DEVELOPING ADDITIVE RELATIONS WORD PROBLEM SOLVING WITH THE USE OF RELATIONAL MODELS

A thesis submitted to the Wits School of Education, Faculty of Humanities, at the University of the Witwatersrand in fulfilment of the requirements for the degree of Doctor of Philosophy

By

Herman M. Tshesane

Student number: 9108968T

Supervisors:

Prof Mike Askew

Prof Hamsa Venkatakrishnan

June, 2021

ABSTRACT

Two issues reported in the literature relating to South African learners' solving of additive relations word problems informed the adoption of a relational model-based approach for the design and implementation of this intervention study. The first issue relates to evidence that learners struggle to set up models of word problem situations that adequately express the sense they make of additive situations. The other relates to evidence of challenges in moving learners beyond the use of rudimentary counting strategies when calculating additive tasks. Given the evidence of both of these issues in the early grades in South Africa, I designed and implemented a relational model-based approach that orchestrates a problem-solving cycle that links models of situations and models of number. The need for designing such an approach was informed by a view of word problem-solving as a cycle involving two critical phases – an interpretation phase in which models of the situation are set up, and a calculation phase in which a model of number is used to calculate the answer – and a third, less critical, evaluation phase in which the answer is rendered discernible as such. To feed this need, I adopted the *Relational Paradigm* as a theoretical base whilst incorporating elements from Realistic Mathematics Education (RME) for both the design of the intervention and the analysis of the data. The *Relational Paradigm* refers to learning and teaching approaches that distinguish between the interpretation phase of the word problem-solving cycle, which involves the construction of a model of the situation described in the problem, and the calculation phase which involves the execution of the chosen operation based on the model of the situation described in the problem, thus making it possible for deliberate attention to be paid to either or both of the two critical phases of word problem-solving, for research and teaching purposes. The RME elements incorporated into the design and implementation of this study related to harnessing the sequential structure of number in order to promote the use of ten-based strategies that can be modelled as jumps on an empty number line.

Findings from this study point to the potential of this relational model-based approach to effect improvements in performance underpinned by improved interpretations of different word problem situations into models (mostly into bar models), and a move away from count-by-one strategies to calculation that makes use of the sequential structure of the number line with ten as a benchmark.

Key words: additive situations; additive relations word problems; relational model-based approach; models of situations; models of number; bar diagrams; empty number lines.

Appendix: detailed descriptions of intervention lessons

DECLARATION

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Herman Makabeteng Tshesane

29th day of June in the year 2021

PUBLICATIONS EMANATING FROM THIS RESEARCH

Takane, T.B., Tshesane, H. & Askew, M. (2017). From Theory to Practice: Challenges in adopting pedagogies of mathematising in South Africa. In M. Graven & H. Venkat (Ed.), Improving Primary Mathematics Education, Teaching and Learning (p. 179-197). England: Palgrave Macmillan.

Tshesane, H. (2018). An emerging framework to inform the learning and teaching of additive relations problem-solving. Proceedings of the 26th SAARMSTE Conference, Gaborone Botswana. 16-19 January. Pp.200-204.

GUIDING PRINCIPLE

We are determined to be masters of our fate; owners of our wealth; and capable of removing the adjective ‘dark’ from the name of our continent. Education – that of the nation and of the individual – is the bastion of this goal.

The solid basis for all lies in education. It is education which allows people to live together, and makes them avoid the pitfalls of immorality and induces respect for the law.

From truth alone is liberty born, and only an educated people can consider itself as really free, and master of its fate.

His Imperial Majesty Emperor Haile Selassie I

DEDICATION

To the thrice majestic Prophet, Priest and King Tehuti for the gift of writing and reading, and for bringing into human awareness the importance of knowing that, while duality is the basis of life, mastery resides in the ability to reconcile polarities.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank, from the deepest reservoirs of my heart, the following people for contributing to the completion of this work. Thank you for your love, guidance, support and encouragement throughout this journey.

My ancestry: Batubatse ba Mohlogopela, Bakgatla ba Mmakau le Ditlou tsa ga Bambo

My parents: the Late and Great Kgoropedi-a-Marumo Johannes & His Lion-hearted Queen, Kutwadi-a-Malope Ngu Gemma Tshesane

My elder brothers and sister: The late Thomas Athanasius and Ernest Kgoloane Tshesane, and the living George Puledi and Mmakgabo Elizabeth Tshesane

My Queen: Esther Lebele

My children: Kholofelo, Mmatefo, Kgatledi, Lesedi and Realeboga.

My supervisors and mentors: Prof Hamsa Venkat & Prof Mike Askew

My family @ the Wits Maths Connect Primary Project: Corin, Thulelah, Marie; Lawan; Lincoln; Samantha; Sameera; Manono (late); Valerie; Nicky; Nomonde & Constance.

My brothers @ the Karabo Centre for Maths: Collin Forka & Maanda Mphohoni

I also would like to thank the principal, teachers, children and their parents in the primary school in Johannesburg for opening their doors and hearts for us to conduct this research with them as willing participants.

The financial assistance of the National Institute for the Humanities and Social Sciences (NIHSS), in collaboration with the South African Humanities Deans Association (SAHUDA) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NIHSS and SAHUDA.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BD, Bar Diagram

CM, Column Model

ENL, Empty Number Line

EDA, Equilibrated Development Approach

LFiN, Learning Framework in Number

NS, Number Sentence

RME, Realistic Mathematics Education

SA, South Africa

TL, Tallies

WMC-P, Wits Maths Connect – Primary

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTORY CHAPTER

1.1 Introduction

In the context of solving additive relations word problems, a dual focus on learners' awareness of the mathematical structure and calculation of the numerical solution provided this study's point of departure. The decision to attend to these foci was based on evidence pointing to particular challenges among South African (SA) learners relating to difficulties with, on the one hand, making sense of mathematical word problems (Sepeng & Sigola, 2013), and, on the other, moving beyond counting based calculations (Ensor, Hoadley, Jacklin, Kuhne, Schmitt, Lombard & Van Den Heuvel-Panhuizen, 2009). This led to this dual focus being at the heart of an intervention study carried out over three months with a SA Grade 3 class.

My hypothesis as I began this study was that learners who have appropriated and mastered the use of the bar diagram (BD) and the empty number line (ENL) as relational models in the context of solving additive word problems would be well placed to use number sense to solve additive tasks. I refer to the BD and the ENL as relational models because these models have been described as having the potential to facilitate reasoning about additive relations and number relations (Gravemeijer & Stephan, 2002) in the context of the solving of additive word problems. In particular, the BD was included in order to support reasoning about the relations between the quantities that make up the mathematical structure of additive word problems, and the ENL to support reasoning about the relations between the specific numbers involved. My theoretical assumption when designing this intervention was that the BD has the potential to support the sense-making (interpretation) of the problem situation and the ENL the efficient calculation of the solution.

In this introductory chapter I set the scene for this study. First I provide a historical overview on the developments in the educational terrain in democratic SA as a background to the current state of mathematics learning in South African government schools. Against this backdrop, I then outline the challenges experienced by learners relating to additive relations and early number working, and juxtapose these challenges with an overview of ways of working for supporting sense-making of the mathematical structure of word problems and efficient calculation, foregrounded by international research. Detailed research-based studies of interventions that may contribute to addressing these challenges in the South African terrain remain limited on the ground (Taylor, 2021). With the research gap that this study looks to fill outlined, I go on to indicate my theoretical formulation and briefly outline the Relational Paradigm (Polotskaia, 2017) as I see it as ideal not only for integrating attention to structure and calculation, but also for serving as the basis of a modelling approach to additive word problem solving. Part of the discussion put forward establishes theoretical links between the two relational models promoted in this study (namely: the BD and the ENL) and the two foci of supporting learners to:

- make sense of the problem situation by reflecting on and modelling the part-whole structure of the additive relationship; and
- calculate the numerical solution to these problems by reasoning about the quantitative relations between the specific numbers involved

Following this theoretical framing, I provide a succinct explanation of the aim of this study, leading into the research questions. There follows an account of the methodological considerations that underpinned the design and implementation of the study, as well as the analysis of the data collected for this study. I conclude with a brief outline of the chapters making up this thesis.

1.2 Historical background to the levels of performance in Mathematics in SA

In 2011 the Wits Maths Connect-Primary project conducted baseline tests with Grade 2 learners and found that three quarters of the cross-attainment sample of learners assessed used rudimentary strategies in early addition problems (Venkat, 2011). By rudimentary strategies, I mean those count-by-ones strategies found by researchers to be inefficient and error-prone (Ensor et al., 2009; Schollar, 2008). Also in 2011, a large scale international study that focused on Grade 8/9 Mathematics and Science – *Trends in International Mathematics and Science Study* (TIMMS) – found that three quarters of Grade 9 pupils in SA still had not acquired a basic understanding about whole numbers, decimals, operations or basic graphs (Spaull, 2013). The quasi-longitudinal view of learner performance provided by these studies suggests that many learners proceed to higher grades without appropriating foundational skills in numeracy and literacy. This lack of grade-appropriate skills has been described as resulting in a 'large cognitive backlog that progressively inhibits the acquisition of more complex competencies' (Schollar, 2008, p. ii) . If left unchecked, the cognitive backlog that learners begin to incur in primary school can become insurmountable at secondary school level. In turn, this is likely to militate against learners' ability to make sense of the content at higher grades, resulting in poor learner performances and high drop-out rates. Taken together, the implications from the larger international comparative studies and the smaller more in-depth studies overviewed above is that the majority of SA learners are achieving performances that are well below those of their counterparts in the Southern and Eastern regions on the African continent as well as internationally.

In particular, one problem identified in the *Report on the Annual National Assessments* compiled by the Department of Basic Education (DBE, 2012) and found by Venkat (2013) to be a contributing factor to the existing infidelity between what the curriculum prescribes and the levels of performance of a great many SA learners is inefficient ways of working with number. What makes this a serious problem is that, not only are these inefficient ways of working with number seen in learners' working in the Foundation Phase (FP), they have been found to persist well into the Intermediate Phase (IP) of their school careers (DBE, 2012).

I now expand on this poor learner performance and the cognitive backlogs that inform this poor learner performance, resulting in high dropout rates and, hence, the need for early identification and remediation of these learning deficits, before I elaborate on these inefficient ways of working with number in the section that follows .

1.2.1 The educational terrain in post-apartheid SA

In the dispensation following the advent of democracy in South Africa, the very first cohort of Grade 1 learners was enrolled in 1995. In 1998, when this cohort of learners was in Grade 4, Outcomes Based Education (OBE) was introduced. OBE represented SA's initial attempt at educational reform as a way into transforming how teachers organise their classes and their teaching in order to optimise learning (Askew, 2019).

Eight years later, when this first cohort of learners sat for Matriculation examinations under the new OBE curriculum, data indicated that of the 1 676 273 learners enrolled in Grade 1 in 1995, less than one third (528 525) completed the 12 grade schooling system to enrol for matric in 2006. Only a fifth (330 513) of the overall Grade 1 cohort actually wrote the National Certificate Exams and less than two percent (25 217) of the learners who wrote Mathematics achieved a pass at a Higher Grade (HG) level (Fleisch, Schöer, Roberts, & Thornton, 2016).

These numbers pointed to widespread drop-out before Grade 12 that contributed to low numbers of learners exiting the school system with a pass in Mathematics. Recently, the Curriculum and Assessment Policy Statements (CAPS) replaced OBE in order to prioritise sequencing and pacing of content, with, in some provinces, scripted lesson plans designed and provided at district level (Fleisch et al., 2016).

1.2.2 Comparative assessments of SA learners' performance

Matric results are no longer the only national measure of the outcomes of the school system. We now also have reliable data on learner performance from two large-scale international assessments of educational achievement in which South Africa participates. These are the *Southern and Eastern African Consortium for Monitoring Educational Quality* (SACMEQ) and the *Trends in International Mathematics and Science Study* (TIMSS). SACMEQ is a study that covers 14 countries in Southern and Eastern Africa, including Zimbabwe, Mozambique, Zambia, Uganda, Tanzania, Kenya and South Africa. Albeit not an annual study, it focuses on Grade 6 numeracy and literacy and there is currently data from 2000 and 2007. TIMSS is an international study that focuses on Grade Five/Eight/Nine Mathematics and Science and has provided data on learner performance for the years 1995, 1999, 2002, 2011, 2015 and 2019.

1.2.2.1 Evidence of poor performance from international comparative large-scale studies

The SACMEQ II and SACMEQ III studies revealed no improvement in South African Grade 6 learners' literacy or numeracy performance over the seven year period from 2000 to 2007, with South African learners ranked 10th for reading and 8th for mathematics out of 14 education systems on the continent. In particular, the study found that 27% of these learners could not read a short and simple text and make sense of it, a deficit with serious implications for reading and making sense of word problems in mathematics.

The TIMSS studies showed that, despite noticeable improvement in Mathematics and Science performance amounting to approximately one and a half grade levels of learning over the 2002-2015 period at the Grade 9 level, the level of performance following this improvement remains low, largely because the baseline results of 2002 were exceedingly low. With achievements reported at four international benchmarks of Low (400), Intermediate (475), High (550) and Advanced (625), only learners in countries with a minimum benchmark of 550 points are deemed able to apply conceptual understanding to solve problems. In the TIMMS 2019 results, South Africa came in at 374 points and top-ranking Singapore at 625 points. In particular, where application of concepts and reasoning in different problem situations was concerned, Singapore had 51% of its 8th grade learners reaching the advanced benchmark, a very high value compared to South Africa's 1%. Compared to learners from other middle-income countries, SA Grade 9 learners were found to perform up to four grade levels lower. Whilst I have focused on assessments up to 2015 as this was the period of my empirical data collection, no change was found in the most recent TIMMS administered in 2019 at Grade 4 level.

With short term trends for Grade 4 showing 14 of the 45 countries that participated in the 2015 study with higher average achievements and 8 countries with declines, South Africa was one of 23 countries that showed no change in average achievement between the years 2015 and 2019. Particular challenges were reported regarding the application of whole number concepts to solving step word problems and working with the number line where only 1% of South Africa's Grade 4 learners attained the advanced benchmark, a very low value compared to Singapore's 54%. Not surprisingly then, for two consecutive TIMMS studies, Singapore has topped the TIMMS rankings whilst South Africa remains amongst the lowest of all participating countries.

Whilst the usefulness of large-scale studies is in providing a broad sense of the low levels of the performance of SA learners' relative to other countries, the evidence from more in-depth small-scale studies is useful for providing the possible reasons for this poor performance.

1.2.2.2 Evidence of poor performance from in-depth small-scale studies

The evidence from small-scale studies conducted within the two SA Numeracy Chair projects, one based at Rhodes University (RU) and the other at the University of the Witwatersrand (Wits) points to inefficient ways of working with number as underlying poor learner performance and the learning deficits associated with it[1].

The Wits Maths Connect-Primary (WMC-P) project was established with the support of the government with the objective of improving the *number sense* of learners' in Grades R to 6. Using the Learning Framework in Number (LFIN) developed by Wright, Martland, and Stafford (2006) to enable assessing learners' proficiency with tackling additive tasks, the WMC-P project administered oral assessments that took the form of interviews on Grade 2

[1] These Numeracy Chairs are research and development projects partially funded by government with the aim of supporting the development of numeracy/mathematics teaching and learning.

learners in the ten schools, drawn from across the attainment range, the results of which were meant as to provide baseline insights into Grade 2 learners' early number sense. The results revealed that 75% of the learners assessed relied on counting (Venkat, Askew, & Morrison, 2021) when required to add or subtract two numbers. In particular, the use of counting as a strategy to add or subtract two numbers was observed to involve 'triple concrete counting' (Anghileri, 2006) where each number is counted out separately and then the total is counted.

In summary, up to and including 2019, research into the learners' performance in numeracy shows the majority of the learners in the system performing significantly below the levels of the stipulated curriculum. With SA Grade 3 learners found to be 1.8 years behind the benchmark, and with these learning deficits found to increase by about a year between Grades 3 and 9 (Spaull, 2013), the evidence points to the learning deficits experienced by learners in secondary school as having their roots in ways of working with number established in primary school. It was necessary, therefore, to identify the specific challenges related to poor performance as a first step to intervening early and appropriately with a view to addressing these learning deficits before they are rendered insurmountable.

1.3 Specific challenges related to poor performance in the school system

While a number of challenges related to primary schooling have been highlighted in the literature on primary mathematics learning (e.g. (Fleisch, 2008)), the key ones relate to issues of progression, teachers' insufficient knowledge of content and teaching (Askew, 2019).

I deliberate here on two main challenges, both of which were observed by Ensor et al. (2009): the limited and limiting use of artefacts by teachers in instruction, and the lack of progression as it relates to the requisite shift from counting to calculating. Whilst the former challenge has been found to lead to limited exposure to models and representations that can support the sense that learners make of mathematical problems, the latter challenge means that learners struggle to move beyond a conception of numbers as outcomes of counting processes towards viewing numbers as objects which are composed of other smaller numbers and which can be decomposed into other smaller numbers.

Given that moves to abstract representations of number underpin moves to abstract methods of working with numbers, I address myself to the former challenge before turning my attention to the latter.

1.3.1 Lack of progression from concrete to abstract forms of representing number

The predominant use of unit-count tally marks by learners' in their solutions not only points to a lack of progression from a conception of numbers as outcomes of counting processes, towards viewing numbers as objects which can be composed and decomposed in accordance with certain established laws; it also suggests a lack of progression towards more abstract/formal models and representations of number. This lack of progression towards more abstract/formal models and representations of number is unsurprising given the evidence provided by Ensor et al. (2009) of limited instances of teaching that presents and encourages learners towards more abstract symbolic representations.

The failure on the part of teachers to support learners to make the shifts from concrete to abstract forms of representation is seen in SA classroom practices that have been described as valuing right answers only, with limited, if any, attention to making sense of situations described in word problems (Barnes, 2010). When Sepeng and Sigola (2013) conducted an investigation into the types of errors that Grade 9 learners make when solving mathematical word problems, they traced the source of these errors to classroom practices that prioritise number-manipulation at the expense of sense-making. This situation contrasts with the growing advocacy for shift in the international terrain towards valuing the ability to represent information gathered from a contextualised problem and to judge the true meaning of information as much as the ability to judge the meaningfulness and the correctness of calculations (Gravemeijer, 1997). In particular, Ensor et al. (2009) observed an absence of teaching moves geared towards increasingly abstract symbol-oriented artefacts, whilst noting the presence and use – up to Grade 3 level – of concrete tools that support counting.

This valuing of the concrete over the abstract for representing number is reported on across a range of South African research conducted in the classroom that shows the prevalence of counting methods as seen in the use tallies by learners when working with number in the early grades (Hoadley, 2007; Schollar, 2008). The teaching that manifests this valuing of the concrete over the abstract for representing number is characterized by a tendency on the part of teachers to accept concrete counting as the primary method for working with number (Venkat, 2013). More generally, it has been described as teaching that is accepting of learners' current ways of working with number, whilst offering little in terms of opening up avenues that lead to holistic understandings of the concepts they are currently learning.

1.3.2 Lack of progression from concrete counting to abstract calculating

With the curriculum and teaching in South Africa presenting counting as a way into introducing numbers and calculation, Ensor et al. (2009) delineate a progression from counting, through calculation-by-counting, to calculation without counting as encapsulating a shift towards more structured calculation. They posit that a shift from using counting towards more structured calculation should accompany a broader shift towards more abstract methods of solving problems. They argue, therefore, that learners' tendency to persist with unit counting – as exemplified by the majority of the grade 5 and grade 7 learners that participated in Schollar's (2008) study who relied on 'simple unit counting' to solve additive tasks – is symptomatic of teaching that pays little attention to modelling calculation-without-counting in the lower grades in order to progress learners towards more abstract methods of solving problems.

Given the widespread prevalence of 'triple concrete counting', more efficient counting strategies have been suggested, including *counting-on*, *counting-down-from* and *counting-down-to* (Carpenter, Fennema, Franke, Levi, & Empson, 1999). Be that as it may, only 24% of the learners tested by the WMC-P project as part of their baseline test showed insufficient ability in using *count-on* strategies in the number range 1 – 20. With even smaller proportions could work with 5s and 10s as benchmarks, the results of the tests returned similar findings to Schollar's (2008) middle grades' finding of learners' overreliance on unit counting as seen in responses to additive tasks dominated by tallies.

My dual interest in learners' ability to discern the mathematical structure of additive word problems and in learners' efficiency in their calculation of the solution to these problems stems, therefore, from insights provided by work done by the WMC-P project over two five year instalments of research into improving primary school learners' number sense. In the first instalment (2011 – 2015), with the use of the LFiN assessments, the work provided insights into how increasing attention to structure in teachers' work can significantly increase the proportions of learners who are able to use count-on, rather than count-all strategies. In the second phase (2016 – 2020) of the project, there was a shift away from an approach focused on supporting reified counting towards an emphasis on number relationships to support calculation that makes use of the structure of numbers for efficient calculation. With this shift in emphasis came substantial shifts away from incidents of counting towards working with number structure (Venkat et al., 2021).

1.4 The theoretical framework adopted in this study

Literature identifies two types of additive tasks that learners encounter in school: addition-and-subtraction tasks and additive word problems (Polotskaia, Savard, & Freiman, 2015). Whilst the former type only requires calculation, the latter requires the solver to first interpret the additive relationship in the problem situation, before attempting to calculate the addition/subtraction inherent in that relationship.

With a view to providing Canadian learners in the early grades with experiences of number that would balance the already strong calculational work with some interpretational work, Polotskaia et al. (2015) developed a conceptual framework in line with the principles of the Relational Paradigm to inform the problem-solving process of additive relations word problems. According to this conceptual framework, the solving of additive word problems is characterised as a cyclic process that takes the problem from the everyday context into the mathematical context and then back to the everyday context. This conceptual framework, which will be discussed in more detail in Chapters 2 & 3, informed my sequencing of the use of the relational models in line with the order of the phases in this cyclic process. Here I provide a brief outline.

1.4.1 The Relational Paradigm for structural and numerical reasoning

For Savard and Polotskaia (2017) the interpretation of additive word problems centres on understanding the problem situation by creating 'a graphical model' of the relationships between the quantities involved. This graphical model then informs the arithmetic operation to be used in the calculation sequence that follows. The calculation sequence is then carried out with an eye for efficient calculation based on the relationship displayed in the graphical model. It was proper, therefore, that any intervention aimed at improving competence with solving additive word problems should attend to improving learners' interpretation of the additive relationship involved, and the calculation implied by the additive relationship. One benefit of interpreting the structural relations into a model was to discourage learners from the widely described move of modelling the action directly into 'guessed' versions of a number sentence.

In the context of solving additive word problems, the required reasoning takes two forms: namely, structural reasoning and numerical reasoning (Polotskaia, 2017). As elaborated upon in section 1.5 below, structural reasoning relates to a holistic understanding of the additive relationship that enables one to express this relationship as a part-whole relationship; numerical reasoning relates to a holistic understanding of part-whole relationships between numbers. An important aspect of numerical reasoning for this study relates to leveraging the base-ten structure (including using the complements of ten) and other number relations to support efficient calculations. In line with the dual focus of this study on developing the necessary structural reasoning to support the interpretation of the additive word problem situation, and numerical reasoning to support calculation that pays attention to relationships between numbers, the two graphical models promoted in this study needed to serve as tools that could support these processes. These models will be discussed in more detail in chapter 2, but for now I provide a sense of the *Relational Paradigm* which integrates this dual focus.

The Relational Paradigm refers to learning and teaching approaches that prioritise quantitative relationships and thinking relationally in mathematics (Freiman, Polotskaia, & Savard, 2017). The Relational Paradigm can be distinguished from traditional approaches in the emphasis that its protagonists place on the primary role of interpreting problem situations as systems of relationships and creating models of those relationships. In accordance with the Relational Paradigm, learners' success with solving additive word problems depends on their ability to make sense of the additive relationship by interpreting it into a model and deriving the operation involved as part of the interpretation (modelling) phase, before using the number sentence incorporating the operation as the basis of calculating the solution to the word problem in the calculation phase. Whilst within the Relational Paradigm the calculation is not the learning target of problem-solving activity (Polotskaia, 2017), in a South African context where counting in ones remains prevalent with little evidence of teaching that exemplifies using relations between numbers for efficient calculation, paying attention to the calculation of the solution was as critical as attending to the interpretation of the situation in the intervention at the centre of this study.

Given my concerns over the lack of attention by teacher to using models to bring out structural relations in additive working, I used the BD model in this study to support the interpretational move from the everyday context of additive word problem situations to the mathematical context. In the mathematical context, an addition/subtraction statement was then used to introduce the arithmetic operation required by the additive relationship captured in the BD. Given my concerns with the lack of progression towards using number relations and base-ten structure, the ENL was introduced to calculate the numerical solution. In moving from the mathematical context back to the everyday context, the numerical solution would then be evaluated in light of the context in the word problem before it is offered as the solution to the problem.

To summarize, in line with my theoretical assumption that relational models can support the sense-making and efficient calculation necessary to solve additive word problems I chose to selectively promote models that can support making sense of the mathematical structure of additive situations and models that can support efficient calculation of the numerical solution. In particular, I identified the BD as a model that can support making sense of the additive relationship that constitutes the mathematical structure of additive word problems, and the ENL as a model that can support the efficient calculation of the numerical solution, with an appropriate number sentence derived from the BD acting as a bridge between the BD and the ENL.

1.4.2 Looking within and between phases of the problem-solving cycle for coherence

In moving from the interpretation of the problem situation as expressed in the BD, into the calculation of the solution with the use of the ENL the idea was for the models to coalesce in the process of solving a given additive word problem. The underlying assumption was that both models would have had to be appropriately configured to reflect the same additive relationship for a solution process to yield the correct answer. This necessitated guiding learners operations of thought towards grasping the additive relationship involved in the problem situation (structural reasoning) in order to derive the arithmetic operations necessary to calculate the solution (numerical reasoning) in an appropriately coordinated manner (Giroux & Ste-Marie, 2001) during intervention lessons. So in a modelling approach to solving additive relations word problem, a successful engagement would have to reflect models that are appropriately configured to reflect the same additive relationship, i.e. in relational models that were coherent, not only within the individual phases of the problem-solving cycle, but also between them, with the models reflecting the same correctly interpreted additive relationship.

