60	.002
44	13	13.9	<.001	.041	10.8	<.001	.158	.008	.930	<.001
45	21	20.5	<.001	.093	4.92	.009	.077	1.78	.184	.015
*46	69	.111	.740	<.001	2.59	.080	.046	2.42	.122	.021
47	21	2.80	.097	.010	2.38	.097	.040	0.05	.823	<.001
48	10	20.7	<.001	.087	6.26	.003	.097	0.48	.490	.004
49	19	23.2	<.001	.130	1.60	.206	.026	1.99	.161	.017
50	3	42.9	<.001	.179	9.93	<.001	.145	1.92	.168	.016
51	47	.102	.750	.001	2.34	.101	.039	.242	.624	.002
52	16	22.1	<.001	.074	13.9	<.001	.189	12.2	.001	.096
53	13	27.0	<.001	.109	11.3	<.001	.163	2.13	.147	.018
54	20	5.19	.025	.017	4.59	.012	.072	6.03	.016	.050