1.5 The aim of this study

The study, therefore, was aimed at determining the contributions to additive relations problem solving competence that can be effected through an intervention that focuses on the bar diagram (BD) and the empty number line (ENL) as relational models. Whilst an ancillary motivation for embarking on this study was the lack of South African data on learners' performance on early number approaches as noted by Ensor et al (2009), the main motivation was to determine the extent to which the use of BD and the ENL can respectively support the interpretation of additive relations problem situations and the calculation of additive word problems by structuring the numbers involved.

In designing and implementing this intervention study, my interest was in the extent to which learners at the end of the Foundation Phase (FP) could be supported to make sense of additive relations situations and to calculate the solution to the additive word problem through the use of relational models chosen specifically for those two purposes. The intervention study adapted and implemented the Equilibrated Development Approach formulated by Polotskaia and colleagues within the Relational Paradigm (Polotskaia, 2017) to incorporate attention on the calculation phase alongside these authors' emphasis on the interpretation phase whilst using word problem tasks from the Big Books' series by Askew (2004) designed support a generic view of the different situations seen in different additive relations word problem types.

The approach was preferred, firstly, for the greater emphasis placed in the Relational Paradigm on interpretations of problem situations into models, and; secondly, for the explicitly distinguishing between the interpretation phase which involves the interpretation of the problem situation into a model and the calculation phase which involves the execution of the chosen operation as 'the key element' of working within the Relational Paradigm (Polotskaia, 2015). This enabled incorporating explicit instructional focus on BD and ENL as key relational models, as well as to including dedicated time at the beginning of the lesson to focus explicitly on introducing and discussing combinations of 10 and 20 in BD and ENL formats. Ultimately, the exploration of this study was into the extent to which the BD and ENL combination can serve as a meta-cognitive tools in a relational model-based approach designed to support the solving of additive word problems with structural and numerical sense.

1.5.1 The two dimensions of the part-whole concept and their relevance to this study

In order to provide the study with a lens that could help me infer the extent to which learners' workings reflect structural and numerical reasoning, I appealed to Ellemor-Collins & Wright's (2001) work on structuring. In the interpretation phase, my focus was on the *part-whole construction of the additive relationship* to operationalise making sense of the additive relationship in a word problem. Within this phase, as noted above, I became interested in the extent to which learner responses included appropriately configured models. In the calculation phase, Ellemor-Collins & Wright's (2001) work on *structuring number* as 'a bridge from counting-based calculation to formal, facile calculation' (p.56). I draw from their work on structuring numbers around 10 as a reference point, and I use their notion of a *part-whole construction of number* to operationalise the type of working that relates numbers to other numbers instead of counting. In this study, therefore, the ultimate goal was to develop in learners a sense of the notion of parts within a whole as an idea that can be operationalised in two ways: structurally (quantities embedded within quantities) and numerically (numbers within numbers). In assessing this development, learners' responses in the pre-test, post-test and the delayed post-test were analysed.

With the broad aim outlined above in mind, my investigation was driven by the following specific research question that is elaborated into three sub-questions:

1.6 The research questions

The overarching research question is:

What are the contributions to additive relations word problem-solving competence that can be effected through an intervention that focuses on the bar diagram (BD) and the empty number line (ENL) as relational models?

Sub-questions within this are:

a) What were the overall patterns of learner performance across different types of additive relations problems prior to the intervention, and what shifts in performance were observed from pre- to delayed post-test?

b) What nature and patterns of model use underlay the difference in learner performance across the three tests, and to what extent was there evidence of models that reflected appropriate interpretation of the mathematical structure and correct calculation of the numerical solution?

c) What was the nature of the calculation strategies used by the learners across the three tests and to what extent did the use of these strategies contribute to efficient calculation of the numerical solution?

1.7 Methodology

To roll out the intervention I approached a teacher from one of the ten schools participating in the WMC-P Project for permission to work with her Grade 3 class of 39 learners. The teacher, having participated in the broader WMC-P development activity, was willing and able to avail her Grade 3 class, and Grade 3, being the exit point of the Foundation Phase, was an appropriate grade to work with. Having got her agreement, I received consent from the Principal to work in the school, subsequent to which I approached the learners in the class, and their parents, for their consent to participate in the study. The school was located in one of the suburbs in Johannesburg, and, English – which was not the learners' home language – was the medium of instruction. More importantly, the school had been categorized by education authorities as 'underperforming' based on poor learning outcomes in language and mathematics.

The intervention was styled as a 'teaching experiment' (Steffe & Cobb, 2012) with a class of Grade 3 learners. An individually written pre-test was administered followed by ten two-part intervention lessons with this Grade 3 class. The lessons were each about an hour long, with, on average, two lessons taught per week. Following the intervention lessons the learners sat for a post-test. Two months later, and two weeks into the first term of their Grade 4 year, the learners sat for a delayed post-test. I describe this process in detail in Chapter 4.

The data that I analyzed came from the pre-test, post-test and delayed post-test that the learners wrote. In analysing the data for answers to my first research sub-question, my focus was purely on the answers provided by the learners. This allowed for delineation of changes in learner performance across the tests and across the additive word problem types that the learners had to solve in each of the three test sittings. For my second research question, my focus was on the models used by the learners in their responses, with particular interest in the extent to which the learners were able to correctly interpret the problem situation based on the appropriateness of their part-whole construction of the additive relationship, as well as the extent to which they could correctly calculate the sum or difference implied in the additive relationship. In this regard, the appropriateness of the part-whole construction of the additive relationships was used as an indicator of learners' ability to make sense of the additive relationship involved. Where my third research question was concerned, I looked at learners' part-whole construction of number in their calculations. In this regard, the appropriateness of the part-whole construction of number was used as an indicator of learners' ability to operate on the additive relationship in the additive word problems without the need for counting.

Additionally, given the support provided to learners during intervention lessons for engaging in the practice of constructing relations not only *within* individual models of situations but also *between* the part-whole construction of the additive relationship and the calculation of the solution to the problem, I also explored learner responses for appropriateness of the coordination of the interpretation and the calculation. With literature pointing to the importance of coordinating one's interpretation and calculation for a solution process to yield the correct answer (Giroux & Ste-Marie, 2001), the extent to which this coordination could be seen in learners' responses was a measure of the extent to which the intervention was a success. In other words, based on the prevalence (or lack thereof) of responses with problem situations correctly interpreted into models and correct calculations culminating in correct answers I could infer the extent of success of the intervention.

Ultimately, I reflect on the contributions to additive relations problem solving competence that were effected through this intervention study give the learners' differential performance over time and the nature and patterns of model and strategy use that underlay such performance. By implication, the success of this intervention study is a function of the extent to which learners' responses in the post- and the delayed post-test were dominated by solutions with models that reflected representational coherence between the models used, with the models reflecting the same correctly interpreted additive relationship. Despite the short-term nature of the intervention, the results reported in this study suggest that, in the context of a relational approach that promotes the use of relational models in order to support structural and numerical reasoning, improvements in performance can still be achieved. This study will show that these improvements in performance were underpinned by improvements in interpretation, calculation and the coordination of structural and numerical reasoning brought to bear during problem-solving as supported by relational models. This has implications both for teaching and learning additive relations word problems, as well as for future research.

1.8 Outline of thesis

The purpose of writing this chapter was to introduce the area of interest in my research, provide the background of the study, the rationale as informed by literature, as well as the research questions that guided the investigation. Deliberations in this chapter have highlighted the importance of a dual focus on learners' awareness of the mathematical structure and calculation of the numerical solution in the context of solving additive relations word problems. In providing a rationale for conducting an intervention study with a dual focus on developing in learners the necessary structural reasoning to support the interpretation of additive word problem situations, and numerical reasoning to support calculation that pays attention to relationships between numbers, I identified particular challenges among SA learners relating to difficulties with making sense of mathematical word problems, and moving beyond counting based calculations. This rationale in turn informed the theoretical framework in which the current study was conducted.

In chapter 2, I outline my literature-informed theoretical orientation, focusing on how literature defines and promotes the use of models for interpretation and efficient calculation strategies for calculation in the context of additive word problem-solving, as well as the theoretical stance taken regarding models and how they can support learning and teaching. In this regard, I attend to the ways in which I conceptualize the use of relational models for teaching and learning with as part of my theoretical framework.

After presenting the literature base and theoretical orientation and discussing the necessary adaptations to it, chapter 3 details the methodology that I adopted for my study, providing a description of the design of the research including the design of the instruction during intervention lessons, and culminating in the analytical framework for this study. Accompanying these descriptions are explanations of the techniques used for data collection and for coding responses for analysis. Also attended to in this chapter are matters of validity and reliability, where issues relating to ethics are also addressed.

The analysis of the data is reported in Chapters 4, 5, and 6 where I also discuss the findings that emerged in light of the relevant literature. Finally, in Chapter 7, I provide a synopsis of my thesis and what I found to be the contributions to knowledge and limitations of the study, while highlighting implications of the findings, with recommendations for future directions.

CHAPTER 2

LITERATURE BASE AND THEORETICAL ORIENTATION

2.1 Introduction

I believe that the interpretation of the problem situation and the calculation of the solution to the problem are two critical and complementary dimensions that constitute the solving of additive relations word problems that need to be supported if learners are to solve additive word problems with both structural and numerical sense. National and international research reports on two key findings where learners' solving of additive word problems is concerned; firstly, research reports on a tendency by learners to suspend the sense they need to make of situations described by the text of these problems (De Corte, Verschaffel, & Greer, 2000; Sepeng & Sigola, 2013), and; secondly, challenges in moving beyond the use of rudimentary counting strategies (Ellemor-Collins & Wright, 2009; Weitz & Venkat, 2013).

Giroux and Ste-Marie (2001) identified the processes of structuring the relations between the quantities in a word problem and the arithmetizing of this structural relationship as two critical processes that play complementary roles in a successful engagement with additive word problems. Drawing from Vergnaud's (1982) notion of a *relational calculus* that is concerned with the structuring of the relationships between the quantities in the problem situation, and a *numerical calculus* that facilitates the *arithmetizing* of these relationships, they posit that a successful engagement with additive word problems requires a dialectic interplay between these two distinct processes:

> By "numerical calculus" I mean ordinary operations of addition, subtraction, multiplication, and division. By "relational calculus" I mean the operations of thought that are necessary to handle the relationships involved in the situation (Vergnaud, 1982, p. 40).

Giroux and Ste-Marie (2001) argue that the difficulties that learners experience when solving additive word problems emanate from an inappropriate coordination of the operations of thought required to handle the relationships involved in the problem situation (the relational calculus) and the arithmetic operations necessary to calculate the solution (numerical calculus). The position taken in this study is that the relational calculus is primarily brought to bear whenever the need to interpret (or make sense of) the mathematical structure of the problem at hand – and to transform the additive relationship into a number sentence incorporating an arithmetic operation – arises, and that it manifests as structural reasoning. The numerical calculus is brought to bear whenever the need to calculate the numerical solution arises, and manifests as numerical reasoning. I therefore, define structural reasoning as reasoning on the structural relationships involved in the problem situation in attempts to make sense of, and arithmetize, the additive relationship; and numerical reasoning as reasoning on the relationships between the numbers involved, in the process of calculating the numerical solution to the problem.

Given evidence of issues with both interpretation (sense-making) and calculation in the South African early grades (see Takane, 2021), it was necessary to review both the literature base that emphasises interpretation into models of the structural relationships between knowns and unknowns in problem situations (Chan, 2009; Lesh & Caylor, 2007; Ng & Lee, 2005; Polotskaia, 2017; Verschaffel, Van Dooren, Greer, & Mukhopadhyay, 2010; Xin, 2012) as well as the literature base that emphasises working with the relational structure of number, and how this could be leveraged for progress towards increasingly efficient calculation of answers (Baroody, 2006; Blöte, Klein, & Beishuizen, 2000; Carpenter et al., 1999; Ellemor-Collins & Wright, 2009). Literature on the solving of additive relations word problems also reports on suggestions by researchers for a change towards a model-based approach to problem-solving (Gravemeijer, 2007; Greer, 1993; Lesh & Harel, 2003; Lesh & Lehrer, 2003; Mukhopadhyay & Greer, 2001; Van den Heuvel-Panhuizen, 2003; Verschaffel et al., 2010; Xin, Wiles, & Lin, 2008).

In studying this literature, I became interested in conducting this intervention study within a theoretical framing that incorporates an approach that pays explicit and deliberate attention, both in the design and implementation, not only to learners' calculations of their solutions, but primarily to their interpretation of the problem situation into models. Moreover, I was interested in a theoretical framing that makes use of models to support the structural reasoning required for interpretation and the numerical reasoning required for calculation. For these reasons, and for consistency in referring to learners' interpretation of the problem situation and efficiency in their calculation of the solution to the problem, I followed Polotskaia's (2017) conceptualisation of the problem-solving process and identification of these two critical and complementary dimensions of working with additive relations word problems as the 'interpretation' and the 'calculation' phases respectively. In the interpretation phase, a model expressing the structural relations between the quantities is constructed and a number sentence incorporating an arithmetic operation is derived from the model, and; in the calculation phase, the quantities in the additive relationship expressed in the number sentence are processed as numbers unencumbered by the problem situation, to produce the solution to the problem.

For their potential to facilitate reasoning about additive relations and number relations (Gravemeijer & Stephan, 2002), I refer to the models promoted in this study as ***relational models***. In particular, the spatial part-whole structure of the BD facilitates reasoning about the additive relations between the parts and the whole that make up the mathematical structure of additive word problems, while the sequential structure of the ENL facilitates reasoning about the number relations between the numbers involved. The BD and ENL models were introduced and promoted during the intervention with the deliberate intention that they should mediate the interpretation and the calculation phases of the solving of additive word problem respectively.

Having described the interpretation of the problem situation and the calculation of the solution to the problem as two critical and complementary phases of word problem-solving, I review in the first half of this chapter, the antecedent literature relating to each of them. I also describe key mathematical models engaged with in the literature as particularly useful for focus on additive relations across the two different phases. This first section of the chapter provides useful pointers relating to the different views regarding emphases placed by different literature bases on interpreting problem situations into models for sense-making, and working with the relational structure of number for efficient calculation. I also look at the key models of additive relations that were incorporated into classroom instruction for supporting structural and numerical reasoning. On this account, I begin with the interpretation phase, reviewing a series of studies that have attended to the challenges that learners experience with interpretation. I then turn my attention to groups that have focused on the challenges that learners experience with calculation. Theoretical positions on word problem-solving are interwoven with the literature-based claims and findings emanating from groups working in the main 'camps' that are discussed in this chapter.

In the second half of the chapter, I deal with the literature on the additive word problem types. This literature base provided the basis for my selections of tasks for pre- and post-tests, and also for the intervention lessons. This literature base also provided evidence of further aspects of complexity and progression within learners' additive working with word problems that provided additional lenses for interpreting differences in learners' responses across the pre and post-tests.

2.2 The interpretation phase

As outlined above, the focus of attention in the interpretation phase is on the structural relations between known and unknown quantities in a problem situation. While there is broad agreement in the literature that this kind of interpretation is at the heart of what it means to 'make sense' of a problem, there are differences between research groups on how interpretation should be supported. In this section, I overview the areas of agreement and dissonance, as well as the implications of positions on how interpretation occurs for the kinds of instruction that are advocated. Within this, I attend to the part-whole bar model as a key relational model that is suggested as particularly useful for the interpretation of structural relationships between quantities in additive situations.

2.2.1 Literature-based views on the interpretation phase

I have already noted the agreement across several bodies of work on word-problem solving that interpretation of the relationships between quantities in the problem is a critical precursor to being able to calculate the answer (Ng & Lee, 2005; Baroody, 2006; Gravemeijer, 2007; Verschaffel et al., 2010; Xin, 2012). However, there are important differences across research groups on how interpretation occurs and the extent to which it is seen as requiring explicit support through instruction.

The extent to which interpretation requires explicit support through instruction is the key distinguishing factor between groups that are in general consensus regarding its importance in a problem-solving situation. On the one hand, there are those research groups which – while acknowledging interpretation as important – see it as a largely 'natural' process that learners engage in when presented with problem situations, and, therefore as not requiring of explicit support during instruction (Carpenter, Fennema, & Franke, 1996; Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Klein, Beishuizen, & Treffers, 1998; Treffers, 1991; Van den Heuvel-Panhuizen, 2003). On the other hand, there are research groups that have provided evidence pointing to widespread evidence of learners suspending the sense they need to make of situations described by the text of these problems (De Corte et al., 2000; Sepeng & Sigola, 2013; Verschaffel, Greer, & De Corte, 2000). This evidence is the basis of the growing view (held for instance by Peled and Balacheff (2011)) that interpretation is not natural, and of the advocacy for the need to pay explicit and deliberate attention to supporting learners' interpretations of additive word problems.

2.2.1.1 Research groups that see interpretation of word problem situations as a largely 'natural' process

Two groups that have extensively studied learners' experiences with solving additive relations word problems are the Cognitively Guided Instruction (CGI) group (Carpenter, Ansell, Franke, Fennema, & Weisbeck, 1993; Carpenter, Fennema, & Franke, 1996; Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Carpenter & Moser, 1984; Carpenter, Moser, & Bebout, 1988) and the Realistic Mathematics Education (RME) group (Gravemeijer, Bruin-Muurling, Kraemer, & van Stiphout, 2016; Gravemeijer & Doorman, 1999; Klein, Beishuizen, & Treffers, 1998; Treffers, 1991; Van den Heuvel-Panhuizen, 1998; Van den Heuvel-Panhuizen & Drijvers, 2014). These groups tend to value learners' intuitive ways of thinking, and look to harvest these intuitive ways of thinking as starting points of instruction on the solving of additive relations word problems. Their proponents argue for the learning-teaching of additive relations problem-solving to build on the intuitive strategies that emerge from learners' attempts at solving these problems, with a view to transforming these intuitive ideas into more sophisticated ways of thinking and working. Accordingly, they delineate a learning-teaching trajectory that takes learners from directly modelling problem situations (Carpenter et al., 1996) towards flexibility in the use of calculating strategies.

These groups agree that the progression from the use of strategies involving counting by ones to the use of more facile strategies that do not involve counting is a critical one. In this regard, the RME group has promoted strategies that harness the sequential structure of numbers. Since these sequence-based strategies can be modelled as jumps on a number line, the empty number line (ENL) has been promoted over the past few decades, as a model that can facilitate such a progression (Ellemor-Collins & Wright, 2009; Ellemor-Collins & Wright, 2007; Treffers & Moor, 1990).

At this point, what is important to highlight, is that the CGI and RME groups tend to suggest that the setting up of direct or emergent models in order to make sense of the problem situation is 'a basic process that comes relatively naturally to most primary grade learners' (Carpenter et al., 1993, p. 440). The pedagogic implication of this view is that instruction needs to provide real or realistic situations that are set in contexts familiar to learners, encourage learners' engagement with these situations and monitor and draw on the models that the learners construct.

Some research groups working with the perspective that interpretation cannot be presumed as occurring naturally have provided, in contrast to the groups mentioned above, evidence that realistic considerations were not sufficient for 'activating a situation analysis as a [problem-solving] habit' (Peled & Balacheff, 2011, p. 5). According to these researchers, it is problematic to view the analysis and modelling of the situation as 'an automatic act' (ibid, p.6).

2.2.1.2 Research groups that advocate for explicit and deliberate attention to be paid to supporting learners' interpretations of additive word problems

For the most part, it has been linguistic considerations that have provided the chief motivation behind research groups' turning the focus of their attention to how instruction might explicitly support learner construction of models of situations. As early as the 1960's, researchers have been aware that the presentation of word problems as a text which tells a story, requires certain linguistic skills in order to make sense of the story and of the question posed (Huttenlocher, Eisenberg, & Strauss, 1968). Since then, different research groups have been involved in extensive investigations into the influence of the semantic structure of word problems on learners' solution strategies. Four of the main groupings conducting parallel research into the influence of the semantic structures of word problems on learners' solution strategies are the *CGI group* (Carpenter et al., 1993; Carpenter et al., 1996; Carpenter et al., 1998; Carpenter & Moser, 1984; Carpenter et al., 1988), the Vergnaud group (Vergnaud, 1982, 2009; Vergnaud & Durand, 1976), the *Lewis & Mayer group* (Hegarty, Mayer, & Monk, 1995; Lewis, 1989; Lewis & Mayer, 1987; Mayer & Hegarty, 2012; Mayer, Lewis, & Hegarty, 1992), as well as the *Nesher group* (Greeno, 1980; Nesher, Greeno, & Riley, 1982; Riley, 1979, 1984). In the second half of the chapter, where I deal with the literature on the additive word problem types, I also attend to the contributions made by each of these groupings to the categorization of additive relations word problems. I introduce them here to foreground the fact that, whilst these groups recognise the importance of paying attention to interpreting relationships and choosing the correct operation, 'all emphasized semantic analysis of the problems by which students make distinctions among [the different] problem types and then map the problem into [their] respective schematic diagrams' (Xin et al., 2008, p. 4).

More recently, Peled and Balacheff (2011) have also observed that, because word problems are often presented to learners by means of a text which tells a story, in the initial treatment of the problem, linguistic considerations will inevitably come into play. In this regard, they posited that linguistic skills are always needed to make sense of the story and of the question posed. Peled and colleagues have subsequently proposed that whilst the language of some word problems may be sufficiently accessible for the model of the situation to be 'induced naturally by experience and/or cultural reference', there is a need for a broadening of the goals of the learning-teaching of problem-solving to go beyond simply teaching learners how to fit a mathematical model to a given situation (by mapping particular problem types into their respective schematic diagrams), into the development of learners' modelling skills.

These views have been echoed by Polotskaia and colleagues who have conducted parallel research into the reasons behind learners' struggles with constructing sustainable knowledge related to word problems (Polotskaia et al., 2015). With the introduction of word problems in the early grades in Canada in 2001, there arose an opportunity to do some work on supporting learners' interpretation of additive word problems. This presented Polotskaia & colleagues with an opportunity to look into how learners learn to analyse the situations in additive relations word problems, and, help provide insights into how best to balance the already strong calculational work seen in the Canadian context with some interpretational work.

In light of the observation made by Hegarty et al. (1995) that learners' success in solving a word problem is a function of the choice of reasoning they make when approaching the problem, Polotskaia (2017) proposes two types of reasoning; namely: *sequential numerical reasoning* and *holistic relational reasoning.* She observed that within the context of a curriculum that only promotes numerical sequential reasoning, the development in learners of a 'relational holistic vision' of a situation can be stifled. This meant that, ideally, the teaching of additive relations problem solving should be organised in such a way that learners are afforded the opportunity to develop sequential numerical reasoning and holistic relational reasoning 'in harmony and coherence with each other' (p.166).

In their investigations into learners' struggles with interpreting additive relations word problems, Polotskaia and colleagues came up with a *reasoning duality hypothesis* which forms the basis of their explanations regarding these challenges. They conjecture that the challenges that learners experience with solving word problems are rooted in traditional teaching approaches that do not do enough to foreground the additive relationship with a view to supporting learners to 'see a situation as a whole and [to] better coordinate relationships between the quantities involved' (Polotskaia et al., 2015, p. 257). With a view to helping learners develop their interpretational skills, Polotskaia and colleagues propose an approach that draws on the *Theoretical Thinking* approach of Davydov (1999). The fundamental principle behind this approach is that learners' initial encounters with problem solving should not be about finding a numerical solution, but rather, they should be about analysing and modelling the situation in terms of the relationships involved (Polotskaia, 2017, p. 170). They describe this as the Equilibrated Developmental Approach (EDA).

The pedagogic implications of Polotskaia's ideas are that the main goal of instruction on word problem-solving must be reformulated from an emphasis on finding the numerical answer to the word problem into an activity of analysing and modelling additive relationships. Consequently, they call for dedicated time to be allocated to getting learners into the habit of first interpreting the situation into what Peled and Balacheff (2011) have referred to as a *model of the situation.* This includes time for transforming the semantic meaning of the text into an additive relationship by constructing a situational model to represent all the important elements, known and unknown quantities, according to their constitutive roles in the additive relationship. They also propose that time be made available to learners for manipulating the situational model and transforming the relationship expressed in its configuration into a mathematical expression that reveals the necessary arithmetic operation. In the classroom discussion that would then follow, the teacher 'engages pupils in an explicit analysis of the mathematical structure of a problem as a system of quantitative relationships' (Polotskaia & Savard, 2018) , thus focusing learners' attention on the situation as a whole. This approach, they argue, supports learners seeing the relationship between knowns and the unknown quantities as a part-whole relationship.

2.2.2 Relational models advocated as useful for modelling additive situations in interpretation phase

Several studies have recommended different part-whole models as highly useful for supporting the analysis of the structural relations in word problems. The literature provides support for part-whole models as highly useful representations of the structural relations in a word problem, particularly the works of Davydov (1982) but also in more recent studies (Polotskaia et al., 2015; Xin et al., 2008). These part-whole models are promoted to compel learners 'to take into consideration the relationships between the quantities involved in order to understand the situation mathematically' (Savard & Polotskaia, 2017, p. 825).

Whilst engaging learners in an explicit analysis of the mathematical structure of a problem as a system of quantitative relationships, Polotskaia and colleagues recommend the use of line segment diagrams named Arrange-All (A-A) diagrams as models that can facilitate thinking about additive relationships. The chief motivation behind this choice is that A-A diagrams allow for the configuring of relationships without necessarily representing numerical values. This, they argue, is in line with Davydov’s (2008) conceptualization of the *Theoretical Thinking* approach in which he argues that the notion of an additive relationship – which is ‘the law of composition by which the relation between two elements determines a unique third element as a function’ (Davydov, 1982, p.229) – should develop from learners first playing with physical objects such as strings, and describing relationships between quantities without referring to numbers. Only later, when they are acquainted with the fundamental principle of composition in a general form (part + part = whole), are they introduced to abstract quantitative relations through solving additive problems with the use of A-A diagrams (Polotskaia, 2017).

2.2.2.1 The Arrange-All (A-A) diagram

Polotskaia (2017) offers the diagram in figure 1 to illustrate the use of the A-A diagram to model the additive relationship in the following apples and pears problem:

There are 25 apples and some pears in a basket. There are 7 more apples than pears. How many pears are there in the basket?

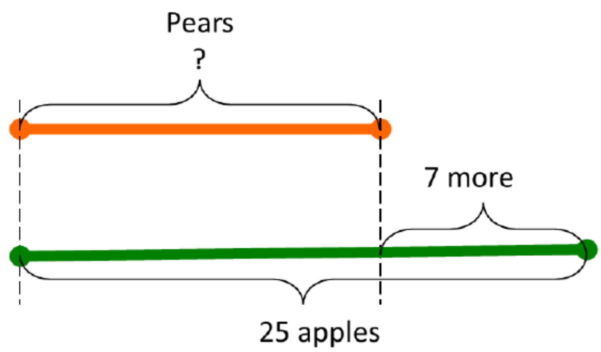


Figure 1: An A-A diagram modelling the additive relationship in the apples and pears problem (Polotskaia, 2017)

This interpretation of the apples and pears problem into the A-A diagram in figure 1 transforms the ‘more’ relationship in the apples and pears problem into a part-part-whole relationship that interprets the 25 apples as the whole and the ‘7 more apples’ as part of the 25 apples in the basket. In the interpretation phase, therefore, where ‘the central activities that students need to engage in is the unpacking of the meaning of the system’ (Lesh, Carmona, Hjalmarson, & Mason, 2006, p. 216), the A-A diagram is meant to facilitate the allocation of each of the three additive elements in their constitutional positions as part, part and whole.

Around the world, studies with an interest in explaining the consistently high performance of learners in the Singapore Mathematics Curriculum (SMC) on international benchmark tests have attributed the success to the use of models for supporting sense-making of word problems (Murata, 2008). In their Conceptual Model-based Problem Solving (COMPS) intervention program designed and implemented in the US, Xin et al. (2008) emphasize 'representing the word problem in a defined mathematical model [for] conceptual understanding of additive word problems before considering the *choice of operations*' (p. 12)[emphasis mine] and representing the additive relationship in an equation. In the same way that A-A diagrams are used in the Equilibrated Developmental Approach (EDA), bar models are used to 'bridge the conceptual gap between concrete modeling and abstract representation of mathematical models' (p.11) in the Conceptual Model-based Problem Solving (COMPS) intervention program. When covering additive word problem solving, the COMPS program recommends the use of (what they refer to) as *Part-Part-Whole (PPW) model equations* as abstract representations of mathematical models or a generalizable conceptual model.

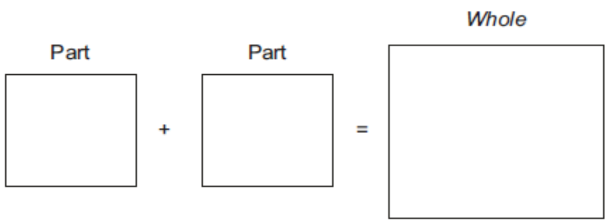


Figure 2: A Part-Part-Whole (PPW) model equation diagram (Xin et al., 2008)

2.2.2.2 The bar diagram (BD)

The model made popular and used extensively by the SMC is referred to around the world as the *bar diagram*. As outlined in the Singapore Curriculum Framework, the use of a 'diagram or model' is regarded as a useful heuristic.

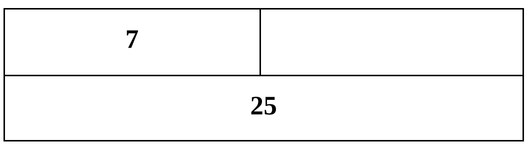


Figure 3: A bar diagram modelling the additive relationship in the apples and pears problem

The bar diagram is a rectangular bar partitioned spatially to represent the whole and the parts that make up the whole, where the combined length of the parts is equal to the length of the bar representing the whole. Its structural (spatial) distribution into two rectangles that together are as long as the third rectangle allows for the parts and the whole to be conceived of simultaneously. It is the static nature of the bar diagram that makes it the ideal artifact for handling problems whose language does not suggest an action but rather a static relationship between the component parts of a whole (Carpenter et al., 1999), as is the case in the apples and pears problems.

More importantly, what research into the use of BDs has shown is that, regardless of which part of the world the research was conducted and the name used to identify it - tape diagrams in Japan (Murata, 2008), strip diagrams in US (Beckmann, 2004), or bar models in Singapore (Hoven & Garelick, 2007) - the power of this relational model resides in its potential for representing 'all the information in word problems as a cohesive whole, rather than as distinct parts' (Cai, 2004, p. 8). In particular, Hoven and Garelick (2007) found that, in Singapore, where curriculum prescribes building complexity into the use the bar model from Grade 3 (where learners have to solve 'simple problems') through grades 4 & 5 (where they have to solve 'multi-step problems') to grade 6 (where they are required to solve 'hard problems'), provides them with a 'solid foundation' for transitioning into algebra in later grades. It is for this longevity in application that the BD was promoted during the intervention in this study.

2.2.2.3 Implications for my intervention design in terms of interpretation phase

In the South African context, there is evidence that learners struggle to set up models of word problem situations (Sepeng & Sigola, 2013; Webb & Sepeng, 2012). The major reason offered by these studies for learners' struggles with setting up models of word problem situations was that classroom practices 'put emphasis on computational proficiency at the expense of interpretive skills' (Sepeng & Sigola, 2013, p. 332). In Takane's (2021) study with Grade 2 learners, she notes that even when presented with familiar situations in a home language medium of instruction, learners did not 'naturally' move to creating emergent models. She notes too, limited incorporation of realistic situations/word problems in early grades' instruction. Given this, instruction in her intervention had to devote substantial time to engendering interpretation activity. Broader evidence points more generally to the prevalence of rote learning approaches, set within cultures of teacher-led instruction (Hoadley, 2012; Tabulawa, 2013).

In light of this evidence, I decided to follow an approach that would attend explicitly and deliberately, not only to learners' calculations of their solutions, but primarily to their interpretation of the problem situation into models. Such a modelling approach has been developed and used by Polotskaia and colleagues to teach additive word problem-solving. In section 2.4.7 I detail this modelling approach – the Equilibrated Developmental Approach (EDA) – along with the necessary adaptations incorporated into it in order to meet the aims of my study. Suffice it for now to highlight the fact that Polotskaia and colleagues designed the EDA following Relational Paradigm principles, and on the basis of their *reasoning duality hypothesis*, to integrate 'the simultaneous and equilibrated development of the relational and sequential ways of thinking'(Polotskaia & Savard, 2018, p. 76) .

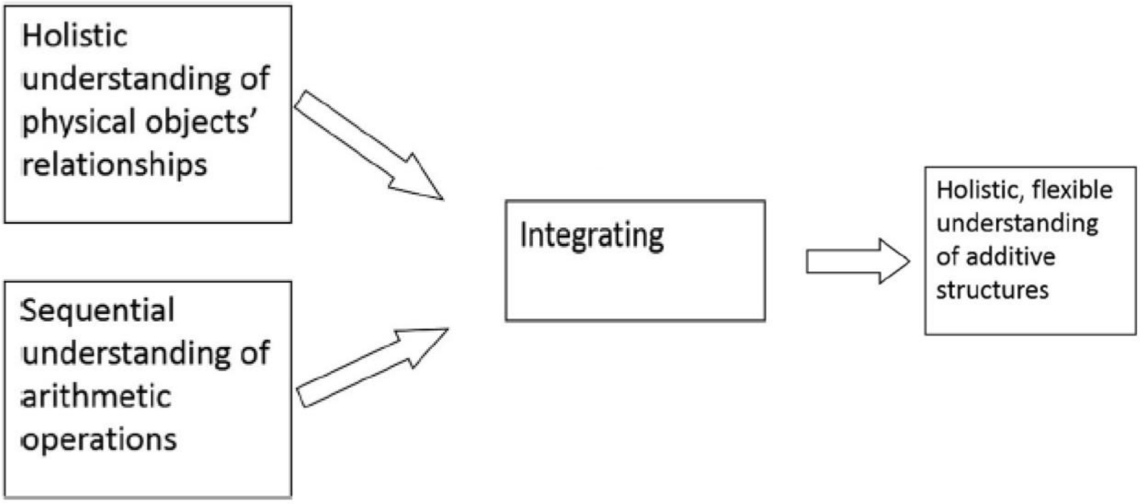


Figure 4: The development of thinking in the Equilibrated Development Approach (Polotskaia & Savard, 2018, p.)

Given that the short nature of the intervention did not allow for progressive explorations beginning with concrete and physical objects (strings) and culminating in abstract part-whole diagrams, an important adaptation to foreground here is the substitution of the word 'additive' for the phrase 'physical objects' in figure 4. Also, given that this study was conducted in the context of a South African curriculum that introduces learners to number through counting, and then swiftly proceeds to expose the learners to symbolic representations of number, it made sense to promote models that incorporated numbers in symbolic form. Realising that the learners who participated in this study were in Grade 3, and that working with abstract material quantities was very distant from Grade 3 learners' experiences in South Africa, I opted to have learners model the part-whole structure of the additive relationship on a BD (instead of an A-

A diagram)[2]. Important to highlight is that, whilst the A-A diagram is primarily used to model relationships between physical quantities by representing two parallel segments (often of different lengths) and in some other cases only one segment composed of two parts, the BD is mainly used to model relationship between numbers (and unknowns) representing two parts which together have the same length as the bar representing the whole.

Like the A-A diagram, however, the BD provides learners with a structural view of the additive relationship as a part-whole relation, with the added advantage of using the numerical values of the quantities involved as confirmation of the whole as bigger than either of the two parts. Whilst the numerals representing the known quantities can be annotated in the relevant slots in the BD, the unknown quantity can be represented with an empty slot in a way that relates it to the two known quantities as a function of the additive relationship. As a spatial representation of the additive relationship, the BD was meant to provide learners with a model of the part-whole structure of the additive relationship, thus supporting the structural reasoning required of the solver in the interpretation phase. The part-whole structure of the BD facilitates reasoning about the relations between the quantities in the problem as parts and whole. It, therefore, lends itself to supporting the interpretation of the structural relations in the interpretation phase of additive word problem-solving.

With this study conducted in the Relational Paradigm, priority was given to transforming the semantic meaning of the text into an additive relationship via explicit analysis of the mathematical structure of a problem as a means to understanding the problem situation holistically. In order to support this analysis, the construction of a model that represents the part-whole relationship between the knowns and the unknown in the problem situation was the central activity from which a mathematical expression in the form of a number sentence in standard mathematical form could be derived.

This deriving of the mathematical expression from a model of the additive relationship is the learning target of problem-solving activity implemented in the Relational Paradigm. Problems are regarded as 'successfully solved if the pupil provided the correct mathematical expression in the required standard form' (Polotskaia & Savard, 2018) .

[2] Rationales and justifications for adaptations to the Equilibrated Development Approach for this study are detailed in section 2.4.7.

This apparent relegation of the calculation to a secondary role stems from the consensus held within the Relational Paradigm that 'knowledge about relations described in the situation can be studied and learned independently from calculation methods' (Polotskaia & Savard, 2018, p. 80) . Although from a methodological point of view the emphasis of the Relational Paradigm approach is on the interpretation phase, it is agreed that during problem-solving, once the correct mathematical expression has been 'deduced from the identified relationship, [the calculation phase] can then be carried out using a variety of calculation tools' (p. 74).

2.3 Calculation phase

Alongside some experiences of interpretation of the quantitative relations in place, the attention of this intervention study also needed to focus on the calculation phase. Here, the emphasis in reviewing the relevant literature is on moving learners' beyond counting-based strategies to calculation that is based on the relational structure of number, and how this way of working with numbers could be leveraged for progress towards increasingly efficient calculation of answers. As already pointed out above, however, much less prominent in the Relational Paradigm is an emphasis on developing increasingly efficient and flexible calculation strategies, and in fact, more broadly on the calculation phase overall.

Given the South African context of problems both with interpretation and calculation, and hence the focus of this study on both issues, there was a need to structure this study such that attention to efficiency during calculation feeds into calculation phase working. A key literature base that emphasises working with the relational structure of number and how this structure can be leveraged for progress towards increasingly efficient calculation of answers is located within Realistic Mathematics Education (RME) writing. Not only did this literature base provide this study with a model for supporting learners' progress towards increasingly efficient calculation of answers, but by delineating hierarchies of increasing sophistication amongst calculation strategies, this literature base also provided this study with a further basis for considering the effectiveness of the intervention. Thus, I detail in this section, RME-based rationales for attention to sophistication, as well as key models for supporting calculation that is based on number relations around ten and multiples of ten.

2.3.1 RME attention to the calculation phase

Researchers conducting their studies within RME argue for a clear distinction to be made between two broad categories of procedures for solving addition and subtraction problems; namely: column-wise and non-column-wise processing of numbers (Beishuizen, 1993). Column-wise processing is commonly referred to as the column addition/subtraction method, and non-column-wise processing of numbers can be further elaborated into the two sub-categories of sequential and decomposition procedures known as the N10 (jump strategies) and the 1010 (split strategies) respectively. Jump strategies refer to sequence-based strategies in which one number is kept whole, and the other one is added or subtracted in a series of jumps, the accomplishment of which requires knowledge of jumping forwards and backwards in tens from any given number. Split strategies, in contrast, involve both operands being partitioned, and split, into tens and units (in the case of 2-digit numbers) which are then added or subtracted separately, the results of which are then added or subtracted (Beishuizen, 1993) depending on whether the task involved addition or subtraction.

It is worth noting that studies conducted in the Dutch and the US systems indicate that, the N10 strategy is the more effective of the two non-column-wise strategies, and that the 1010 strategy, which, like column-wise processing, was found to be prone to errors (Klein et al., 1998), especially in subtraction tasks with numbers in which the digits of the subtrahend are larger than those of the minuend (Kilpatrick & Swafford, 2002). The reason that the 1010 strategy and the column method are prone to error relates to the fact that they both involve treating and processing the tens and units separately, instead of working with both simultaneously.

2.3.2 Relational models advocated as useful for modelling additive situations in calculation phase

A key part of supporting numerical reasoning relates to the use of models that can support the necessary moves away from counting-based strategies towards deriving answers based on the relational structure of number. In this regard, the empty number line (ENL) was developed and has been successfully used for over three decades in the emergent model instructional design of RME based at the Freudenthal Institute for supporting the progression towards increasingly efficient calculation of answers.

Given the promotion of mental arithmetic in the lower grades as an extension of learners' informal strategies (Beishuizen, Van Putten, & Van Mulken, 1997) in the RME approach, researchers working with this approach have extensively studied the development of increasingly flexible and efficient approaches to additive calculations. These researchers (Beishuizen et al., 1997; Ellemor-Collins & Wright, 2009) emphasize developing learners' facility with adding and subtracting numbers between 1 and 20 as a basis for tackling numbers bigger than 20. To this end, Ellemor-Collins and Wright (2009) propose adding and subtracting that does not rely on counting as a critical goal of facile calculating.

Research conducted within RME identified persistent counting as the reason behind the failure on the part of a great many learners to make the qualitative change in number thinking that will allow them to achieve the critical goal of solving additive tasks without resorting to counting. To the end of facilitating this qualitative change in learners' thinking where operating on numbers is concerned, in the *number structuring framework* (Ellemor-Collins & Wright, 2009) developed within the RME tradition, the main activity is 'numerical restructuring' (Beishuizen et al., 1997) of the numbers involved according to the quantitative relations between them on an ENL.

2.3.3 The empty number line (ENL)

The empty number line is a graphic representation of the logical structure of numbers which relates closely to the counting sequence (Anghileri, 2006). From its structural make up, it provides a sequential view of numbers with values increasing from left to right, and thus connects movement to the right to addition, and movement to the left to subtraction. Connecting with arithmetic operation through the structure of the number line with jumps to the right/left across all additive tasks gets learners involved with their own actions, allowing them to track their processing during calculation. Continuing to work in this way facilitates reified counting (Gray & Tall, 1994). In view of the evidence of learners' predominant use of count-all based strategies in early addition problems in the broader WMC-P project (Venkat, 2011) and in South African evidence more generally (Schollar, 2008), and the role that the ENL can play in facilitating the qualitative change in learners' thinking pursued across all realistic approaches, the ENL was the model promoted in the calculation phase.

Given that, where the calculation phase is concerned, the goal of this study was to have the learners appropriately decomposing and recomposing the numbers involved instead of counting, particular non-counting strategies were introduced during intervention lessons for calculating with efficiency. For its potential to dynamically support different non-counting strategies (Treffers & Goffree, 1985), these calculation strategies (detailed later) are best modelled as jumps on an ENL, hence the choice of the ENL to mediate the calculation phase. The ENL was, therefore, a good relational model for supporting the calculation implied in the additive relationship expressed in the model of the situation constructed in the interpretation phase. In particular, its promotion was in line with the principle of associating arithmetic operations with sequential processing in the Relational Paradigm.

The rationale behind the promotion of the ENL, therefore, was based on the assumption that its availability in learners' tool-kit for solving additive relations problems would mediate the choices regarding best strategies for solving the problem in two ways: firstly, by serving as a model that can facilitate the disruption of the immediate impulse to count in ones given its potential for supporting non-counting strategies, and, secondly that, in fact, the strategies used are of the kind that require the type of reasoning that makes use of partitions and multiples of ten as reference points.

Recognizing that 'choice exists only where there are possibilities' (Levitin, 1982, p. 315), defining *strategy* as involving a choice to be made amongst different possibilities (Beishuizen, 1997) meant that it was imperative that the learners be introduced to models that could support strategies that would increase their repertoire of calculation strategies. Given the objective of moving learners away from using unit counting strategies, the strategies that were promoted had to be of the non-counting type, hence the promotion of the ENL.

Following the introduction of the ENL in Dutch schools in light of poor results in the National Evaluation Tests conducted in 1987, one significant outcome was a strong improvement on subtraction problems, with problems that require subtraction-with-carrying reaching an about 80% correct level amongst 2nd-graders (Year 3) when modeled on an ENL. Ever since the Evaluation Tests were conducted in 1987, studies conducted around the world have continuously confirmed the ENL as a highly useful model 'for presenting a visual structure which can inform number calculations' (Mason, 2018, p. 332). More recently, in post-assessments of 204 low-attaining 3rd and 4th graders conducted across the state of Victoria in Australia, Ellemor-Collins and Wright (2007) found that the sequential structure of the ENL and its use as a calculation tool can facilitate shifts away from counting in ones towards the use of the base-ten structure of number during calculation.

2.3.4 Calculation strategies

Given the implementation of this intervention in the Relational Paradigm and its requirement to work relationally, ideally, the strategies had to invoke the base ten structure of the number system. Of the two types of base-ten strategies identified by literature, *jump strategies* were promoted over *split strategies* (Ellemor-Collins & Wright, 2007) in this study. In particular, for longevity in application and following the observations made by international research regarding the effectiveness of the N10 over the 1010 strategy (Blöte et al., 2000), it was the N10 and its variations that was promoted during the intervention in this study.

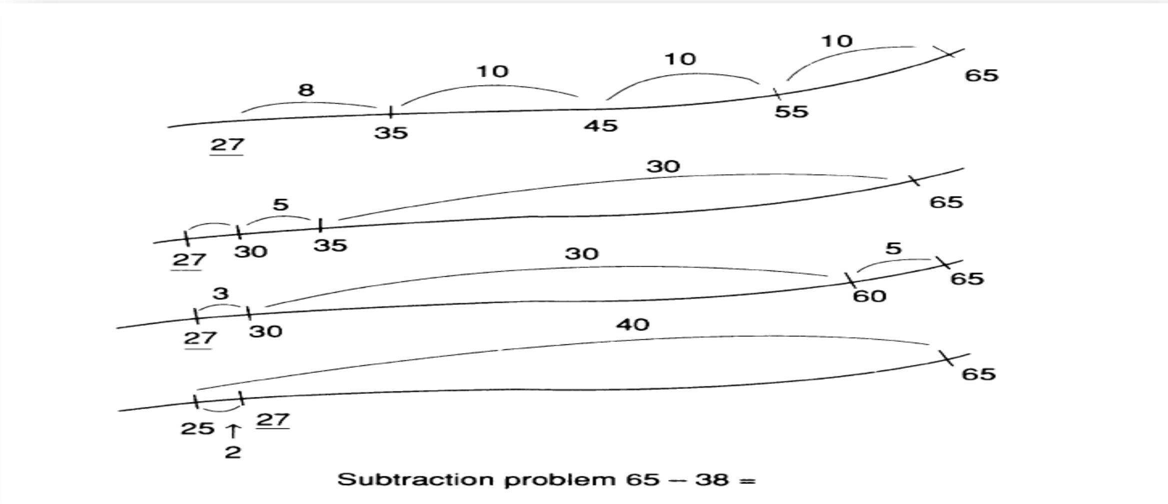


Figure 5: Various non-counting N10 strategies supported by the empty number line (Beishuizen, 1993)

Given that the N10 strategy and its variations are sequential procedures, they can be modelled on the linear character of the number line. In distinguishing between the different jump strategies promoted during the intervention as supported by the ENL, and the different ways in which each one exhibits a part-whole construction of number, figure 5 drawn from (Beishuizen, 1993) was instructive.

The strategy used in the top ENL and the one below it (in figure 5) is what is commonly referred to as the N10 strategy for subtraction. What sets the two ways of working apart is the compression of jumps of ten into a single jump of 30, and the use of multiples of ten as convenient landing positions ('adding through ten': A10) in the latter compared to the former. What sets the procedure in the second ENL from the procedure in the third ENL is that in the latter the A10 strategy is performed in the very first jump, whereas in the former the A10 strategy is performed in the second jump. As a result, the second and the third number lines reveal the same degree of structuring. The central engagement in the use of the A10 strategy is the partitioning of the units of the addend/subtrahend in such a way as to add/subtract exactly what is needed to land on a (partial sum/difference) of a multiple of ten. When 'adding through 10' one is often required to appropriately partition single digit numbers. Consequently, the A10 strategy (adding through ten) is regarded in this framework as a more advanced strategy from a number structuring point of view as compared to the N10.

The fourth ENL records the use of the N10C strategy. The N10C also involves using the partitions of 10 and its multiples, and thus facilitates further compression of the jumps to only two jumps (from a total of four jumps recorded in the top ENL using the N10 strategy, to three jumps with the use of the A10).

The differences between the four different ways of processing illustrated in figure 5 above as seen from the point of view of a part-whole construction of number are illustrated in the following table:

ENL 1 (N10 strategy: 38 as composed of three 10's and an 8)	ENL 2 (*A10* strategy: 38 as composed of 30, 5 and 3)	ENL 3 (*A10* strategy: 38 as composed of 5, 30 and 3)	ENL 4 (*N10C* strategy: 38 as 2 less than 40)
$65 - 38$	$65 - 38$	$65 - 38$	$65 - 38$
$= 65 - (10 + 10 + 10 + 8)$	$= 65 - (30 + 5 + 3)$	$= 65 - (5 + 30 + 3)$	$= 65 - (40 - 2)$
$= 65 - 10 - 10 - 10 - 8$	$= 65 - 30 - 5 - 3$	$= 65 - 5 - 30 - 3$	$= 65 - 40 + 2$
$= 55 - 10 - 10 - 8$	$= 35 - 5 - 3$	$= 60 - 30 - 3$	$= 25 + 2$
$= 45 - 10 - 8$	$= 30 - 3$	$= 30 - 3$	$= 27$
$= 35 - 8$	$= 27$	$= 27$	
$= 27$			

Table 1: The different ways in which the base-ten strategies illustrated in figure 5 demonstrate a part-whole construction of number

What is not illustrated in figure 5 and table 1 above is the Connecting Arc (CA) strategy (Beishuizen et al., 1997), which, like the N10C, equally compresses the calculation into two jumps, thus revealing a similar degree of structuring as the N10. Typically, the CA strategy is recorded as forward jumps that add up to the difference between the two subtrahends. Using the CA strategy for the same subtraction task (65 – 38) would involve incrementally adding-on to 38 until one lands on 65, with the sum of the jumps constituting the answer. The CA strategy was not explicitly taught during intervention lessons.

2.3.5 Implications for my intervention design in terms of calculation phase

The *Report on the Annual National Assessments* compiled by the Department of Basic Education (DBE, 2012) identified a gap between the prescribed curriculum and the performance of the majority of SA learners. The major reason offered by research for this gap relates to ongoing inefficient ways of working with number (Venkat, 2013). In particular, Schollar (2008) and Ensor et al. (2009) point to a prevalence of inefficient and error-prone tally-based counting strategies. What makes this a major problem is evidence of this reliance on tally-based counting strategies persisting well beyond learners' engagements with additive tasks in the Foundation Phase, into their engagements with multiplicative tasks in the Intermediate Phase (Venkat & Mathews, 2019).

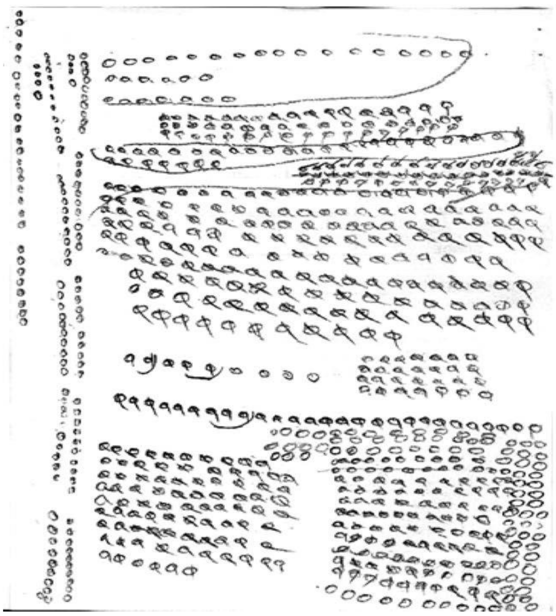


Figure 6: Examples of Intermediate Phase learner work on multiplicative tasks from Schollar (2008)

Given that this study was conducted in the context of a South African curriculum that introduces learners to number through counting, it was important to work toward disrupting the cycle of persistent counting and facilitating a qualitative change toward learners attending to the relations between numbers as a means to increasingly efficient and flexible calculation.

With this study conducted in the Relational Paradigm which, at least from a methodological point of view, prioritizes the analysis of the mathematical structure of an additive word problem as an indispensable feature of the interpretation phase, it was necessary to maintain the relational character of the engagement going into the calculation phase. On this account, I opted to work from a theoretical base that acknowledge, not only to promote a part-whole construction of a model that represents the additive relationship between the knowns and the unknown in the problem situation, but also with a part-whole construction of number as indicative of learners' understanding of number as composed by smaller numbers. In particular, for getting learners into the habit of attending to the relations between numbers and for longevity in application, the model of choice was the ENL. As a linear model, the ENL lends itself to supporting sequence-based strategies that involve keeping one number whole and conveniently decomposing and recomposing the other one for efficient calculation, thus supporting the numerical reasoning required of solvers in the calculation phase.

Whilst the foregoing discussions attended to the interpretation and calculation phases separately, in the design and implementation of the study attention was also on linking the two phases, and hence the two relational models. With this in mind, I look to detail the theoretical framing that I used to design the contours of the intervention for this study in the section that follows.

2.4. Combining attention to interpretation and calculation: RME-interwoven-Relational Paradigm

One concern raised by mathematics educators was that 'the current practice of word problems in school mathematics does not at all foster in students a genuine disposition towards… treating the text as a description of some real-world situation to be modeled mathematically' (De Corte et al., 2000, p. 1). In light of this concern, De Corte and colleagues proposed the mathematical modelling view of the teaching of word problem solving.

According to the mathematical modelling view of the teaching of word problem-solving, word problems can serve as an arena for the application of mathematics for solving word problems in the real world:

> According to the mathematical modeling view, students have to translate problem situations into mathematical expressions that can function as models (Gravemeijer, 2002, p. 2)

According to Gravemeijer (1997) there are two distinct ways of modelling within mathematics education: *modelling as organizing* and *modelling as translation*. Given that in the Relational Paradigm the central activity is interpreting the structure of the additive relationship in a problem into a graphical model, and then transforming this relationship into a mathematical expression involving an arithmetic operation to serve as the basis of a calculation sequence (Polotskaia & Savard, 2018), it can be argued that the inclination in the Relational Paradigm is towards *modelling as translation.* Whilst the interpretation phase of problem-solving involves organizing the quantities implicated in the text of the word problem into a part-whole relationship, moving within the interpretation phase from the BD to the mathematical expression incorporating an arithmetic operation involves translating the additive relationship in the BD model into a number sentence. In general, modelling as translation is an indispensable feature of the mathematical modellin*g* approach to additive relations word problem-solving. In this section, I argue that a systematic adoption of a cyclic view of modelling in the Relational Paradigm is, effectively, a tacit acknowledgement of the need for word problems to be conceived of as exercises in mathematical modelling (De Corte et al., 2000).

As explained in section 2.3, this study attended to both the interpretation and calculation phases, seen as important and necessary in South Africa. Also alluded to was my decision to follow an approach that would attend explicitly and deliberately, not only to learners' calculations of their solutions, but primarily to their interpretation of the problem situation into models. Given the greater emphasis placed in the Relational Paradigm on interpretations of problem situations into models, I adapted the Equilibrated Developmental Approach conceptualized by Polotskaia and colleagues. One major adaptation was that, given the need to also attend to learners' calculations of their solutions, I had to structure this study such that attention to efficiency during calculation feeds into calculation phase working. This necessitated locating the study in the Relational Paradigm, and drawing from the number structuring frameworks developed within RME for efficiency during calculation. In this section, both Polotskaia's overall model of working seen in what she has called the Equilibrated

Developmental Approach implemented within the Relational Paradigm, and the overall RME model of working are briefly outlined in order to highlight their overlaps and contrasts.

2.4.1 A view of modelling as progressively reorganizing situations in the RME tradition

Freudenthal (1973) lamented the use of models in a top-down instructional approach in which 'static models are derived from crystallized expert mathematical knowledge' (Gravemeijer & Stephan, 2002, p. 146) as anti-didactical. The main dissatisfaction associated with the use of these models, particularly for solving contextual problems, is that they leave little or no room for learners' informal strategies (Klein et al., 1998). This realization spawned a search for dynamically evolving models that could, not only foster learners' use of their informal strategies, but also allow them to enact more sophisticated strategies as their early number learning becomes progressively mathematized. Since then, from a pedagogical point of view, the approach was to treat modelling as organizing.

When modelling is treated as organizing, everyday-life reality becomes the starting point and the models are treated as tools that can mediate the appropriation of particular aspects of mathematical knowledge. In such an approach, the link to everyday-life situations is maintained through examples and tasks that are set in realistic contexts. The reasoning behind an emphasis on realistic contexts in the emergent modelling approach is that when learners learn mathematics severed from the real-world, they make little sense of it because the instruction disintegrates into teaching ready-made axiomatics (Freudenthal (1973), thus making it difficult for learners to apply in the real world in turn. By beginning in context and having the situation co-evolve with the model, learners are afforded the opportunity to participate in 'the activity of organizing matter from reality or mathematical matter' (Van den Heuvel-Panhuizen, 2003, p. 11) From this perspective, the process of constructing models is one of 'progressively reorganizing situations' (Gravemeijer, 2007).

2.4.2 Models and progression in the RME tradition

In the RME approach to learning and teaching, models are interpreted broadly as vehicles to elicit and support the progression from an informal understanding connected to the 'real' or imagined reality to the understanding of formal systems (Van den Heuvel-Panhuizen, 2003). Defined in this way, a model can assume the form of materials, visual sketches, paradigmatic situations (like repeated subtraction), schemes, diagrams and even symbols.

The term *emergent model* (Gravemeijer, 2002) was coined in recognition of the dynamic character of the process by which the model (of division as repeated subtraction, for instance) emerges in specific contexts, and later forms the basis for the emergence of more formal mathematical knowledge (like long division). In other words, the model and the meanings that learners are making of the model emerge in the process of structuring (organizing) a problem situation. In this conceptualization, 'the model and the sense made of the situation modelled are mutually constituted in the course of modelling activity' (Gravemeijer, 2002, p. 1). This mutual constitution of the model with the sense that learners are continuously making of the situation is particularly relevant in learners' early encounters with modelling word problems. In fact, the proponents of the emergent modelling approach acknowledge that emergent modelling can serve as a precursor to mathematical modelling (Gravemeijer, 2007).

Viewed from this modelling perspective, a repeated subtraction example highlights a shift in attention, on the part of the learner, from the original situation – where the model facilities sense-making through the connection that the learner maintains with the problem situation – to the mathematical relations involved. This is understood as a shift from repeated subtraction serving as a *model of* real-life division to curtailed repeated subtraction that serves as a *model for* mathematical reasoning about formal division (Gravemeijer, 1997). Important to notice is that there is little direction here on how to promote informal understanding – it is presumed that learners will do that naturally. Instead, the emphasis is on the informal *emergent model*, and how it can be progressed.

2.4.3 RME and mathematizing

According to Gravemeijer and Terwel (2000), Hans Freudenthal viewed Mathematics as 'a human activity', so he coined the word *mathematizing* in staying with the view of mathematics as something that humans do. Treffers (1987) elaborated on the two dimensions of mathematizing as:

> [Horizontal mathematization] involves converting a contextual problem into a mathematical problem, [vertical mathematization] involves taking mathematical matter onto a higher plane (Gravemeijer & Terwel, 2000, p. 782).

What makes it problematic to work with these constructs, however, is that, unlike the clearly defined interpretation and calculation phases used in the Relational Paradigm, the point of demarcation between the two dimensions of mathematizing, are by no means transparent (Gravemeijer & Terwel, 2000). Whilst Treffers (1987) posits *progressive mathematization* as a sequencing of these two types of mathematical activity, the notion has come to be understood as referring to the development of 'mathematical sophistication' over time (Rasmussen, Zandieh, King, & Teppo, 2005). Where progression in number is concerned, it is a shift that is in line with a learning-and-teaching trajectory that moves from calculation by counting, through calculation by structuring, to formal calculation and automatisation (Treffers, 2001). In the domain of additive tasks, an example of mathematical sophistication is when learners who initially relied on counting in ones begin to structure their counting around 10 as a reference point (Ellemor-Collins & Wright, 2009).

2.4.4 RME emphasis on efficient calculation

Working within the RME tradition which emphasises developing learners' facility with adding and subtracting numbers between 1 and 20 in order to reify their counting, Ellemor-Collins and Wright (2009) posit adding and subtracting that does not rely on counting as a critical goal of facile calculating. They identified persistent counting as the reason behind the failure on the part of a great many learners to make the qualitative change in number thinking that will allow them to achieve the critical goal of solving additive tasks without resorting to counting. To the end of facilitating this qualitative change in learners' thinking where numbers are concerned, they advocate for instruction that aims at having learners' facility in addition and subtraction constituted as an activity of *structuring number*. In this regard, Ellemor-Collins and Wright (2009) advocate for structuring numbers around 10 and its multiples as reference points, and offer the empty number line (ENL) as a key model for supporting calculating by structuring numbers around 10 and its multiples as reference points.

2.4.5 A view of problem-solving as modelling in the Relational Paradigm

In light of concerns over practices of word problem-solving that do not foster a genuine disposition towards appropriately interpreting problem situations, Verschaffel et al. (2010) proposed a paradigm shift towards presenting the solving of additive word problems as mathematical modelling exercises. This proposal has since been taken-up in the RP.

Instead of a view of the process of solving word problems as a linear process that starts with knowns and culminates in the calculation of the unknown (De Corte et al., 2000), proponents of the Relational Paradigm would like for the process of solving word problems to be thought of as a cyclic process involving both interpretation and calculation. More importantly, in the Relational Paradigm the two sub-processes of interpretation and calculation are seen as distinct, thus making it possible to speak of an interpretation phase that culminates in a mathematical expression incorporating an arithmetic operation and a calculation phase that produces the numerical solution to the problem. In their different conceptualizations of the problem-solving cycle, different researchers have proposed different steps or phases in line with the foci of their studies. After reviewing two key renditions, I detail such a cycle as I build-up towards presenting, in the section that follows, the model of problem-solving in line with the Equilibrated Developmental Approach adopted and adapted for this study.

2.4.6 The Relational Paradigm and translating problem situations into models

Viewing problem-solving as a cycle necessitates delineation of steps to be taken by the solver during problem-solving. This endows the problem-solving process with pedagogical hooks that allow for learners to focus, not only on the immediate outcomes of each stage/phase, but also with a 'modelling infrastructure' (Galbraith, 2012) or 'framework for modeling' (Xin, 2013) for problem-solving as a whole applicable to an entire host of similar problems.

Recorded in table 2 is one seminal rendition of a framework for modelling by Polya (1945), and another more recently offered by Blum and Leiss (2005).

Polya (1945)	Blum & Leiss (2005)
1) Understand the problem 2) create a plan 3) execute the plan, and 4) look back to check the results	1) Read and understand the task 2) structure the task and develop a real situational model 3) connect it to and/or represent it with a relevant mathematical model 4) solve and obtain the mathematical results, 5) interpret the math results in real problem context, and 6) Validate the results, and re-modify as the need arises.

Table 2: Two different renditions of a framework for modelling

A few things are worth highlighting regarding the two frameworks for modelling illustrated in table 2. Firstly, they both make explicit the step of the reading of the task/problem in order to emphasis sense-making; secondly, they both require of the solver to construct a model that will facilitate this sense-making; thirdly, they are similar in that they both adhere to the cyclic nature of problem-solving; and lastly, they differ according to how many steps are needed to complete the cycle. In the Equilibrated Development Approach, the shift towards explicit inclusion of the reading of the text of the problem constitutes an important adaptation introduced by to Savard's (2008) framework for modelling, which did not include the reading/interpretation part of their framework for modelling.

2.4.7 A framework for modelling in the Equilibrated Development Approach (EDA)

Drawing from the notion of problem solving as a cyclic process, Polotskaia and colleagues (Freiman et al., 2017; Polotskaia, 2017; Savard & Polotskaia, 2017) designed an approach to problem solving which they termed the Equilibrated Development Approach. The Equilibrated Developmental Approach, was developed based on the hypothesis that the ability to reflect simultaneously on the parts and on the whole is the basis upon which one can distinguish between an additive relationship on the one hand that can be represented by a range of number sentences, and selection of a specific number sentence for ease of calculation on the other. The Equilibrated Developmental Approach was, therefore, designed to reflect more closely the necessary balance that needs to be struck between structural and numerical reasoning for successful additive word problem-solving.

In order to support the implementation of the EDA in the classroom, Polotskaia and colleagues designed and developed the Ethno-Mathematical (EM) model of the problem-solving process following Relational Paradigm principles. As used in the Relational Paradigm, the EM model was initially conceptualised as a seven-stage cycle that included a *mathematical context* in which a model of the mathematical structure is constructed, a *graphical model context* in which the model is used to propose a calculation plan, and a *numerical context* in which the actual calculation is effected and the numerical result is produced (Freiman et al., 2017).

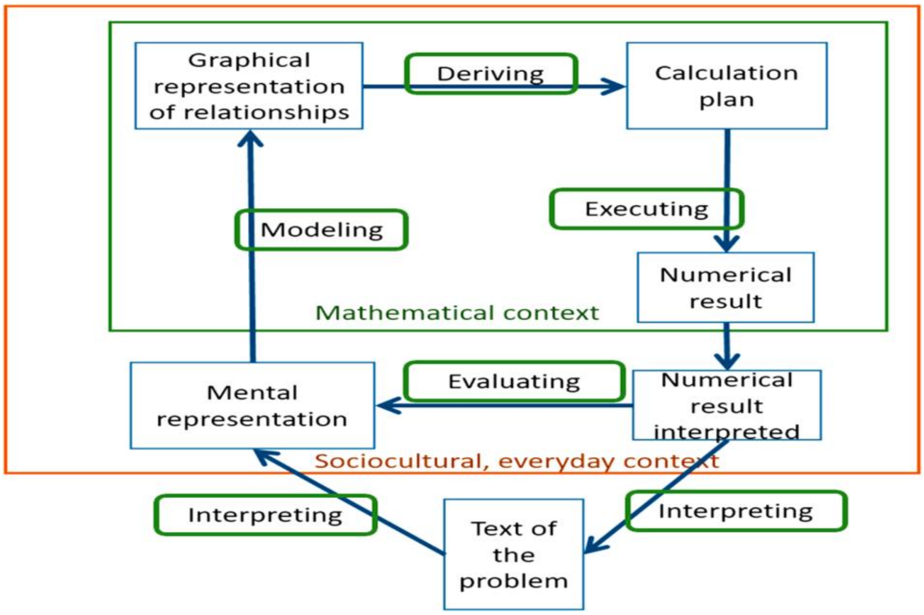


Figure 7 : Ethno-Mathematical model of problem-solving (Polotskaia & Savard, 2018)

Subsequently, the seven stages were compressed into four stages as captured by Polotskaia and Savard (2018). In this four-stage rendition, the EM model (illustrated in figure 7) provides a view of problem-solving activity as proceeding from the everyday socio-cultural context (represented by the outer orange rectangle), moves into the mathematical context (represented by the inner green rectangle) and back into the everyday socio-cultural context of the word problem from whence the process began.

The four steps that constitute the problem solving cycle as conceptualised in the EM model in its current compressed form are:

1. Understanding the situation within its everyday context
   (i.e. accessing the socio-cultural context through reading of the text of the problem),
2. Creating a model to express an understanding of the situation within the mathematical context , thereby producing a holistic view of the quantitative relationship involved,
3. Deriving the necessary arithmetical operation from this model and transforming the holistic view into a calculation sequence,
4. Caring out the calculation by using a convenient calculation strategy, then
5. Making sense of the calculation results in terms of the socio-cultural context, evaluating it in relation to the initial understanding of the problem.

By characterising problem solving as a cyclic process made up of clearly defined steps, the EM model distinguishes a little more clearly the boundaries between the steps to be taken when solving a word problem. This makes it more amenable to serving a didactical function, as a 'mental infrastructure or scaffolding for students learning the modelling process' (Galbraith, 2012, p. 8)

In the Relational Paradigm, flexibility is seen in terms of reasoning about additive structures and moving between mathematical expressions in what Polotskaia and Savard (2018) describe as standard and non-standard mathematical forms. In particular, Polotskaia and colleagues argue that 'the transformation of [non-standard] number sentence containing an unknown as addend or subtrahend ("$3 + ? = 8$") into the standard form ("$8 - 3 = ?$") requires holistic thinking' (Polotskaia & Savard, 2018, p. 80). In the Relational Paradigm, therefore, priority is given to interpreting the semantic meaning of the text into a model of the part-whole additive relationship as a means to understanding the problem situation in a structural way. In order to support this analysis, the construction of a model that represents the part-whole relationship between the knowns and the unknown becomes the central activity from which a mathematical expression in the form of a number sentence in standard mathematical form can be derived. The deriving of this number sentence from a part-whole model of the additive relationship constitutes transforming the additive relationship into a mathematical expression.

2.5 Implications for my intervention design in terms of the framework for modelling

Polotskaia and colleagues argue that the modelling of the additive relationship is the critical part of the solving of an additive word problem. As a result, the Relational Paradigm approach prioritises the interpretation of the mathematical structure over the calculation of the solution. On the other hand, RME advances the calculation of numerical solutions to additive word problems. Consequently, with a view to placing equal priority on matters of interpretation as well as on calculational matters, I designed this intervention study following Relational Paradigm principles, and interwoven with RME's focus on increasingly sophisticated calculation. This meant that whilst my framework for modelling made use of the phases of interpretation and calculation, it proceeded into RME focus on increasingly sophisticated use of calculation strategies. By doing so, the approach adopted in this study integrated interpretation and calculation as two critical phases that are required during additive word problem-solving.

2.5.1 The dual focus on interpretation and calculation and the reasons for adapting the EM model for purposes of this study

In the Equilibrated Development Approach (EDA) designed by Polotskaia and colleagues, the understanding of additive relationships and the modelling thereof is the primary engagement during additive word problem-solving. In their methodology, Polotskaia and colleagues transform the traditional word problem solving task into an activity of analysing and modelling an additive relationship. In the Relational Paradigm, within which the EDA was implemented, however, the problem-solving process was seen as a cyclic process, one that proceeds from the everyday socio-cultural context, moves into the mathematical context and back into the everyday socio-cultural context of the word problem from whence the process began (Polotskaia & Savard, 2018) . Whilst ideally the solving of a problem should have the general structure outlined by the Ethno-Mathematical Model, the actual solving process and the written response may not necessarily include nor explicitly reflect the five steps outlined by the EM model. For methodological and analytical reasons, therefore, the EM model had to be adapted to suit the dual focus of this study on interpretation and calculation.

In proceeding from the everyday socio-cultural context of the word problem, the implicit step "understand the problem" is transformed into an explicit activity of analysing and modelling of the additive relationship, in line with the EM model promoted for solving word problems in the Relational Paradigm. In other words, the sense that the solver makes of the problem situation is expressed in the model s/he constructs[3] of the structural relations in the problem situation. So, whilst the EM model of solving additive word problems in the Relational Paradigm identifies five stages that the solving process goes through, in the EDA, the implicit step 1 – *understanding the situation* – can be seen in the explicit step 2 – *creating a model to express an understanding of the situation* – effectively collapsing the two steps into one. The combination of *understanding the situation* and *creating a model to express an understanding of the situation* with *deriving the necessary number sentence from the model* then constitutes the interpretation phase of additive relations problem-solving for this study. In this way, the interpretation phase takes the problem from the everyday context, which is expressed in the word problem in ordinary everyday language, into the mathematical context.

[3] In the RME tradition, Gravemeijer (2002) would say that the sense that the solver makes of the situation and the construction of the model of the mathematical structure of the word problem are mutually constituted in the course of interpreting the problem situation.

The mathematical expression derived from the model of the additive relationship then serves as the basis of the calculation of the unknown element during the calculation phase which takes place entirely within the mathematical context. The requirement is for learners to generate a calculation method from the number sentence. As a result, the calculation phase is expressed in arithmetic language and modelled using mathematical objects. Viewing additive relations problem-solving in this way positions structural reasoning (about additive relationships) and numerical reasoning (on the relative sizes of the numbers for efficient addition and subtraction) as two complementary parts of additive relations problem-solving knowledge; which knowledge can only be developed in learners through the use of relational models in a dialectic approach implemented within the Relational Paradigm.

The view taken in this study, then, is that the problem solving process goes through three phases:

1. **Interpretation** of the problem situation, which involves making sense of the problem situation and creating a model to express the sense made of the problem situation, and deriving the necessary number sentence in standard form with the appropriate arithmetical operation from the model created;
2. **Calculation** of the unknown quantity, which proceeds by transforming the number sentence into a calculation method within the mathematical context aimed at producing the numerical solution, and
3. **Evaluation** of the numerical solution, which involves making sense of the calculation results in context, thus evaluating it in relation to the initial sense made of the problem situation.

In the interpretation phase, the solver's attention is on the constitutive roles of the knowns and the unknown in the problem situation. In the interpretation phase, the values of the numbers play a secondary role in that they only serve as 'place holders' of the parts and the whole that constitute the additive relationship; hence, the structural reasoning required in the interpretation phase. In the calculation phase, the solver's attention is on the number relations between the known quantities in the additive relationship, hence the numerical reasoning required in the calculation phase. In moving from the interpretation phase into the calculation phase, the transformation of the number sentence into a calculation sequence amounts to a shift from structural to numerical reasoning on the part of the solver. For this reason, success in solving additive relations word problems is viewed in this study as a function of the extent to which the solver can interpret the additive relationship involved, and calculate the sum or difference implied by the additive relationship. The prioritizing of the interpretation and the calculation phases over the evaluation phase deserves an explanation.

2.5.2 Motivation for truncating the evaluation phase of the numerical solution into simply identifying the numerical solution as the answer to the problem

Given that this study accorded equal value to the interpretation and calculation phases in light of the problems identified in SA, the evaluation phase did not form a substantial part of intervention. Instead, the answer produced through interpretation and calculation phases formed the more limited focus. Whilst the evaluation of the numerical solution is an important phase of problem-solving for purposes of making sense of the problem, the reasoning developed and the numerical answer, in this study the evaluation phase was truncated, both in my instructional methodology as the teacher of the intervention, and in my role as the researcher analysing the results, to simply identifying the numerical solution as the answer to the question. So instead of profiling problem-solving as an *interpretation-calculation-evaluation* cycle, I worked in this study with an *interpretation-calculation-answer* cycle, illustrated in figure 8.

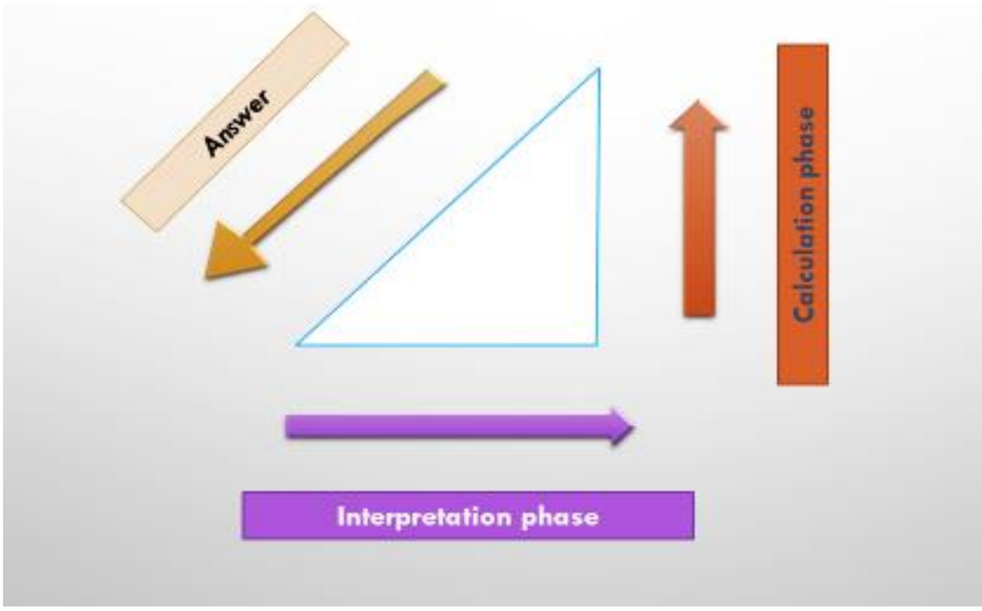


Figure 8: The interpretation-calculation-answer problem-solving cycle for this study

Profiling the problem-solving cycle in this way was in keeping with the equal priorities placed in this study on supporting learners to correctly interpret the structure of an additive relationship (by modelling the structural relations) AND efficiently calculate the numerical solution to an additive relations word problem (by modelling the number relations).

Indicators of success with interpretation are drawn from Polotskaia and other work dealing with setting up an appropriate model of a situation, and indicators of success with calculation are drawn from RME's trajectories of flexibility and efficiency in the use of calculation strategies. During the intervention lessons, I promoted the BD as a relational model that is ideal for supporting the interpretation of the problem situation and the ENL as a relational model that is ideal for supporting the calculation of the solution to the problem. However, given that the intervention lessons were interspersed with 'normal' teaching, all models presented by learners in their written solutions were considered in the analysis, recorded and coded for appropriateness. The indicators of success in evaluation were correct answers.

2.6 Semantic classification of additive word problems

Traditionally, additive word problems are defined by the arithmetical operation implied in their semantic structure. As a result, differences in difficulty were accounted for in terms of the semantic structures and the operations required to derive the meaning structures from the problem texts (Nesher et al., 1982).

Researchers agree on three additive relation situations that give rise to three classes of *semantic structures*; namely: *Change*, *Combine* and *Compare* (Nesher et al., 1982) . Prior research has used different nomenclature for these headings. In particular, Carpenter and colleagues in the CGI group identified *change increase* and *change decrease* as *Join* and *Separate* respectively. Nevertheless, research is in broad agreement on these classes.

In table 3, I present the three semantic categories according to how they are identified in three major camps: the CGI, Vergnaud and Nesher groups. Whilst the titles listed in the last column of table 3 are the titles previously used by the Nesher group, the titles that I will be using to identify the categories are the ones used by Nesher et al. (1982), and are captured in the first column.

Name of category used in this study	**Characteristics of each category according to Nesher et al.(1982)**	**CGI group**	**Vergnaud group**	**Nesher group**
Change	Describes increase or decrease in some initial state to produce a final state	**JOINING AND SEPARATING**: Carpenter, Moser, and Hiebert (1981)	**TRANSFORMATION LINKING TWO MEASURES:** Vergnaud & Durant (1976); Vergnaud (1981)	**CHANGE:** Greeno (1980) **DYNAMIC:** Nesher and Katriel (1978); Nesher et al. (1982)
Combine	Involves static relationship between sets. Asking about the Union set or about one of the two disjoint subsets	**PART-PART-WHOLE:** Carpenter et al. (1981)	**COMPOSITION OF TWO MEASURES**: Vergnaud & Durant (1976); Vergnaud (1981)	**COMBINE:** Greeno (1980); Heller & Greeno (1978); Riley, Greeno & Heller (1981). **STATIC:** Nesher et al. (1982)
Compare	Involves static comparison between two sets. Asking about the difference set or about one of the sets where the difference set is given	**COMPARE:** Carpenter et al. (1981)	**A STATIC RELATIONSHIP LINKING TWO MEASURES:** Vergnaud & Durant (1976); Vergnaud (1981)	**COMPARE:** Greeno (1980); Nesher & Katriel (1978); Nesher et al. (1982)

Table 3: The different nomenclature used to identify the three semantic categories of additive word problems

2.6.1 A hierarchy of difficulty identified in research studies

Traditional research addressing the different levels of difficulty associated with additive word problems in line with the abovementioned classification necessitates, that the classification according to semantic structure speaks to 'addition' separately from 'subtraction word problems' (Riley, 1979). In particular, research has found that, in order of difficulty, Compare problems are the hardest, followed by Combine problem and Change problems (Nesher et al., 1982). In other words, provided that the implied operation is an addition and the number range is similar, the *change* class of additive word problems has been found to be the least challenging for young learners. The implication here is that the same is not necessarily true for *subtraction word problems*. Traditional research is also in general consensus regarding the relative difficulty of Compare problems when measured against that of combine and Change problems (Carpenter et al., 1999; Nesher et al., 1982; Vergnaud, 1982; Verschaffel et al., 2000).

These studies also show that problems within the same category also vary in difficulty depending on what the CGI and the Nesher group refer to as the 'position' of the unknown in the associated number sentence and the Vergnaud group refer to as the 'nature' of the unknown according to its role in the semantic of the problem. Whilst learners' early encounters are often with additive word problems of the type where the final set is the unknown, the unknown can also be the initial set or the change set. The consensus view is that Change problems with the initial quantity unknown continue to prove more difficult than problems with the final quantity or the change quantity unknown (Giroux & Ste-Marie, 2001). In particular, Change problems in which there is a decrease in the initial quantity and the initial quantity is the unknown have been found to be particularly challenging for young problem-solvers relative to any other type of Change problems.

Whilst this difficulty has been observed for Change problems, more recent research, upon which this study is based, shows that the difficulty identified by traditional research, is not a direct function of the class, but depends more on the correspondence (or not) of the wording to the necessary operation to use. This difficulty has been found to be more pronounced where Compare problems are concerned given the static nature of the relationship expressed in them. As a result, it is not surprising that Compare problems have been the focus of many more researcher investigations than Change and Combine problems (Hegarty, Mayer, & Green, 1992; Hegarty et al., 1995; Pape, 2003; Verschaffel, 1994; Verschaffel, De Corte, & Pauwels, 1992).

The consensus view then is that the three semantic categories of additive word problems can be further distinguished into 14 different types. The 14 types are made up of six change, two combine and another six Compare problems. The six Change problems fall into the two subcategories of 'change increase' and 'change decrease' depending on the type of change effected in the statement of the problem.

2.7 Further challenges with correct interpretation of the problem situation

In their research, Nesher et al. (1982) distinguished between *ordinary language* and *arithmetic language*, and the roles that each one plays during problem-solving. They hypothesized levels of development towards the ability to solve word problems in order to explain the differences in performance on the different kinds of additive relations word problems by learners at different levels of development. They analysed semantic structures for their complexity and how this complexity can militate against the correct identification of the operation required (Nesher et al., 1982). Their work draws from Piaget's postulation that the ability to solve additive relations word problems involves two distinct structures of knowledge: the semantic knowledge of ordinary language and the semantic knowledge of arithmetic language. They concur with Piaget (1965) that, in order for the child to make the correct choice of arithmetic operation, he/she must first make sense of the ordinary language as formulated in a text of a word-problem. In particular, they posit that a child's interpretation of the situation (i.e. the sense that the child makes of the story formulated in ordinary language) depends on his/her understanding of the semantics expressed in ordinary language in the text of the word problem, which understanding enables the child to choose the correct arithmetic language operation, + or –.

Nesher and colleagues posit that there is an inherent difficulty that learners experience with solving word problems that establish a static relationship or comparison between quantities. They proposed that, in order to provide more nuanced explanations for different levels of performance in additive word problems, there is a need to further breakdown the requisite arithmetic knowledge into knowledge of *logical* and *mathematical operations*. This differentiation, they argue, allows for a more nuanced explanation of the reasons behind the challenges experienced by learners with solving word problems that establish a static relationship or comparison between quantities, as compared to problems that implicate an action. According to them, the difficulties that learners experience with problems implicating a static relationship or comparison between quantities can be explained in terms of a part-whole schema that may not be sufficiently developed to allow seeing the parts and on the whole at the same time. From a logical operations point of view, this means that the learner has not yet appropriated the additive relationship among the three quantities, and mathematically it means that the learner is not yet able to express this additive relationship in a model.

Nesher and colleagues hypothesize that, relative to change and Combine problems, solving Compare problems requires more advanced understanding of numbers as composed of other numbers, rather than as outcomes of counting processes. Consequently, Nesher et al. (1982) propose that, to successfully solve Change problems, attention must be paid to the direction of the change, as well as to the nature of the change.

Other researchers have identified other further elements that contribute to the difficulty of additive word problems.

2.7.1 Inconsistency in the language of the text

The research begun by Lewis and Mayer (1987) into learners' comprehension of word problems makes a distinction between what they refer to as *consistent language* (CL) and *inconsistent language* (IL) problems as its basis. In particular, they found that some word problems have inconsistencies between the terms that denote the action or the relation, and the arithmetic operation required to produce the solution to the problem, and that learners were more likely to misinterpret the IL problems than CL problems. For example, the correct interpretation of the word problem '*Joe has 8 marbles. Tom has 5 marbles. How many more marbles does Joe have than Tom?'* is that in the number of marbles that Joe has, the number of marbles that Tom has is included, and so we must subtract the latter from the former.

This observation led Lewis and Mayer (1987) to develop what they coined *the consistency hypothesis* according to which, CL problems are identified as problems in which there is consistency between the relational term (*more/less than*) in the statement of the problem and the arithmetic operation necessary to successfully solve the problem. In particular, the two instances of CL problems relate either to the case where the relational term *more* is used in the statement of the problem and the necessary operation is addition, and; conversely, where the relational term *less* is used in the statement of the problem and the proper operation is subtraction.

IL problems are identified as problems in which there is inconsistency between the relational term in the statement of the problem and the arithmetic operation necessary to successfully solve the problem. In IL problems the relational term (e.g., *more than)* conflicts with the necessary arithmetic operation (e.g., subtraction), and vice versa. Nesher and colleagues found that when learners encounter these IL problems, they have a tendency to transform the problem into a consistent language (CL) with which they are familiar as part of their solution strategy. Sometimes the transformation is into a CL problem with an equivalent relational structure, in which case the transformation maintains the relational structure of the original problem. For example '*John has 8 apples. He has 2 apples more than Mark. How many apples does Mark have?'* can be transformed into '*John has 8 apples. Mark has 2 apples less. How many apples does Mark have?'* and the relational structure would be maintained. If the learner understands: '*John has 8 apples. Mark has 2 more. How many apples does Mark have?',* and substitutes the relational structure of this problem for the relational structure of the two former problems, then, in the Relational Paradigm, his phenomenon is referred to as *structure substitution,* and it is understood as a special case of a *misinterpretation* in which the mathematical structure of a word problem is completely replaced with that of another (Polotskaia, Savard, & Freiman, 2016), resulting in a different problem being solved.

Over the past three decades, the implications of Lewis and Mayer's (1987) consistency hypothesis have been examined among learners of different age groups by different research groups (Hegarty et al., 1992; Hegarty et al., 1995; Pape, 2003; Verschaffel, 1994; Verschaffel et al., 1992). In reviewing this literature, it became apparent that the consensus view is that IL problems are more difficult for learners than CL problems. Whilst this is true for Change problems where there may be incongruences between direction of the change and the arithmetic operation (Giroux & Ste-Marie, 2001), it has been found to be more pronounced in the case of Compare problems where there is inconsistency between the relational term and the arithmetic operation. By arriving at the same conclusion as Lewis and Mayer (1987), these studies did not only confirm, but also elaborated on the already established hierarchy of difficulty between different additive relations word problems: change-increase problems in which the result is the unknown are the easiest, and Compare problems in which the referent set is the unknown are the most difficult. Table 4 summarizes the hierarchy of difficulty and links this hierarchy to examples provided by different researcher.

Levels according to Nesher et al. (1982)	Additive word problem-solving ability that relates to the identified level according to Nesher et al. (1982)	Problem category and nature of unknown according to Nesher et al. (1982) and Giroux & Ste-Marie (2001)	Example of Change problem	Example of Combine problem	Example of Compare problem
1	Reference to sets, adding members to and removing members from sets. Understanding 'put' 'give 'take', etc. as denoting change in location or possession.	Change Increase/Decrease Result Unknown; Combine Result Unknown	*John has 6 marbles. He lost 2 of them. How many marbles does John have now?*	*There are 3 boys and 4 girls. How many children are there altogether?*	(None)
2	Ability to link events by cause and effect. Reference to the amount of change. Understand sequence of events ordered in time in a non-reversible manner.	Change Increase/Decrease Change Unknown;	*John had 6 marbles. He lost some of them. Now John has 4 marbles. How many marbles has John lost?*	(None)	(None)
3	A reversible part-part-whole schema is available, and can be used to find unknown part in any slot in a sequence of events. Understanding class-inclusion.	Change Increase/Decrease Start Unknown; Combine Part Unknown; Compare problems mentioning 'more'/'less' and the difference / compared set is the unknown.	*John had some marbles. He lost 2 of them. Now John has 4 marbles. How many marbles did John have at the beginning?*	*There are 7 children at the park. 3 are boys. How many are girls?*	*Joe has 8 marbles. Tom has 5 marbles. How many more marbles does Joe have than Tom?*
4	Reversibility of non-symmetrical relations. Ability to handle directional descriptions (more/less), and quantify a relational set (relative comparison).	Compare problems mentioning 'more'/'less' and the referent set is the unknown.	(None)	(None)	*Joe has 8 marbles. He has 5 more than Dan. How many marbles does Dan have?*

Table 4: The levels of difficulty associated with each additive word problem type (semantic category and nature of unknown) according to Nesher et al. (1982)

Important to note is that what Giroux and Ste-Marie (2001) refer to as problems with a incongruences between direction of the change/relational term and the arithmetic operation are exactly the types of problems that Nesher et al. (1982) refer to as requiring a reversible additive structure in order to solve successfully. In this regard, Nesher et al. (1982) demonstrated that a reversible part-whole schema affords the learner the ability to read the relational term 'more than' in ordinary language, and yet perform a subtraction operation in arithmetic language. Conversely, the learner is able to read the relational term 'less than' in ordinary language, and still perform the requisite addition operation in arithmetic language. At that point in their development, the learner can reflect simultaneously on the parts and on the whole in a flexible way.

2.8 The Relational Paradigm and the disruption of the hierarchy of difficulty

In an attempt to account for the tendency for learners to choose the opposite operation, the view in the Relational Paradigm is that the challenges that learners experience when solving additive word problems relate either to a failure to discern the additive relationship or a failure to calculate the solution (Polotskaia & Savard, 2018) . On this premise, Polotskaia and colleagues assert that 'if both sequential and systemic thinking are developed simultaneously, pupils' reasoning will be better coordinated, and they will have similar success with all types of problems (Polotskaia & Savard, 2018, p. 76). This bold assertion provided this study with a secondary interest: the extent to which the hierarchy of difficulty across different types of problems was disrupted or altered by learners' working relationally in this study as hypothesized by Polotskaia and colleagues.

The position taken in the RP concerning the classification of additive word problems as consistent and inconsistent language problems is simply to 'not speak about the inversion of an operation, but about deriving an operation from the additive relationship' (Polotskaia, 2015). The reason for this is that when the position taken is that 'all additive word problems should be equally difficult if solved in two steps: transforming the semantic meaning of the text into an additive relationship, which is followed by transforming the relationship into the operation' (Polotskaia, 2017, p. 16), "operation inversion" looks different. Instead, the inversion, should it happen, would be an inversion of the additive relationship, and thus located in the interpretation phase, before an arithmetic operation could be derived. In such a case, it would be more appropriate to speak of a misinterpretation of the additive structure in the problem situation which resulted in an inappropriate configuring of the additive relationship as a system of three elements.

2.9 Concluding remarks

In reviewing the large body of research that has studied learners' problem-solving activity, two lines of investigation were distinguished. One line of investigation relates to learners' interpretations of additive relations in word problems. Here, the growing evidence associates successful solving of a word problem with the ability to analyse its mathematical structure and to model this structure in a part-whole diagram (Hegarty et al., 1995; Polotskaia, 2017; Xin, 2013). The other line of investigation reviewed observed and analysed the kinds of strategies that learners use to solve word problems (Beishuizen, 1993; Ellemor-Collins & Wright, 2009; Thompson, 1999). Here, scholars argue for the importance of choosing an efficient calculation strategy. Given the dual focus of this study on making sense of the mathematical structure of the problem and calculating the numerical solution, both lines of investigation were reviewed.

In an attempt to identify an approach to problem solving that warrants the value of attending to both foci, literature offered the Relational Paradigm. This is a mathematical modelling approach which recognises the importance of first interpreting the relational structure of the word problem into a model and deriving from this model a mathematical expression incorporating an arithmetic operation before calculating the numerical solution to the problem. In a South African context where challenges with making sense (interpretation) of additive situations and moving beyond counting based calculations loom large, it made sense to have the Relational Paradigm interwoven with the RME focus on increasingly sophisticated calculation.

Literature also revealed a growing advocacy for a modelling approach to problem solving. In this regard, literature reveals two types of modelling: modelling as translation and modelling as organizing. In reviewing these, I argued that a systematic adoption of a cyclic view of modelling in the Relational Paradigm is, effectively, a tacit acknowledgement of the proposal made by Verschaffel et al.'s (2010) proposal for word problems to be understood and used as mathematical modelling exercises. On these grounds, my framework for modelling catered for the phases of interpretation and calculation, with RME focus on increasingly sophisticated use of calculation strategies. By doing so, the approach adopted in this study integrated interpretation and calculation as two critical phases that require during additive word problem-solving. In this regard, the BD and ENL models were introduced and promoted during the intervention with the deliberate intention that they should mediate the interpretation and the calculation phases of the solving of additive word problem respectively.

In the following chapter, I present the relevant methodology and analytical frameworks for the issues raised in this chapter, appropriate for an intervention that sought to develop learners' structural and numerical reasoning in accordance with the Relational Paradigm.

CHAPTER 3

METHODOLOGY & ANALYTICAL FRAMEWORKS

3.1 Objectives of the research study

This study was conducted to take an in-depth look into learners' approaches to solving additive relations word problems, and in changes in these approaches following an intervention designed to support moves to more structured problem solving of additive word problems. The aim, in particular, was to determine the contributions to additive relations problem solving competence that can be effected through an intervention that focuses on the use of the bar diagram (BD) and the empty number line (ENL) as relational models

Whilst the main motivation for embarking on this study was to determine the extent to which the use of the BD and the ENL can support the interpretation of additive word problem situations and the calculation of their solutions by structuring the quantitative relations and the numbers involved respectively, the ancillary motivation was the lack of South African data on learners' performance on early number approaches as noted by Ensor et al. (2009). The research gap that insights from this study would serve to inform relates to the nature of, and possible shifts in, South African Foundation Phase learners' interpretations of additive word problem situations and calculations of solutions to these problems following an intervention that promoted the BD for supporting the structural reasoning required for interpreting problem situations, and the ENL for supporting the numerical reasoning required for calculating the solutions.

In the sections that follow, I provide a rationale for the outline model of the intervention lessons, the contours of instructional attention, as well as the frameworks used for analysing learners' responses to additive relations word problems.

3.2 Context

The Wits Maths Connect-Primary (WMC-P) project was launched in 2011 with the idea of working with ten government primary schools in one district with the overarching objective of improving the *number sense* of learners' in Grades R to 6. In order to establish Grade 2 learners' proficiency on number problems, a baseline assessment was conducted through administration of tests drawn from the work of Wright et al. (2006) structured on the basis of their Learning Framework in Number (LFiN). An overview analysis indicated that almost three quarters of the 238 learners assessed used count-all based strategies in early addition problems (Venkat et al., 2021). This strategy is often referred to as 'triple concrete counting' (Anghileri, 2006) to highlight the fact that a learner first counts out each addend and then counts the total – an inefficient strategy to use. In short, across the ten schools that initially partnered with the WMC-P project, counting approaches like 'triple concrete counting' dominated learners' working.

Investigations conducted into learning under the auspices of the WMC-P project reported on treatments of tasks by learners in which count-all procedures were the default approach to solving additive tasks (Venkat & Naidoo, 2012) and multiplicative tasks (Venkat et al., 2021). In both investigations, it was found that, invariably, these unit-counting approaches by learners occurred against a backdrop of teaching that paid limited attention to supporting learners' awareness of part-whole additive relationships and/or calculating that makes use of ten-based relationships between numbers, as well as a general rarity in instruction of connected sequences of examples (Askew, Venkat, Abdulhamid, Mathews, Morrison, Ramdhany & Tshesane, 2019).

Given these factors, the overall attention of the WMC-P project – over a nine year period – has been increasingly shifting towards balancing an emphasis on number relations for efficient calculation with 'more overt attention to creating and discussing informal models of situations that allow for setting up and comparing models, and the approaches linked to these models' (Venkat et al., 2021, p. 107). The outcomes bear testimony to how, in a South African context of low performance and low resources, an approach based on 'key models and connections' for improving sense-making can improve results substantially. It was in this evolving environment that the idea of this intervention study was seeded.

More broadly, research in South Africa and sub-Saharan Africa points to evidence of largely traditional teaching formats, in which learners tend to wait for instruction rather than exploring problems independently (Hoadley, 2007; Tabulawa, 1997). As noted earlier, this cultural norm has implications for the extent to which sense-making and emergent modeling can be seen as 'natural' activities that learners will engage with (Takane, 2021). This point was important within the instructional model selected for the intervention lessons, detailed later in this chapter.

While interventions focused on additive problem-solving have been trialed within the WMC-P project, more attention had been paid to ‘bald’ number additive problems than to word problems. Thus, the foregrounding of relational models as tools, firstly, for making sense of additive relationships in problem situations, and secondly, for arithmetizing additive relationships once the additive relationship has been appropriately interpreted had not been tried in combination. With this research gap in mind, I was interested in the patterns of gain that could be associated with an intervention designed to work from a theoretical base that values a modelling approach to solving additive relations word problems. To this end, this study implemented a short-term intervention programme in which, based on principles advanced in the Relational Paradigm, the two relational models were introduced.

Whilst the learners in Polotskaia’s (2015) study also participated as a whole class in a specially designed teaching experiment conducted by their regular teacher, they used individual task-based semi-structured interviews with their participants to collect data on learners’ reasoning. My interest was in discerning learners’ reasoning from their written responses. The motivation for this position stemmed from the norm of large classes in South Africa, which makes it almost impossible to equitably attend to learners individually. Consequently, I was interested in an intervention study, the outcomes of which could provide insights into the patterns of interpretation and calculation on a group of learners, as well as their reasoning, purely from learners’ written work.

3.3 Epistemological justification

On a research methodology continuum that has positivism and constructivism on the extreme poles, the analysis of a research study falls into one of two broad categories, or some semblance of the two: the qualitative, the quantitative or mixed methods respectively. With the focus of my research not only on the improvements (if any) in learners’ performance on the different additive word problems, but more importantly, also in the appropriateness of the interpretation and efficiency in the calculation of the solution to these additive word problems following the intervention, mixed methods were required. Whilst the nature of the changes in performance can be expressed in purely quantitative terms, the shifts in structural and numerical reasoning, the analysis of which was conducted on the basis of how learners' models were configured and worked with, can only be expressed in qualitative terms, and hence, required the use of qualitative methodology.

The founders of the Relational Paradigm and RME agree that human activity is *goal-oriented* activity (Davydov & Markova, 1982) and mathematics is a human activity (Freudenthal, 1986). They would also agree that, given its orientation to study human intellectual development in terms of concepts and approaches established in the 'hard sciences', positivism is unable to reveal 'the essence of things fixed in such concepts as "activity"' (Fellus & Biton, 2017, p. 145). Constructivism, on the other extreme, initially sought to explain human intellectual development in cognitive terms. In his theory of cognitive development, Piaget (1973) ventured to explain how humans construct meaning by integrating new ideas into their experiences as separate human individuals. He therefore argued for instruction to be 'developmentally appropriate', meaning that at each stage of instruction, the child should only be offered knowledge that he can understand right away. Thus, according to cognitive constructivism, a major factor shaping what is to be taught to a child at any given time is what the child already knows.

It was Lev Vygotsky who singled out social interaction as the greatest motivating force in human development, and by so doing, laid the foundation for a socio-cultural theory of human intellectual development. Socio-cultural theory stands opposed to cognitive constructivist 'theories that proclaim a naturalistic understanding of human psychical development based on features of the separate human individual whilst ignoring the initial sociocultural and semiotic nature of the human psyche' (Fellus & Biton, 2017, p. 140). On the contrary, socio-cultural theory admits of instruction and upbringing – where each individual is continually interacting with other more knowledgeable individuals – as, not only essential, but also necessary for human intellectual development. In other words, Vygotsky initiated a paradigm shift from cognitive constructivist theories that thought of a child's development as preceding learning, to socio-cultural theories that perceive of the development of the child as resulting from teaching and learning as social interaction processes.

Vygotsky's tenet of collaboration among learners in the process of their upbringing and teaching provided the theoretical base for Davydov's (2008) *Theoretical Thinking* approach. The fundamental principle behind Davydov's approach is that learners' initial encounters with problem solving should not be about finding a numerical solution. Davydov envisioned a learning trajectory in which learners' initially play with physical objects such as strings, and describe relationships between quantities without referring to numbers. Only later, when they are acquainted with physical qualitative relations between continuous quantities in a general and abstract form (A+B=C), bar or segment diagrams, are learners introduced to solving additive problems with the use of those abstract models. Inversely, based on Davydov's theory and the concept of theoretical thinking, Polotskaia (2015) proposes that the additive word problem-solving activity should be about analysing and modelling the additive relationship.

Davydov (2008) defined the notion of an additive relationship as 'the law of composition by which the relation between two elements determines a unique third element as a function' (Davydov & Markova, 1982, p. 22). In his theory of developmental instruction, Davydov (2008) argues that an understanding of the principle of composition is something that should be brought into learners' early awareness of quantitative relationships and developed as a fundamental concept of their learning if it is to develop into a flexible understanding of additive structures over time. Building on this conviction, Davydov (1982) showed that, following an appropriate instructional design, young learners who begin their formal mathematical instruction, are capable of grasping this qualitative relationship. In this regard, Davydov (1982) proposed an approach that focuses explicitly on developing learners' awareness of structural relations between quantities through the use of non-numerical representations of quantities.

Given that this study was conducted in a SA curriculum context where learners in Grade 3 are already operating on number, and where Davydov's approach is unfamiliar, my attention was on structural relations involving numerical representations of quantities. By conducting this study within the Relational Paradigm, which is underscored by Davydov's vision of mathematics as the study of quantitative relationships (Polotskaia, 2017), my plan was to adopt a methodology that would open up spaces for collaboration and interaction during instruction, using word problems set in familiar contexts to encourage the sense-making needed for interpretation phase working. As a result, the instruction took the form of lessons conducted under as 'normal' South African teaching conditions as possible, with myself as the teacher. After all, ‘all forms of human intellectual activity, including scientific activity, are realised by social processes rather than by isolated individuals’(Davydov & Markova, 1982, p. 94).

Design-based research methods have been touted for providing insights on learning naturalistic intentional settings, particularly where the study involves ‘designing an intervention that reifies a new form of learning’ (Tabak, 2004, p. 226). With the overarching exploration of this study into the uptake of the BD and ENL as relational models that can support structural and numerical reasoning, the naturalistic setting for this study was the Grade 3 classroom and the ‘theoretical intent, therefore, [was] to identify and account for successive patterns in student thinking by relating these patterns to the means by which their development was supported and organised’ (Cobb, Confrey, DiSessa, Lehrer, & Schauble, 2003, p. 11). Given the exploratory nature of the design of this study, and the decision to conduct the study under as ‘normal’ South African teaching conditions as possible so as to contribute meaningfully to both research and practice, design-based research methods were employed.

In light of my dual role as researcher into the competence that can be effected through this intervention study and teacher during intervention lessons rolled-out as part of the study, I attend in section 3.8 to the ways in which I ensured objectivity. For the moment though, I turn my attention to the sources of data used in this study.

3.4 Data sources

With an interest in the patterns of gain that could emerge from an intervention designed to work from a theoretical base that values a modelling approach to solving additive word problems, I worked with one WMC-P Grade 3 class in one of the ten schools in the broader project. The school is a suburban school serving a historically disadvantaged population, with English as the language of learning and teaching. The focal class had 40 learners, with matched initial and post test data for 39 of these learners.

In an exploration of the affordances made possible by an intervention using relational models to solve additive word problems, the answer to my first research sub-question

What were the overall patterns of learner performance across different types of additive relations problems prior to the intervention, and what shifts in performance were observed from pre- to delayed post-test?

- came from a comparative analysis of learners' performance in each of the three written test sittings. This comparison allowed for delineation of changes in learner performance across the tests, particularly as it related to the additive word problem types that the learners had to solve in each of the three test sittings. With this study conducted within the Relational Paradigm, this afforded me with an opportunity to analyse the extent to which the hierarchy of difficulty asserted by Nesher and colleagues to exist amongst different additive word problem types could be disrupted by learners' working relationally, as hypothesized by Polotskaia and colleagues.

The intervention was guided by what literature had to say regarding models that researchers advocated as useful for modelling additive situations. Data on learners' work with models and the patterned differences in the use of these models were drawn from my analysis of learner responses in the pre-, post- and delayed post-test. This analysis enabled the answering of my second research question:

What nature and patterns of model use underlay the difference in learner performance across the three tests, and to what extent was there evidence of models that reflected appropriate interpretation of the mathematical structure and correct calculation of the numerical solution?

Performance in the pre-test led to the identification of three sub-categories of items, detailed later in Chapter 4. In order to understand how patterned differences in the use of these models played out across the different sub-categories of test items based on pre-test performance, the following two questions were instructive:

- *What is the nature of the shifts in interpretation across the responses in each sub-category?*
- *What is the nature of the shifts in learners' co-*ordination *of interpretation and calculation across the responses in each sub-category?*

Given my complementary interest in the nature and extent to which the rudimentary strategies found to dominate learners' early working with additive word problems were transformed following the intervention, I also analysed learner responses in the pre-, post- and delayed post-test for efficiency in calculation, in line with my third research question:

What was the nature of the calculation strategies used by the learners across three tests (pre-, post- and delayed post) and to what extent did the use of these strategies contribute to efficient calculation of the numerical solution?

Across all three research questions, therefore, learner responses in the pre-, post- and delayed post-test comprised the key data sources. In order for the intervention to be deemed successful, however, learners had to demonstrate increased facility in setting up the models – from which structural reasoning could be inferred– and in using more efficient strategies– from which numerical reasoning could be inferred – to arrive at the correct solution to the problem following the intervention.

3.4.1 The test

In order to make the test encompassing of the different kinds of word problems that learners can encounter in the South African school curriculum, the test items were compiled following the classification of word problems generally agreed upon by research into learners' processing of additive relations word problems. Eventually, 12 questions were selected from across the three semantic categories of change, combine and compare, and nature (or position) of the unknown for inclusion in the test administrations. The selections also took into consideration possible variations with respect to consistency and inconsistency in the language of the text of the word problem, noting – as detailed in the literature review – that this feature interplays with the level of difficulty of items.

The pre-test, post-test and delayed post-test were administered on the 13th August 2015, the 26th November 2015 and the 25th January of the following year respectively (after the long end of year break). In administering the tests, each problem was read out twice for the learners by the teacher (taking into account the broader evidence of low levels of reading proficiency in South Africa), following which learners were given 5 minutes per question in order to allow for the use of tally counting if that was what learners opted to use.

Tabulated in table 5 below are the 12 additive word problem test items:

Q#	Word problem	Category and nature of unknown	Consistent or Inconsistent language
1	On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether?	Change increase Result unknown	Consistent
2	Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether?	Combine Whole unknown	Consistent
3	52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether?	Combine Whole unknown	Consistent
4	Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost?	Change decrease Change unknown	Consistent
5	Chris gave away 19 stickers and was left with 34 stickers. How many stickers did Chris have at the beginning?	Change decrease Start unknown	Inconsistent
6	69 children travel by bus. 77 children travel by train. How many more children travel by bus?	Compare Difference unknown	Inconsistent
7	James had 8 marbles. John gives him some more marbles. James now has 77 marbles. How many marbles did John give to James?	Change increase Change unknown	Inconsistent
8	There were 40 people at the park. 17 were adults and the others were children. How many children were at the park?	Combine Part unknown	Consistent
9	Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles?	Compare Difference unknown	Inconsistent
10	18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls?	Combine Part unknown	Consistent
11	Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick?	Combine Whole unknown	Inconsistent
12	There were 77 people on the school bus during the trip. 69 of them were children and the others were teachers. How many of them were teachers?	Combine Part unknown	Consistent

Table 5: The 12 additive relations word problems in the test

3.5 The intervention

I have already noted that prior research has found that word problem tasks have been useful in supporting sense-making, but also pointed out the caveat that expecting learners to engage in the sense-making required for interpretation naturally and independently was unlikely in the South African context. This fed into the need for some careful choices for a sequence of 10 contact session intervention lessons if interpretation and calculation-phase working were to be supported.

Askews's (2004) Big Book of Word Problems resource delineates an instructional sequence that features word problems (examples and worksheets), ENL models, teacher orchestrated discussions of links between problems, whole class, pair and independent working on similar problems. These materials, developed in UK, tend to be aligned with the RME view of emergent modelling as natural, and so the intervention lesson structure needed some adapting to both incorporate attention to BD as a key structural reasoning model but also to some explicit instructional focus on BD and ENL as key relational models. An important and explicit part of the task design in each of the sets of problems is that they are intended to help move the learners from treating 'each problem in isolation from other problems that they have worked on' to treating 'each problem as being one of a generic class of problems' (Askew, 2004, p. 5). Askew's (2004) generic classes of additive word problems fit the detailed analysis of additive word problem types (semantic categories and nature of unknown) set out in chapter 2.

3.5.1: The intervention process, resources and rationales

In order to get learners to treat each problem as being one of a generic class of problems, each semantic category of additive word problem was introduced, identified and exemplified as each lesson unfolded. In particular, Change problems were introduced in the very first lesson, Compare problems in the fourth lesson, and Combine problems only in the 9^{th} lesson. Also, Compare problems received the most attention (lessons 4 – 10), followed by Change problems (lessons 1, 2, 3, 9 and 10) and Combine problems (lessons 9 & 10). The motivation for spending a disproportionate amount of time on Compare problems related to the consensus amongst researchers regarding the relative difficulty associated with this category of problems when measured against that of combine and Change problems (Carpenter et al., 1999; Nesher et al., 1982; Vergnaud, 1982; Verschaffel et al., 2000). Also, as already mentioned in chapter 2, the choice of the static bar diagram as a relational model best suited for supporting the requisite structural reasoning in the interpretation phase resonated with Davydov's (1982) idea that 'the additive relationship does not include any action, but only describes the relationship between elements and how they comprise the whole and the parts' (Polotskaia, 2017, p. 165). In turn, this made it the ideal artifact for handling problems whose language suggests a static relationship between the component parts of the additive relationship (Carpenter et al., 1999), in particular, Compare and Combine problems.

Extending the use of the bar diagram to include supporting the interpretation of change problems was in line with the Relational Paradigm advocacy for instruction to facilitate a systemic view of the additive relationship as 'the law of composition by which the relation between two elements determines a unique third element as a function' (Davydov, 1982, p. 229) encompassing all additive relation problems, regardless of their type. Consequently, all semantic categories of additive relations word problems were included across lesson example, learner exercise and test item spaces.

3.5.1.1 The Big Book of Word Problems

The additive word problems that were used as examples and tasks during intervention lessons were taken from the Big Book of Word Problems by Askew (2004). Each of the ten lessons had a different theme while preserving the same basic form so that learners could quickly get used to 'what is expected of them in the lesson and they can then concentrate on thinking about the mathematics' (p. 6). The Big Book intervention lesson model suggests four stages to each lesson, namely:

- Solving the introductory contextual Big Book problems
- Linking up the introductory contextual problems
- Follow-up problems, and
- Wrap-up

A key lesson structure adaptation that I made involved the inclusion of an initial section of focused explicitly on introducing and discussing combinations of 10 and 20 in BD and ENL formats. This was intended to address the concerns raised about learners' lack of familiarity with emergent modelling as a way of working, and the concerns relating to inefficient counting-based approaches through attention to base 10 thinking. This produced a 5-stage lesson rather than the four stage model outlined by Askew (2004). More generally, this initial part of the lesson (henceforth referred to as the connecting models stage) was used – in out-of-context ‘bald’ number engagements – to demonstrate translating between a BD, a number sentence (NS) in standard form and an ENL, before applying it to the solving of additive word problems in the *solving the introductory contextual Big Book problems* stage of the lesson in accordance with the problem-solving cycle.

3.5.1.1.1 The connecting models stage

Savinainen and Viiri (2008) define *representational coherence* as the ability to use multiple representations and move between them. In an intervention study that sought to promote the BD and ENL models to support the interpretation and calculation of additive word problems respectively, there was a need to open a space in which the models would be introduced and the connections between them foregrounded. The connecting models stage was therefore a space opened up during the lesson to provide learners with an opportunity to hone their *representational coherence* skills for using and moving between a BD, a number sentence (NS) in standard form and an ENL. In helping learners’ get acquainted with each model and the connections between the models, in my role as the teacher during this stage of the lesson I would introduce and discuss the relational models and the accompanying notation, demonstrate how to set up the relational models, and to associate particular meanings to particular configurations of the BD and the ENL models, whilst attempting to show how the different models can reflect the same additive relationship.

In setting the scene for the introduction of the ten-based calculation strategies promoted in this study for efficient calculation during the calculation phase of the problem-solving cycle, a key idea flagged during the connecting models stage of the first lesson was how each configuration of a bar model can be associated with a family of two addition and two subtraction number sentences. When making explicit the connections between the addition and subtraction number sentences, the part-whole structure of the BD served as a visual mediator of the commutative property of addition and the inverse relationship between addition and subtraction.

The connecting models stage of the intervention lessons was also used to introduce learners to different ways of configuring number triples in a BD. Part of the engagement was to relate different configurations of number triples in a BD to different families of number sentences, thus preparing learners' to grasp the additive relationship as a composition of three quantities during the interpretation phase of the problem-solving cycle, as shown in figure 9.

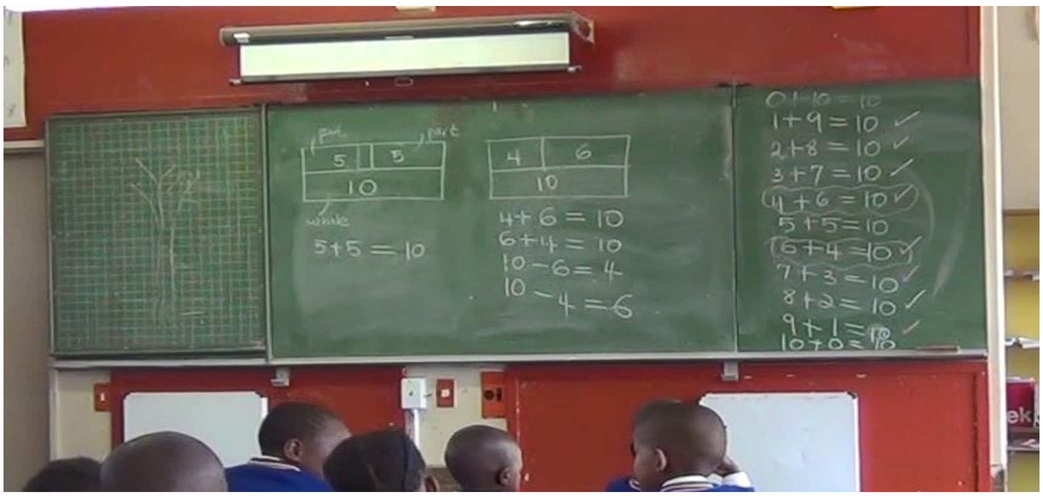


Figure 9: Relating configurations of number triples involving combinations of 10 in lesson #1

An important connection made in the connecting models stage of lesson #1 was between the combinations of 10 and how these could be modelled on a bar diagram (BD) and empty number line (ENL). Whilst modelling the combinations of 10 on a BD was useful for foregrounding the part-whole constructions of 10, modelling them on a number line provided preliminary encounters with partitions of 10, with the aim that knowledge of these partitions would later serve as recalled number facts when the need to use the adding through ten (A10) strategy promoted in the study eventually arose.

Following these decontextualized engagements designed to introduce and acquaint learners with the necessary models as objects that I wanted to serve as tools for solving additive relations word problems with structural and number sense, learners were given the opportunity to experience how these translations between models could serve the problem-solving process.

3.5.1.1.2 Solving of the introductory contextualised Big Book problems stage

The first 15 minutes of the *solving of the introductory contextualised Big Book problems* stage of the lesson were used to introduce the theme for the lesson and to use the three introductory contextualised problems as examples of how to solve additive relations word problems. In figure 10 is as an example from lesson #4 of the Year 4 Big Book of the kind of the introductory contextualised problems used during lessons, and of the contexts (and themes) that formed the basis of the problems encountered by the learners during the intervention:

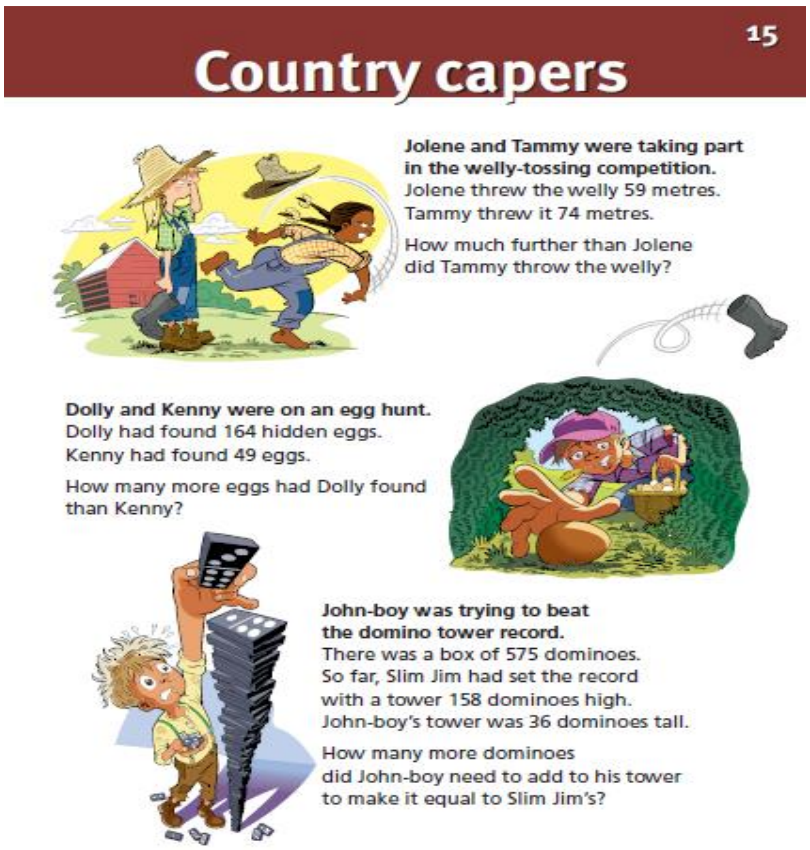


Figure 10: Lesson #4: introductory contextualized problems (Askew, 2004)

These introductory contextualised problems functioned as a base for exemplifying the protocol followed when solving additive relations word problems. In this stage of the intervention lessons, I, as the teacher, orchestrated discussions revolving around the sense that learners made of each problem, which model to use when and why, and the strategies they used to solve them.

My role as the teacher during the solving of the introductory contextualised problems was to direct learners' attention to the problem-solving cycle designed to guide their reasoning. This was managed through asking guiding questions that kept learners attention on the sequencing of the models by referring back to the big idea in the connecting models stage of the lesson of ensuring that the models coalesced one with the other. As shown in figure 11, the problem-solving cycle prescribed reading the text of the problem and expressing their understanding of the structure of the additive situation as a part-whole relationship in a BD, translating the structural relations (as expressed in the BD) into a mathematical expression that then serves as the basis of a calculation sequence (Polotskaia & Savard, 2018). This calculation sequence was then set in motion using the ENL model that harnesses the sequential structure of number.

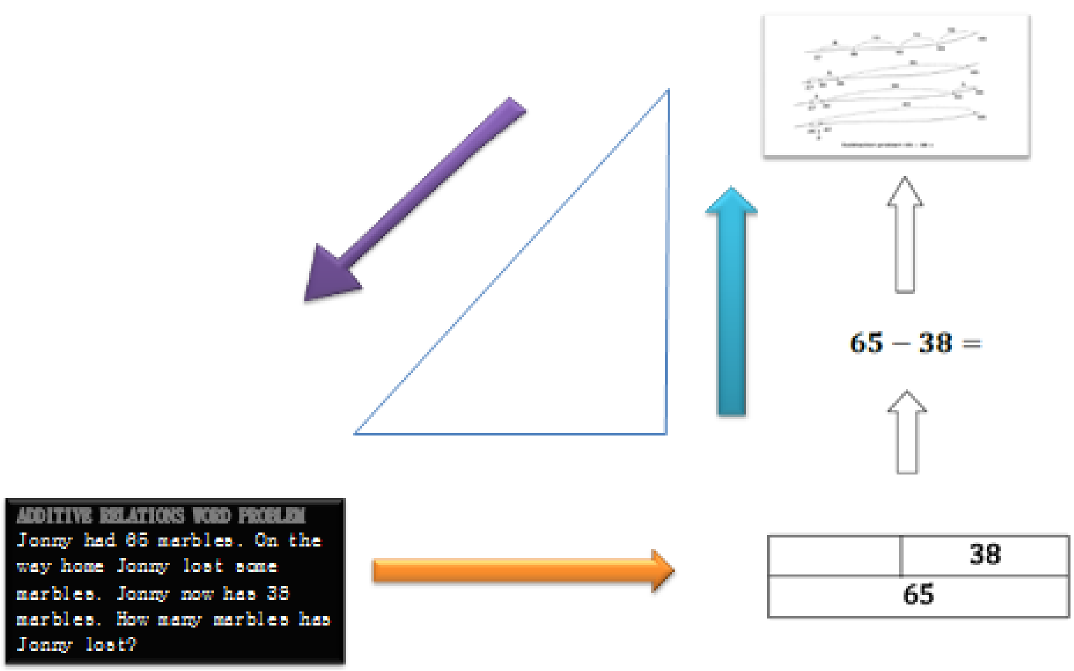


Figure 11: the problem-solving cycle and the models promoted in this study

After solving the introductory contextualised problems on the board, I then refocused the attention of the lesson on a whole class discussion in which the links between the three introductory contextualised problems, and between the models constructed for problem-solving, were explicated.

3.5.1.1.3 The linking-up the introductory problems stage

In order to explicate the links between the introductory contextualised problems exemplified in the connecting models stage of the lesson, I orchestrated the discussion to revolve around bringing out the character of the three problems that had just been solved, and drawing learners' attention to what (Askew, 2004) describes as the 'common mathematical structure underlying the problems' (p. 7). This was achieved through introducing and engaging with each of the three different semantic categories of the problems provided by literature.

During discussions of the sense that learners were making of the word problem in question, I foregrounded the semantics (words and phrases) that were indexical – not of the operation implied in the wording of the problem, but instead – of the class to which the problem belonged. In particular, learners were advised to look out for phrases such as 'how much further' and 'how many more' in the question of the word problem as indicative of a Compare problem. From the point of view of deducing the arithmetic operation from the BD, knowledge of the different semantic categories was used – in lesson #4 – for instance, to establish a generic mapping from a Compare problem situation, through a part-whole BD, to a number sentence implicating a subtraction operation.

3.5.1.1.4 Follow-up problems

Following the linking-up of the problems in a whole class discussion, the learners were then afforded an opportunity to work more independently on the ideas just discussed in class. This was achieved through an exercise provided to learners in the form of a worksheet whose problems had the same underlying structure as the three introductory problems that had just been discussed. This individual engagement with the worksheet formed the basis of the sharing session in pairs that constituted the second part of the follow-up phase.

Following the introduction of Compare problems in lesson #4, an important variation that related to calculation strategies was introduced to calculate the solutions to two contextualised problems, as shown in figure 12. In the former, the use of the N10 was demonstrated, and in the latter, it was the use of the A10 that was demonstrated.

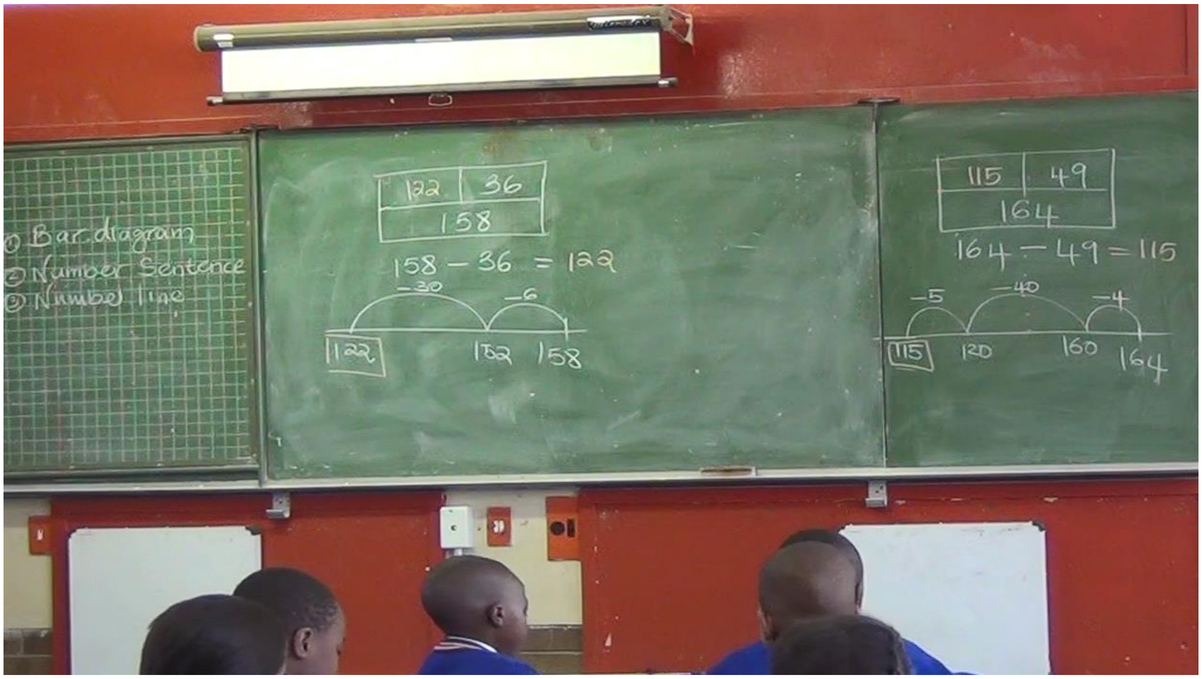


Figure 12: Two calculation strategies for solving contextualised problems demonstrated on the board during the linking-up the problems and the follow-up problems stages in lesson #4

The idea of leaving the two illustrations simultaneously visible on the board during the follow-up problems stage meant that both exemplifications, were available to learners to refer to while working on the follow-up problems, thus making it possible to explicate the similarities and differences in the use of the N10 relative to the A10 strategy.

This kind of teaching that emphasizes simultaneity of examples and connections within and across examples and representations was found by Venkat, Ekdahl, and Runesson (2014) to have the potential of closing gaps in learning outcomes seen in pre-tests, and thus improving performance. In emphasizing making connections within and across examples and representations, this teaching approach is in line with the relational view promoted within the Relational Paradigm. With the BD, NS and ENL simultaneously visible in both examples, and the sequencing of the models coded to the left of the two examples serving as a reminder of the protocol to first interpret the mathematical structure of the word problem into a BD, which is then translated into a NS, and finally into an ENL, it was possible to make explicit connections between the similarities in the configurations of the BD in each example, and the subtraction operation in the NS.

3.5.1.1.5 Wrap-up

Towards the end of each intervention lesson, the class was reconvened into a whole class discussion for further exploration of the nature of the problems in the worksheet and consolidation of the key ideas. Often this stage of the lesson was used to reiterate the big ideas flagged in the linking-up stage of the lesson. In particular, two key messages were shared with learners. These related, firstly, to establishing linguistic connections between problems in the same semantic category for more appropriate interpretation of situations, and, secondly, paying attention to the different ways of working with the three calculation strategies as well as to the numbers involved so as to use the most efficient strategy. With contexts, problem types and calculation strategies continuously varying, the relational models remained as a constant: 'invariant across examples, thus emphasizing their general usefulness' (Venkat et al., 2014, p. 2). In table 6, I provide a brief description of what was covered in each of the ten intervention lessons, as well as the dates on which the lessons took place[4].

[4] In Appendix A, I relate the specifics of the descriptions of each of the 10 lessons to the five stages outlined above.

	Descriptions of the lessons	Date
Lesson #1	Out of context **introduction of the BD & ENL** with partitions of 10. Families of number facts associated with a single configuration of a number triple in a BD. Translations from BD to ENL exemplified. Two Change Increase result unknown problems used to exemplify the interpretation of a Change Increase Result Unknown additive word problem into a BD, and completing the model-based solution cycle. **N10 & A10 strategies introduced**, compared and contrasted. Learners get first opportunity to apply the problem-solving protocol.	20 Oct
Lesson #2	Out of context **recap on the structures of the BD & the ENL** with partitions of 20. Families of number facts associated with a single configuration of partitions of 5 & 6 in a BD. Translations from BD to ENL exemplified. One Change Increase Result Unknown problems used to consolidate making sense of such problems and a reminder of the model-based solution process. A10 used as a calculation strategy. Learners solve two Change Decrease Result Unknown Problems. Teacher uses the A10 to solve the first problem, and uses the second problem to **introduce the N10C strategy**.	21 Oct
Lesson #3	Out of context **recap on part-whole structures of BD**. Same number triple used to recap implications of different configurations of the BD for the choice of arithmetic operation. Resulting number sentences used to recap implication of the different arithmetic operations for the direction of the jump on the ENL during calculation. One Change Increase and one Change Decrease Result Unknown problem solved in whole class demonstration on the board. Differences in interpretation highlighted. Opportunity used to highlight efficiency of a single-multiple-of-ten jump as compared to multiple jumps with 10 as a unit. **Efficiency of N10C over the N10 also highlighted**.	27 Oct
Lesson #4	One Change Increase and one Change Decrease Result Unknown problem solved in whole class demonstration of the difference in interpretation of the two into the BD and calculation using the N10 strategy, with emphasis on the direction of the jump and **efficiency of single -multiple-of-ten jumps.** Teacher used practical example of differential heights of learners in the class to **introduce Compare problems, and recaps on the problem-solving protocol** as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL**. Learners get first opportunity to apply the problem-solving protocol to two Compare problems. Language of Compare problems foregrounded for sense-making.	28 Oct
Lesson #5	Reminder of models used in the protocol. Interpretation into a BD and calculation with ENL of two Compare problems with additional irrelevant data demonstrated on the board. Learners' attention directed to the presence of **additional irrelevant data** which learners need to factor out of the interpretation and calculation. Language indicative of Compare problems highlighted for sense-making.	03 Nov

Lesson #6	Out of context recap on part-whole structures of BD. **Same number triple used to recap implications of different configurations of the BD for the choice of arithmetic operation**. Resulting number sentences used to recap implication of the different arithmetic operations for the direction of the jump on the ENL during calculation. One Change Increase and one Compare problem solved in whole class demonstration on the board. Semantics of Change problem distinguished from semantics of Compare problem highlighted along with the implications for interpreting each one into a BD. Follow-up problems used to emphasize the appropriate interpretation of individual problems in an exercise set of six additive word problems. Reiterated that Compare problems invoke subtraction.	04 Nov
Lesson #7	**Recap on the problem-solving protocol** as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL.** Another practical example involving heights of two learners in the class used. Recap on the difference between semantics reflecting a **Compare** problem contrasted from the semantics of a **Change** problem. Teacher does Change problem on the board, learners do two Compare problems individually/ in pairs. Language of Compare problems highlighted as important to recognise if sense to be made of the problem is to be appropriate. Teacher uses opportunity to revisit efficiency of a single-multiple-of-ten jump over multiple jumps with 10 as a unit. Efficiency of N10C over the A10 also highlighted.	10 Nov
Lesson #8	More practice with Compare problems. All six Compare problems in exercise set 17A done by learners' individually\in pairs, and corrected by teacher on the board. Each time teacher solicits sense that learners made of the problem being solved based on the text of the problem for appropriateness of learners' interpretations. First problem done using the N10 & the A10, and follow-up discussion emphasises relative efficiency of adding/subtracting through 10. Language of 'how much further' and 'how many more' re-emphasized as indicative of a Compare problem.	11 Nov
Lesson #9	**Recap on the problem-solving protocol** as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL**. Language associated with each category revisited. Family of number facts associated with a single combination of number triple modelled on an ENL. **Combine problems introduced.** All three *intro context problem* done on the board. With each one representing a different category, emphasis on appropriate interpretations. Two problems from the follow-up exercise set done by learners in pairs and then by the teacher on the board. One of them used to reiterate efficiency of N10C over N10.	18 Nov
Lesson #10	Family of number facts associated with a single combination of number triple modelled on ENLs. Recap on the difference between semantics reflecting a **Compare** problem contrasted from the semantics of a **Change** problem. **Combine problems recapped.** One Change Increase Result Unknown problem done by teacher on the board, and another done by learners individually/in pairs, and then by the teacher. In both solutions, the use of the A10 & the N10C was promoted.	24 Nov

Table 6: Brief descriptions of what was covered in each of the ten intervention lessons

3.6 Analytical processes

Following the collection of the necessary data from the three sittings of the test, I conducted three analyses that would allow me to answer the three research questions outlined above relating to the contributions to additive relations problem solving competence that came to be effected through an intervention that focused on the use of the bar diagram (BD) and the empty number line (ENL) as relational models. Overall, I conducted three micro analyses relating to the Interpretation, Calculation and Evaluation Phases as promoted during intervention lessons, and one macro analysis relating to the extent to which leaners' responses reflected appropriate coordination of structural and numerical reasoning.

3.6.1 The micro analyses

At the micro level my interest was in the extent to which the use of models reflected the correct interpretation of the additive relationship, the correct calculation of the solution to the problem and the correct answer. In the first micro analysis, given the truncated nature of my interest in the Evaluation Phase detailed in the theoretical framing, the primary interest was in the extent to which there was an improvement in learners' performance following the intervention. Consequently, the data analysed in chapter 4 were the answers that the learners provided to the items in the pre-, post- and delayed post-test: learners' written test responses therefore constituted the unit of analysis. This allowed also for an analysis of the extent to which the evidence of change in performance supported the conjecture made in the Relational Paradigm that working relationally allowed learners to move towards being as effective in solving the theoretically 'difficult' problems as in solving the theoretically 'easy' (Polotskaia, 2017) problems. To this end, I made a comparison between the empirical ranking-order of performance on individual items (relevant to this cohort of learners) with the expected performance on the items based on the difficulty levels theorised by research. The use of the hierarchy of 'levels of difficulty' hypothesised by Nesher et al. (1982) served as a useful reference point.

As outlined in chapter 2, researchers (including Carpenter et al. (1996); Verschaffel, De Corte, and Lasure (1994); Hegarty et al. (1995); Pape (2003) and Giroux and Ste-Marie (2001)) are in general consensus on the existence of a hierarchy of difficulty associated with different categories and types of additive word problems, found by Nesher et al. (1982) following investigations based on their hypothesis of developmental levels of learners' performance on additive word problems at different ages. These researchers agree that the differences in difficulty could be accounted for in terms of the 'semantic structures and the operations required to derive the meaning structures from the problem texts' (Nesher et al., 1982, p. 373). In particular, they posit that change-increase problems in which the result is the unknown are the easiest, and Compare problems in which the referent set is the unknown are the most difficult. Table 7 summarizes the hierarchy of difficulty that was introduced in the previous chapter, and links this hierarchy to items in the test administrations.

Levels according to Nesher et al. (1982)	**Additive word problem-solving ability that relates to the identified level according to Nesher et al. (1982)**	**Problem category and nature of unknown according to Nesher et al. (1982) and Giroux & Ste-Marie (2001)**	**Problems in the test**
1	Reference to sets, adding members to and removing members from sets. Understanding 'put' 'give 'take', etc. as denoting change in location or possession.	Change Increase/Decrease Result Unknown; Combine Result Unknown	Qs 1, 2, 3 & 11
2	Ability to link events by cause and effect. Reference to the amount of change. Understand sequence of events ordered in time in a non-reversible manner.	Change Increase/Decrease Change Unknown;	Qs 4 & 7
3	A reversible part-part-whole schema is available, and can be used to find unknown part in any slot in a sequence of events. Understanding class-inclusion.	Change Increase/Decrease Start Unknown; Combine Part Unknown; Compare problems mentioning 'more'/'less' and the difference set is the unknown; Compare problems mentioning 'more'/'less' and the compared set is the unknown.	Qs 5, 6, 8, 9, 10 & 12
4	Reversibility of non-symmetrical relations. Ability to handle directional descriptions (more/less), and quantify a relational set (relative comparison).	Compare problems mentioning 'more'/'less' and the referent set is the unknown.	(Not included amongst test items)

Table 7: The 'levels of difficulty 'associated with the different additive word problem types

In the second micro analysis, I attend to the Interpretation Phase, and I analyse the models in learners' responses for sense making, with a view to determining the extent to which there was an improvement in learners' interpretations of additive situations into models. A distinguishing feature was a model configured to reflect relational structuring of the quantities in their respective roles, representing the parts and the whole accordingly. It is in an analysis of learners' interpretations that what has been referred to as a 'reversal error' (Pape, 2003) to denote an 'erroneous choice of arithmetical operation' (Fischer et al., 2019) is addressed. Whilst the secondary analysis in chapter 4 relates to the hierarchy of difficulty as expressed in terms of the semantic categories and nature (position) of the unknown, with each item also identified as consistent or inconsistent problem in table 5 in section 3.4.1, the shifts in terms of consistent or inconsistent problem could be inferred.

In the third micro analysis I attend to the strategies that learners used to calculate the solutions to the problems in the Calculation Phase, and the shifts in these strategies from one test to the next. Whilst the feature of a part-whole construction of an additive relationship in a model provides me with a basis for commenting on the sense made by the learners of the different problem situations, a distinguishing feature of sophistication in calculation was a part-whole construction of number (Ellemor-Collins & Wright, 2009). A part-whole construction of number is a feature of non-counting calculation strategies that distinguishes them from counting strategies as it indicates a conception of numbers as composed of other smaller numbers. Firstly, the notion of a part-whole construction of number enabled differentiating between models that support calculation strategies from those that support only counting strategies. Secondly, it provided the study with an analytical framing informed by the hierarchy of calculation strategies developed in the RME tradition (Beishuizen et al., 1997; Blöte et al., 2000) in an attempt to differentiate between calculation strategies according to their efficiency during calculation.

The calculation strategies promoted during intervention lessons were those that are supported by the ENL, namely: the N10, A10, and the N10C strategies, which are all strategies that use 10 and its multiples/partitions as reference points during calculation. The use of the ENL with non-unit partitions, therefore, provided a marker of the use of non-counting strategies, whilst the use of tallies reflected the use of the count-on or the count-all strategy. Whilst the column addition/subtraction algorithm was regarded to be the strategy used whenever a learner modelled the problem columnally, the inclusion of a BD or NS without a calculation sequence provided no overt evidence of calculation strategy. All responses reflecting only a BD or NS with no overt calculation seen were interpreted as involving mental calculation. In table 8, the hierarchy of sophistication of different calculation strategies is presented from mental strategies which for this study represent the most sophisticated calculation strategy, to the use of tallies (TL) which, for being devoid of number structuring, represents for this study the least sophisticated calculation strategy.

Models	**Strategies**	**Nature of strategy**	**Descriptions of the calculation strategies**
Bar diagram (BD)	Mental	Mental	Mental
Number Sentence (NS)			
Empty Number Line (ENL)	Connecting arc (CA)		Adds-on the difference instead of taking away the second operand, or subtracts the missing added. For example 51 – 49 will be treated as 49+? = 51 (and the answer is 2)
	N10C	Strategies that use 10 and its multiples/partitions as reference points	Adds-on (or takes away) a value that is a multiple of ten closest to the second operand and then **compensates** for the amount by which the jump was exceeded
	A10		Stands for **adding through 10** and is a variation of N10 in which the second operand is decomposed in such a way as to enable landing on a multiple of ten
	N10		'N' represents the initial number that is kept whole while '10' represents the series of tens and then units that are then added or subtracted. For example, 38 + 24 might be solved: 38 + 20 = 58, 58 + 4 = 62.
Column Model	Column addition or subtraction algorithm		Add (or subtract) the units followed by adding (or subtracting) the tens.
Tallies	Counting	Counting	Count each and every tally mark or count-on from first quantity

Table 8: Hierarchy of sophistication in calculation strategies

3.6.2 Interpretation-Calculation-Evaluation: macro analysis overview and coding scheme

The macro analysis was undertaken with recognition of the fact that, whilst the individual models were meant as individuated support for the interpretation and calculation phases, the use of the models was meant to be coordinated in such a way that the models functioned as component parts of a closed and reversible system (Rabardel & Bourmaud, 2003). With the BD and the ENL associated with the interpretation and the calculation phases at a micro level respectively, in cases where they were both used, they were also intended to coalesce one with the other for coordination of structural and numerical reasoning at a macro level, thus aligning with the design of the intervention in this study.

An aspect of learner responses that became apparent early in my data capture was evidence of an increase in the extent of inclusion of interpretive models of situations in learners' responses across the three test administrations, likely linked to my attention in the intervention lessons to using the BD to support the interpretation phase, and the ENL to support the calculation phase. With an interest in learners' success in coordinating both structural and numerical reasoning to arrive at the correct answer**,** the macro-analysis of learners' solutions proceeded with the truncated interpretation-calculation-answer cycle as the basis. Profiling learner responses according to the interpretation-calculation-answer cycle provided a vantage point from which to view the success of the models in mediating the solution process at a macro level.

The analytical focus on the Interpretation Phase allowed for an exploration of the extent to which the increasing number of models appropriately reflected the problem situations presented in each item. Every learner response across the three test administrations was coded in one of three ways for the Interpretation Phase: * (no model), 0 (inappropriate model), 1 (appropriate model). Moving into the Calculation Phase, each response was analysed for evidence of an appropriate calculation, and coded using the same three options: * (no calculation), 0 (inappropriate calculation), 1 (appropriate calculation). With answers given in every response across the dataset regardless of the nature of the response, the Evaluation phase was coded with a 0 (for incorrect answers) or a 1 (for correct answers). This resulted in a three digit coding system that summarised the success/failure observed in each of the three phases. These *3-digit codes* describe a learner's engagement with an additive relations problem, with particular attention on whether or not the interpretation of the situation was correct or not; whether the calculation of the solution was correct or not, and; whether or not the answer provided was correct. For example, a learner response showing a correct interpretation of the problem situation into a model, followed by an incorrect calculation attempt and an incorrect answer to an item was allocated a '100' code; an incorrect answer with no working shown was allocated a '**0' code. In table 9, I offer a summary of this coding scheme.

Digit	**Interpretation**		**Calculation**		**Evaluation**
***/0/1**	**Full description**	**Short description**	**Full description**	**Short description**	**Description**
*****	No evidence of a model reflecting an interpretation of the problem situation	No model	No evidence of a calculation	No calculation	No answer (not in evidence in dataset)
0	Evidence of an incorrect interpretation of the problem situation as seen in (an) inappropriately configured model(s)	Inappropriately configured model of situation	Evidence of an inappropriate calculation	Incorrect calculation	Incorrect answer
1	Evidence of a correct interpretation of the problem situation as seen in (an) appropriately configured model(s)	Appropriately configured model	Evidence of an appropriate calculation	Correct calculation	Correct answer

Table 9: Key to coding for appropriateness of Interpretation, Calculation & Evaluation

While theoretically, a '3-digit' code with three options for each coding segment can produce 27 coding options, the empirical data analysis revealed 12 different ways of working across the 468 responses and across the three test sittings (see table 10). Of these 12 different ways of working, 2 were 'answer-only' responses, so they could not be analysed for interpretation and calculation. In the absence of any working out, they were coded as simply **1 where the correct answer was provided, and **0 for an incorrect answer. The other ten ways of working demonstrated by the learners related to modelled responses. Quantifying the prevalence of each of the resulting 3-digit codes provided a useful shorthand summary of different profiles of engagement with the interpretation, calculation and evaluation phases of working in learner responses across the pre-, post- and delayed post-tests.

Interpretation phase analysis	**Calculation phase analysis**	**Evaluation phase analysis**	**3-digit codes**
Modelled responses	With calculations	answers correct	111
			001
			101
		answers incorrect	000
			100
			110
	Without calculations	answers incorrect	0*0
			1*0
		answers correct	1*1
			0*1
Answer only responses		answers correct	**1
		answers incorrect	**0

Table 10: The 3-digit codes that emerged from learners responses across all three test sittings

In short, based on the 3-digit codes that emerged from coding learner responses across all three test sittings, learner responses were split between ‘answer-only’ responses (i.e. responses showing no evidence of either an interpretation or a calculation) and ‘modelled’ responses (i.e. responses that included an interpretive model). Whilst there were two possible outcomes where ‘answer-only’ responses were concerned, modelled responses reflected one of ten 3-digit codes. Moreover, ‘modelled’ responses either implicated a calculation or not, and, in both cases, either yielded correct or incorrect answers. Overall, modelled responses could be associated with correct answers in half of the 10 modelled responses codes.

Additionally, each modelled response was analyzed for coordination of structural (interpretation of the situation) reasoning and numerical (calculation of the solution) reasoning. In general, a response including a model configured to reflect either an inappropriate interpretation and/or an inappropriate calculation was deemed to reflect inappropriate coordination of structural and numerical reasoning on the part of the respondent. For example, the 3-digit code for a learner response showing a correct interpretation of the problem situation into a model, followed by an incorrect calculation attempt and an incorrect answer would be ‘100’. Following the same coding scheme, a learner response showing a correct interpretation of the problem situation into a model, followed by an incorrect calculation and the correct answer was allocated a ‘101’ code. More importantly, because both codes have 0’s in the place associated with the calculation phase, I concluded that the structural and numerical reasoning associated with each case was inappropriately coordinated. Similarly, regardless of the appropriateness of the calculation and correctness of the answer, responses coded with a 0 for the interpretation phase were also deemed to reflect inappropriate coordination in the requisite structural and numerical reasoning[5]. This includes response coded 000 because, for the reasoning to be deemed appropriately coordinated, we must first have that the interpretation was appropriate to begin with. The rule of thumb, therefore, was that any response coded with a 0 in either or both of the first two digits of the 3-digit code, would be deemed to reflect inappropriately coordinated reasoning on the part of the learner concerned. In particular, any response coded with 0 in the interpretation phase was necessarily read as reflecting inappropriately coordinated structural and numerical reasoning, and only learner responses showing a correct interpretation of the problem situation into a model, a correct calculation and an incorrect answer (110 code) or correct answer (111 code) were regarded as responses with appropriate coordination of the two aspects of reasoning. In the absence of evidence of appropriate interpretation and/or calculations, an inference of appropriately co-ordinated mental structural and numerical reasoning was applied to responses reflecting the 1*1 and the **1 codes as well. These analytical tools provided the overarching lenses for characterizing shifts across the range of additive relations problem classes and types that learners encountered in the pre-, post and delayed post-test. With the macro analysis reported in chapter 5, it is this protocol for coding of learners responses that enabled the answering of my second research question.

In summary, from a developmental point of view, the extent to which learners appropriate the additive relationship as a mental object and transform it in a manner that allows for the most efficient calculation strategy to be used is a function of the extent to which their use of models reflects appropriate coordination of structural and numerical reasoning. By extension, such an

[5] Follow-through responses based on the inappropriately configured model of the situation were marked incorrect.

outcome would provide the basis upon which this intervention would be adjudged to have been successful.

3.7 Ethical considerations

Under the auspices of the broader project, informed consent was sought and granted by the principals of the schools, the teachers and the parents on behalf of the learners involved in the project. Herein the right to privacy of the learners was assured. Also, the participants were guaranteed protection from harm, and the researchers took the necessary precautions to guard against treating participants as objects. On my part I approached the principal, teacher, parents and learners in the Grade 3 class I worked with, sharing information on the study and gaining informed consent regarding the intervention identified in my study. Pseudonyms are used within the findings and analysis chapters that follow. Ethical clearance for the study was also granted from the university.

3.8 Reliability, objectivity and limitations

Given that this study was designed to examine the contributions to additive relations problem solving competence that can be effected on a single group of learners through an intervention that focuses on the use of the bar diagram (BD) and the empty number line (ENL) as relational models, this study can best be describe as a teaching experiment. One limitation of this pre and post-test single-group design is that it does not provide a sufficient basis of comparison for us to unambiguously ascribe the shifts to the intervention, and hence no generalisations will hold as this study was intended as an exploration on the uptake of the BD and ENL as models. Ideally, there should have been a control group against which I could compare the results of the experiment group. Given that my study was exploratory in its design – aimed to look at the feasibility of promoting the use of relational models – the lack of a control group did not centrally impact on my gaining useful insights from the study.

Given my dual role as researcher and teacher I had to take a few precautions in order not to compromise the objectivity of the analysis and via this, the results. Firstly, the theoretical and analytical frameworks that I used were informed by research, thus keeping personal inferences at bay. This is in line with the design-based research methods in which designs are based on principles drawn from theory, the literature in the field, and prior research (Tabak, 2004). Secondly, during the process of designing the test and analysing the data, my approaches and emerging findings were shared with and checked by the broader WMC-P team and my supervisor for interrater reliability (Cohen, Manion, & Morrison, 2013). Finally, I had the teacher whose class I was working with to be present when the tests were administered in order to keep a close eye, not only on the learners, but also on the time allocated to each problem in the test. All of these precautions provided routes for keeping some 'distance' from the data, and ensuring consistent application of the analytical lenses.

In conclusion, this study was designed and implemented within the Relational Paradigm. Whilst building on the Equilibrated Development Approach (EDA) formulated by Polotskaia and colleagues, this study was adapted to incorporate explicit instructional focus on BD and ENL as key relational models, as well as to include dedicated time at the beginning of the lesson to focus explicitly on introducing and discussing combinations of 10 and 20 in BD and ENL formats. Ultimately, the exploration of this study was into the extent to which the BD and ENL combination can serve as a mental tool for supporting the coordination of the interpretation and calculation of additive word problems in a relational model-based approach that can support the solving of additive word problems with structural and numerical sense.

In the chapter that follows I analyse the shifts learners' performance from pre-test to post-test to delayed post-test.

CHAPTER 4

EVALUATION PHASE ANALYSIS OF PERFORMANCE

4.1 Introduction

In this chapter I present findings relating to the evaluation phase of the problem-solving cycle outlined in my literature base and theoretical orientation chapter. For reasons already outlined in that chapter, I have truncated the evaluation phase into just looking at answers. Consequently, the data analysed in this chapter are the answers that the learners provided to the items in the pre-, post- and delayed post-test. Whilst the phases of problem-solving were described in the order of the problem-solving cycle (interpretation, calculation and evaluation), I begin my analysis here with a primary focus on the evaluation phase, because, from a research point of view, it made sense to me to analyse the patterned differences relating to the interpretation and calculation phases following evidence of shifts in performance, whether positive or negative. In this way, positive shifts in performance would point to the success of the intervention, which improvements can then be interrogated for localization. Ultimately, as a teacher and someone interested in development overall, my primary interest is in improving performance.

At the evaluation level, where the data source is simply the answers, two analyses are presented. These analyses relate to a range of features related to word problems that I flagged as salient in the literature chapter. The first analysis relates to overall performance as seen purely from the point of view of the number and related percentage of correct answers provided by the 39 learners to the 12 items in the pre-test. These numbers and percentages provide the grounds for the second analysis in which the focus is on the shifts in the 39 learners' performance on the 12 test items from pre- to delayed post-test, with a particular focus on differences in learner performance on individual items. Thus the purpose of this second analysis is to track and delineate the nature of the shifts in performance on each of these items from pre- to delayed post-test. These analyses demonstrate increases in the proportions of learners who got the items right from pre- to post to delayed post-test.

With the largest gains associated with the most 'difficult' problems, the findings point to how, following the intervention, the gap between learner performance on 'easy' and 'difficult' items narrowed. This finding of a narrowing of the performance gap between 'easy' and 'difficult' items is important as it points to the salience of Polotskaia's (2017) hypothesis that working relationally allows learners to move towards being as effective in solving 'difficult' problem as in solving 'easy' problems. Moreover, as part of the second analysis, I was able to compare and contrast the empirical ranking order of performance on individual items (relevant to this cohort of learners) with the expected performance on the items based on the difficulty levels theorised by research. This analysis reveals how working relationally brought learners' empirical performance on these items into closer alignment with the 'levels of difficulty' hypothesised by Nesher et al. (1982) over time.

4.2 First analysis: overall shifts in learner performance across pre-, post- and delayed post-test.

As summarised in table 11, before the intervention was rolled out, only 134/468 answers provided by the learners on the pre-test were correct.

The tests	**Number and percentage of correct answers in each of the three tests**		**Improvements in the number and proportion of correct answers comparing each of post- and delayed post-test to the pre-test results**	
	Total number of correct answers	Percentage	Increase in total number of correct answers	Increase in proportion of correct answers
Pre-test	134	29%		
Post-Test	173	37%	39	8%
DPost-test	229	49%	95	20%

Table 11: Learner performance and improvement in the pre-, post- and delayed post-test

With an increase in the proportion of correct answers provided by the learners in each test from pre- to post- to delayed post-test (8% points and 20% points increases respectively), the delayed post-test result points to a retention of learning linked to the intervention two months beyond the end of the intervention.

This improvement of 20 percentage points in performance means that the number of correct answers provided by the learners increased, on average, by 7 to 8 correct answers per item in the delayed post-test as compared to the pre-test. What this average overall improvement in performance masks though, are differences in the shifts in performance across items, with larger increases in the number of correct answers on some items than in others. Consequently, disaggregating shifts in performance on individual items constituted the primary engagement in the second analysis.

The focus in the analysis that follows is, therefore, on learners' initial performance on each of the 12 items in the pre-test – that is prior to the rolling out of intervention lessons – with an interest in the shifts associated with these items following intervention lessons. In particular, my interest was in the extent to which learner performance on items predicted to be difficult shifted relative to those predicted to be easy. In this regard, the literature based 'levels of difficulty' hypothesised by Nesher et al. (1982) are of interest because they provide this study with hierarchies that can be used to judge if the intervention was successful or not. The ways in which these hierarchies provide this study with lenses for adjudicating the success of this intervention are detailed below.

4.3 Second analysis: disaggregating the overall improvement in performance in terms of differences in performance on individual items in the tests

As highlighted in the literature review chapter (chapter 2), my selection of the 12 word problem test items was predicated on the 'levels of difficulty' established by literature and the ways in which this literature base provided evidence of further aspects of complexity and progression within the design of items. My study was thus provided with additional lenses for interpreting differences in learners' responses across the pre and post-tests. As elaborated in chapter 2, researchers are in agreement that the semantic category to which an additive word problem belongs – i.e. change, combine or compare – and the nature (position) of the unknown determine the type and, by extension, the 'level of difficulty' associated with it. Lenses for viewing the overall improvement in performance in terms of performance on individual items are, therefore, availed by the literature-based 'levels of difficulty' associated with different types of additive word problems.

4.3.1 Literature-based levels of difficulty

Nesher et al. (1982) established a 'scale of relative difficulty' which enables me to theoretically distinguish between different additive word problem types according to their 'level of difficulty'. They postulate four levels of difficulty which can be used to predict learners' performances on additive word problems. Based on the semantic category to which the problems belong and the nature of the unknown, Nesher et al. (1982) allocate levels of difficulty to different types of additive word problems.

Repeated in table 12 are the 12 additive word problems in the tests, their types (semantic category and nature of the unknown), and 'levels of difficulty' in line with the 'scale of relative difficulty' established by Nesher et al. (1982). As detailed in chapters 2 and 3, problems on level 1 of the 'scale of relative difficulty' are Change and Combine problems in which the unknown is the result and the whole respectively; problems on level 2 are change increase and change decrease problems in which the question is about the change; problems on level 3 are Change problems in which the question is about the initial quantity, Combine problems in which the difference or the compared set is the unknown. Not included in the study were problems on level 4 of the 'scale of relative difficulty', which are Compare problem in which the unknown is the referent set.

Q#	Word Problem	Direct Model	Semantic category	Nature of the unknown	Nesher et al.'s (1982) levels of difficulty
1	On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether?	17 + 23 =[]	Change increase	Result unknown	**1**
2	Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether?	34 + 19 = []	Combine	Whole unknown	**1**
3	52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether?	52 + 49 = []	Combine	Whole unknown	**1**
11	Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick?	[] = 7 + 19	Combine	Whole unknown	**1**
4	Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost?	40 - [] = 32	Change decrease	Change unknown	**2**
7	James had 8 marbles. John gives him some more marbles. James now has 77 marbles. How many marbles did John give to James?	8 + [] = 77	Change increase	Change unknown	**2**
5	Chris gave away 19 stickers and was left with 34 stickers. How many stickers did Chris have at the beginning?	[] - 19 = 34	Change decrease	Start unknown	**3**
8	There were 40 people at the park. 17 were adults and the others were children. How many children were at the park?	40= 17 + []	Combine	Part unknown	**3**
10	18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls?	18 + [] = 40	Combine	Part unknown	**3**
12	There were 77 people on the school bus during the trip. 69 of them were children and the others were teachers. How many of them were teachers?	77 = 69 + []	Combine	Part unknown	**3**
6	69 children travel by bus. 77 children travel by train. How many more children travel by train?	69 + [] = 77	Compare	Difference unknown	**3**
9	Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles?	53 = 19 + []	Compare	Difference unknown	**3**

Table 12: The 12 additive word problems in the tests, their semantic categories and nature of unknown, as well as their 'levels of difficulty' according to Nesher et al.'s (1982)

The analysis of learners' performance on a set of items – item analysis – is usually used to examine the performance of items with the intent of improving tests. I have opted to use it a little differently, to gather evidence regarding learners' knowledge of a particular topic (Bai & Ola, 2017), in this case, their ability to solve different types of additive word problems. Cast in this light, I analyze learners' performance on the items in the tests based on the proportion of learners that got an item correct. Identifying initial learner performance on the 12 items in the pre-test then allows for distinguishing between this particular cohort of learners' performance on items that literature identifies as 'easy' from their performance on items that literature identifies as 'difficult', before intervention lessons could be enacted.

Going forward, learners' initial performance on the 12 items in the pre-test establishes a baseline upon which I analyse and comment on the shifts in performance on the 12 items from pre- to post- to delayed post-test. This way I am able to determine the extent to which learner performance improved on the 'difficult' relative to the 'easy' items following the intervention. This analysis answers my first research sub-question:

What were the overall patterns of learner performance across different types of additive relations problems prior to the intervention, and what shifts in performance were observed from pre- to delayed post-test?

As already highlighted in my review of the literature on additive relations problem-solving, the principled position taken within the Relational Paradigm is that working relationally allows learners to move towards being as effective in solving the theoretically 'difficult' problems as in solving the theoretically 'easy' problems (Polotskaia, 2017). With this study framed within the Relational Paradigm, part of my interest was the extent to which the evidence supports this conjecture.

With a view to determining the extent to which the evidence supports this conjecture, I proceed by looking at performance in the pre-test from the point of view of the number (proportions) of learners that got the individual items right, as well as the differences in performance on the items. I then turn to tracking the shifts in performance with a keen eye on the extent to which the differences in performance on the items diminished on these items from pre- to delayed post-test.

4.3.2 Item by item analysis of learners' performance in the pre-test

Table 13 provides the basis for an item by item analysis of learners' performance in the pre-test. I have placed in this table the items in order, from highest to lowest performance in order to analyse the extent to which more 'difficult' items (based on the Nesher et al.'s (1982) classification) were associated with low performance and 'easy' items with high performance. The distribution of correct answers across items were placed into three initial performance levels, in order to provide an exploratory lens through which changes in performance could be analysed across the three test administrations. The performance levels were developed to provide some spread of items and are:

Item	Problem category	Nesher et al.'s (1982) difficulty level	Pre-test: Ordered empirical performance: # (%)	Pre-test empirical performance category
Q1 17 + 23 =[]	Change-RU	1	29 (74%)	**High initial performance (HIP) items**: [2 items]
Q11 [] = 7 + 19	Comb-WU	1	21 (54%)	
Q2 34 + 19 = []	Comb-WU	1	16 (41%)	**Mid initial performance (MIP) items**: [5 items]
Q4 40 – [] = 32	Change-CU	2	16 (41%)	
Q10 18 + [] = 40	Comb-PU	3	13 (33%)	
Q5 [] - 19 = 34	Change-SU	3	13 (33%)	
Q3 52 + 49 = []	Comb-WU	1	11 (28%)	
Q12 77 = 69 + []	Comb-PU	3	6 (15%)	**Low initial performance (LIP) items**: [5 items]
Q8 40= 17 + []	Comb-PU	3	4 (10%)	
Q7 8 + [] = 77	Change-CU	2	3 (8%)	
Q6 69 + [] = 77	Comp-PU	3	1 (3%)	
Q9 53 = 19 + []	Comp-PU	3	1 (3%)	

Table 13: Correct answers spread across three initial performance levels in the pre-test

A striking feature of this item based distribution in performance was that of large gaps across and within categories. In particular, there was a gap of exactly 8 correct answers between the two HIP items with 29 (74%) and 21 (54%) of the 39 learners getting Qs 1 & 11 right respectively. Also, whilst Q1 was the highest initial performance item with 29 (74%) of the 39 learners getting it right, at the other extreme, only 1 learner provided correct answers to each of Qs 6 & 9, which made Qs 6 & 9 the lowest initial performance items in the pre-test. With Qs 6 & 9 accounting for only two of the total number of correct answers and the other three LIP items (Qs 12, 8, & 7) accounting for 6, 4 and 3 correct answers in the pre-test respectively, Q1 accounted for more correct answers than all five LIP items put together.

Q1 was a Change Increase Result Unknown (Change-RU) problem and Q11 was a Combine Whole Unknown (Comb-WU) problem, both of which are Level 1 problems according to Nesher et al.'s (1982) 'scale of relative difficulty' (1 being the easiest). As table 13 shows, Qs 11 & 1 were, both theoretically and empirically, the two easiest items on the test. The performances on the two other Level 1 items in the test had them placed highest (Q2) and lowest (Q3) in the ranking order of the five MIP items respectively. With 16 (41%) learners getting it right, Q2 had the third highest initial performance in the pre-test, after the two HIP items (Q1 and Q11). Thus, as predicted by Nesher et al.'s (1982) 'scale of relative difficulty', the top three items in the pre-test were all Level 1 items. However, with only 11 (28%) of the 39 learners getting it right, performance on Q3 was associated with almost half the performance associated with Q11. As a result, Q3 was placed lower in the initial performance ranking order (for this cohort of learners) than three MIP items that Nesher et al.'s (1982) 'scale of relative difficulty' would designate as Level 2 and Level 3 problems.

- High initial performance (HIP) items: above 50% of the 39 learners got the correct answer [2 items: Qs 11 & 1]
- Mid initial performance (MIP) items: between 25% and 50% of the 39 learners got the correct answer [5 items: Qs 2, 4, 10, 5 and 3]
- Low initial performance (LIP) items: less than 25% of the 39 learners got the correct answer [5 items: Qs 12, 8, 7, 6 and 9]

The three MIP items with associated performances that were higher than that of Q3 were two Change problems (Qs 4 & 5) and one Combine problem (Q10). To be precise, Q4 was a Change Decrease Change Unknown (Change-CU) problem, Q5 was a Change Decrease Start Unknown (Change-SU) problem, and Q10 was a Combine Part Unknown (Comb-PU) problem. Nesher et al. (1982) found decreasing an initial given quantity by some amount to be less challenging for learners than decreasing an unknown initially quantity by a given amount. In fact, they found that Change problems in which the initial quantity was unknown 'were difficult at all grade levels [one to six]' (p.376). It makes sense, therefore, that Q4, which is a change unknown type of Change problem, should yield an initial performance that is higher than that of the start unknown type of Change problem (Q5). In fact, with 16 (41%) of the 39 learners getting Q4 right, the initial performance on Q4 was as high as the performance on Q2, a Level 1 Combine Whole Unknown problem which came in third in the initial performance ranking order.

Incidentally, according to Nesher et al.'s (1982) 'scale of relative difficulty', the Change Decrease Start Unknown (Change-SU) problem (Q5), and the Combine Part Unknown problem (Comb-PU) (Q10), are both Level 3 problems. Not surprisingly then, they were both associated with a 33% initial performance in the pre-test. Based on Nesher et al.'s (1982) 'scale of relative difficulty', one would have expected the initial performance on Q7 (Change-CU) to be higher than the performance on Q4 (Change-SU). At the very least, one would have expected to find Q7 amongst the MIP items. Instead, only 3 (8%) learners got Q7 right in the pre-test, making it the third item from the bottom of the initial performance ranking order.

The unexpected positions occupied by Qs 3 & 7 can be explained in terms of the relative sizes of the quantities involved. It would appear that number range is playing through as a variable that tends to be absent in Nesher et al.'s (1982) categorizations because they tend to keep number range relatively constant within and across the semantic category/nature of the unknown. In particular, unlike the other three Level 1 problems in the test, Q3 was the only one with a sum exceeding 100. Similarly, whilst the difference between the quantities involved in the only other Change Unknown problem (Q4) was small enough to be counted out (40 – [ ] = 32), the direct model for Q7 would be 8 + [ ] = 77, which tends to render the missing addend too large a number to be counted out.

In summary then, whilst MIP items were a mixed bag, the HIP items were either of the whole unknown or the result unknown types, and the LIP items were all of the part unknown or the change unknown types. In general, the HIP items required adding the quantities and the LIP items required one quantity to be subtracted from the other. Essentially then, the HIP category matched the theoretical position, but the distributions in the MIP and LIP categories indicated other factors at play beyond semantic category and nature of the unknown.

Having established learners' initial performance on individual items and contrasted this initial performance on individual items with the levels of difficulty hypothesised by Nesher et al. (1982), I now turn my attention to learner performance in the delayed post-test with a view to quantifying the shifts in performance associated with the individual test items from pre- to delayed post-test.

4.3.3 Item by item analysis of learners' performance in the delayed post-test and shifts in performance from pre- to delayed post-test

To compare the initial and the final placements of the test items in the pre- and delayed post-test, table 14 ranks the 12 test items in order from highest to lowest performance. For the performance in the delayed post-test to be compared with the performance in the pre-test, the cut-off points of 25% between low and mid final performance items, and 50% between mid and high final performance items were maintained. The distribution of performance in the delayed post-test on individual items was such that there were 5 items for which more than 50% of learners got right (designated high final performance (HFP) items), and 7 items (designated mid final performance (MFP) items) that between 25% and 50% of learners got right.

The first thing worth noting in the distribution of performance in the delayed post-test is the shift from single digits to teen numbers of correct answers on all the items on which learner initial performance was in the low initial performance band (Qs 12, 8, 7, 6 and 9): this resulted in these items now being classified as located within the mid performance band. Secondly, with the number of learners getting Qs 5, 4 and 10 right remaining in the teens, these items remained in the mid performance band. Thirdly, learners' performance on Qs 2 & 3 meant these items shifted from the MIP to the High Final Performance (HFP) band (following 9 and 12 more correct answers in the delayed post-test respectively).

Pre-test						Delayed post-Test			
Item	**Problem category**	**Nesher et al.'s (1982) difficulty level**	**Pre-test: Ordered empirical performance: # (%)**	**Pre-test empirical performance category**		**Item**	**Nesher et al.'s (1982) difficulty level**	**Delayed post-test: Ordered empirical performance: # (%)**	**Delayed post-test empirical performance category**
Q1 17 + 23 =[ ]	Change-RU	1	29 (74%)	**High initial performance items**: [2 items]		**Q1** 17 + 23 =[ ]	1	29 (74%)	**High final performance items:** [4 items]
Q11 [ ] = 7 + 19	Comb-WU	1	21 (54%)			**Q11** [ ] = 7 + 19	1	29 (74%)	
Q2 34 + 19 = [ ]	Comb-WU	1	16 (41%)	**Mid initial performance items**: [5 items]		**Q2** 34 + 19 = [ ]	1	25 (64%)	
Q4 40 – [ ] = 32	Change-CU	2	16 (41%)			**Q3** 52 + 49 = [ ]	1	23 (59%)	
Q10 18 + [ ] = 40	Comb-PU	3	13 (33%)			**Q7** 8 + [ ] = 77	2	19 (49%)	**Mid final performance items:** [8 items]
Q5 [ ] - 19 = 34	Change-SU	3	13 (33%)			**Q4** 40 – [ ] = 32	2	18 (46%)	
Q3 52 + 49 = [ ]	Comb-WU	1	11 (28%)			**Q5** [ ] - 19 = 34	3	18 (46%)	
Q12 77 = 69 + [ ]	Comb-PU	3	6 (15%)	**Low initial performance items**: [5 items]		**Q12** 77 = 69 + [ ]	3	15 (38%)	
Q8 40= 17 + [ ]	Comb-PU	3	4 (10%)			**Q6** 69 + [ ] = 77	3	15 (38%)	
Q7 8 + [ ] = 77	Change-CU	2	3 (8%)			**Q9** 53 = 19 + [ ]	3	14 (36%)	
Q6 69 + [ ] = 77	Comp-PU	3	1 (3%)			**Q10** 18 + [ ] = 40	3	13 (33%)	
Q9 53 = 19 + [ ]	Comp-PU	3	1 (3%)			**Q8** 40= 17 + [ ]	3	11 (28%)	

Table 14: Correct answers spread across two performance levels in the delayed post-test

The net result of these shifts was the presence of 8 items in the MFP band. Interestingly, Q7 moved from being in the middle of the pre-test LIP band to being top of MFP band (following 16 more correct answers in the delayed post-test). Also, whilst in the pre-test there were only two items in the high performance band (Qs 1 & 11), there were four items (Qs 11, 1, 3 & 2) for which more than 50% of learners got the answers right in the delayed post-test. An overall result of these shifts was that there were no longer any items for which 25% of the learners or less got right, meaning that the performance associated with all of the 12 items in the test were either high or mid-performance. Moreover, the large gaps in performance that were evident between and within bands of items in the pre-test were significantly narrower in the delayed post-test. In particular, the difference in performance within the HIP band in the pre-test was such that 8 more learners got Q1 right than Q11. With the performance on Q1 remaining unchanged from pre-test to delayed post-test and 8 more learners providing the correct answer to Q11 in the delayed post-test (as shown in table 15 which records the number of correct answers for each question), these two items were now on par with 29 correct answers each.

Low initial performance, **Mid** final performance items				**Mid** initial performance, **Mid** final performance items				**Mid** initial performance, **High** final performance items				**High** initial performance, **High** final performance items			
Item	Level	Pre-	DPost	**Item**	Level	Pre-	DPost	**Item**	Level	Pre-	DPost	**Item**	Level	Pre-	DPost
Q7	2	3	19	**Q5**	3	13	18	**Q3**	1	11	23	**Q11**	1	21	29
Q6	3	1	15	**Q4**	2	16	18	**Q2**	1	16	25	**Q1**	1	29	29
Q9	3	1	14	**Q10**	3	13	13								
Q12	3	6	15												
Q8	3	4	11												

Table 15: Summary of shifts between bands of the 12 items in the test from pre- to delayed post test

Additionally, not only was the gap between the high and the mid performance bands reduced, but overall, the difference in correct answers between any two consecutive items in the final performance ranking order was not more than 4. At the other extreme, whilst only 1 (2.56%) learner had provided the correct answer to each of the lowest initial performance items (Qs 6 & 9), the lowest final performance item was Q8 with 11 correct answers. Consequently, whilst Q7 was the item associated with the most improvement from pre-test to delayed post-test, it was followed closely by the lowest initial performance items (Qs 6 & 9) with 14 (36%) more and 13 (33%) more correct answers in the delayed post-test respectively. Also, with 29 (74%)

of the 39 learners getting the highest final performance items (Qs 1 & 11) right, the gap between the number of learners who got the highest performance items right and the number of learners who got the lowest performance item (Q8) right was reduced from 28 in the pre-test to 19 in the delayed post-test. This amounts to a contraction of 23 percentage points in the performance range.

In summarising then, the overall effect of the shifts outlined above was, twofold. Firstly, whilst in the pre-test there was a mismatch in the ordering with a Level 1 item (Q3) placed below a Level 2 item (Q4) and the latter placed below Level 3 items (Qs 10 & 5), not only were the whole and result unknown Level 1 items (Qs 1, 11, 2 & 3) placed in the high final performance band in the delayed post-test, but they were now all placed at the top of the performance ranking order, followed neatly by the two change unknown Level 2 items (Qs 7 & 4), and then by the six Level 3 items (Qs 5, 12, 6, 9, 10 & 8). In other words, the performance on individual items in the delayed post-test reveals an ordering that more closely reflects the ordering predicted by the 'scale of difficulty' proposed by Nesher et al. (1982). Essentially the empirical ranking order came into alignment with the theoretical ranking order.

Secondly, and perhaps unexpectedly in a South African context with evidence of low performance on more complex items, with more correct answers in the delayed post-test as compared to the pre-test there came a compression in the spread of the number of learners who provided correct answers from three performance levels – high, mid and low – in the pre-test, to the exclusion of low performance items, as defined in this study, in the delayed post-test. This compression in the spread of the number of correct answers across the items in the delayed post-test, and, in particular, the absence of items with less than 25% of the learners getting the answers right, points to a narrowing of the gap between initially high and initially low performance items.

Finally, considering the increase in performance of 33% or more where the Compare problems were concerned – which, in requiring learners to make sense of the relation between the two quantities involved are the most difficult according to Nesher et al.'s (1982) 'scale of difficulty' – this narrowing of the gap between initially high and initially low performance items was, in effect, a narrowing of the gap between 'difficult' and 'easy' problems. The explanation provided within the Relational Paradigm would be that learners are dealing better with the relational structure of the problems; meaning that they got better at interpreting the additive relationship following the intervention lessons. In the chapter that follows, I interrogate the extent to which this was actually the case.

CHAPTER 5

INTERPRETATION AND CALCULATION PHASE ANALYSES

5.1 Introduction

Data on the overall number of correct answers presented within the evaluation phase analysis in Chapter 4 revealed increases in the proportions of correct answers in learner responses across the three test sittings: from 29% in the pre-test to 37% in the post-test to 49% in the delayed post-test (with each test based on the responses for n=39 learners on each of 12 test items). As reported in Chapter 4, with these increases in the proportions of correct answers in learner responses, there was better alignment with the hypothetical or theoretical hierarchy of item difficulty, with the theoretically hardest Compare problems (Qs 6 & 9) amongst the items showing the largest improvements. These improvements in evaluation phase performance (where the data source is simply the answers) lent interest to understanding whether, and – if so – the extent to which these improvements were associated with shifts in the interpretation phase (with structural reasoning at work in this phase) and/or the calculation phase (with numerical reasoning at work in this phase). As flagged in the literature review and theory sections, structural reasoning refers to the interpretation of the relationships involved in the problem situation into models, and numerical reasoning refers to the calculation of the solution to the problem in light of the model.

My overarching interest was in the nature of the shifts, firstly, within learners' interpretations of situations, and secondly in learners' co-ordination of interpretation and calculation. In each case, I sought to first analyse these shifts at an overview level across the three test sittings, and then secondly, at the level of the three sub-categories of low, mid and high initial performance (HIP, MIP and LIP) that were devised in the evaluation phase analysis chapter. These analyses were guided both by evidence in South Africa relating to learners challenges with interpretation (sense-making), and by an international literature base pointing to, on the one hand, the importance of interpreting the structural relationships between the quantities in additive situations into models (Polotskaia, 2017), and on the other, the requirement for these appropriate interpretations to be complemented with appropriate calculations if learners' engagements with additive word problem are to be successful (Giroux & Ste-Marie, 2001).

Given the attention in the intervention to working with the bar model to support structural reasoning and the number line as a model that can support numerical reasoning, of particular interest within this analysis was understanding the extent to which there was a shift towards modelled responses (using bar, or other models) with calculations, and the success associated with this shift. To this end, I begin by providing, in the next two sections, overview analyses of:

- the shifting balance of answer-only and modelled responses across pre-, post- and delayed post-tests, and how this interplayed with correct and incorrect answers (in section 5.2). This analysis shows how the improvement in performance reported on in chapter 5 was underpinned by a tilt away from the predominance of answer-only responses in the pre-test to a much greater incidence of modelled responses in the delayed post-test
- the contrasting patterns of interplay between modelled-responses-with-calculations and modelled-responses-without-calculations, with correct and incorrect answers. This shift occurred amidst a declining occurrence and influence of answer-only responses (in section 5.3). This analysis showed that a larger proportion of modelled-responses-with-calculations[6] was associated with correct answers in contrast to modelled-responses-without-calculations. This followed a growing willingness on the part of the learners to include written calculations of the solutions.

The finding of an increased willingness by learners, not only to interpret problem situations into models, but also to perform written calculations, provided the grounds for analyses of shifts in learner competence within each of the three sub-categories of HIP, MIP and LIP items. The overview findings led to attention to the interpretation phase, and to co-ordination between the interpretation phase and the calculation phase in learner responses in order to understand how patterns of shift played out across each of the item sub-categories.

[6] In turn, this finding provided further grounds for zooming in on learners' calculations in the calculation phase analysis in the next chapter.

On this account, the following two questions frame the analyses of the sub-categories in this chapter:

- *What is the nature of the shifts in interpretation across the responses in each sub-category?*
- *What is the nature of the shifts in learners' co-ordination of interpretation and calculation across the responses in each sub-category?*

Given that the aim of the intervention was to promote the use of models to support the interpretation and calculation and thus support the production of correct answers to problems in the evaluation phase, I coded for the extent of appropriate response to each question in terms of interpretation, calculation and evaluation phases, producing 'three-digit' codes for every single learner response. This allowed for a delineation of the nature of the shifts in learners' interpretations of situations and co-ordination of interpretation and calculation. For the interpretation and calculation phases, three coding options were possible: *: no evidence; 0: incorrect; 1: correct. For the Evaluation phase, only two coding options were needed to deal with the empirical base: 0: incorrect; 1: correct. A 100 coding, for example, indicated a response with a model that interpreted the situation appropriately, but with an incorrect calculation and an incorrect answer. A 1*1 coding reflected a response with an appropriate model, no evidence of a calculation but with a correct answer. Differences in the patterns of occurrence of these codes across the three test sittings were first studied with reference to the interpretation Phase.

Data summaries leading to the three-digit analyses now follow.

5.2 Shifts between answer-only responses and modelled responses across pre-, post- and delayed post-tests

As reported on in chapter 4, with each test based on the responses for n=39 learners on each of 12 test items, the proportion of correct answers increased from 29% in the pre-test, to 37% in the post-test, to 49% in the delayed post-test. The following analysis reveals that, with learners' acquaintance with the bar model and the empty number line during intervention lessons and a growing willingness to interpret problem situations into these models, a greater proportion of modelled learner responses were associated with correct answers.

The distribution of correct answers ('success') across the modelled-answer and answer-only categories in the three test administrations can be considered in different ways. Table 16 shows the number and proportions of modelled and answer-only responses, and the extent of success in terms of correct answers among these groups. In table 16, percentages of correct answers are proportions of success of each solution type, with this result showing that, in the pre-test, the larger share of correct answers came from answer-only responses, reflecting the large base of answer-only responses in the pre-test pool. In particular, table 16 shows that 351/468 (75%) of the responses were answer-only responses, and 117/468 (25%) were modelled responses. In the post-test, the number of answer-only responses declined to 66, and the number of modelled responses escalated to 402. In the delayed post-test, answer-only responses further declined to 35 whilst modelled responses further increased to 433. Thus, the proportion of modelled responses increased from 25% of the 468 responses in the pre-test to 93% of the 468 responses in the delayed post-test.

	Pre-test		**Post-test**		**Delayed Post-test**	
	Extent	**Success**	**Extent**	**Success**	**Extent**	**Success**
Number (proportions) of modelled responses	117 (25% of 468)	45 (38% of 117)	402 (86% of 468)	158 (39% of 402)	433 (93% of 468)	218 (50% of 433)
Number (proportions) of answer-only responses	351 (75% of 468)	89 (25% of 351)	66 (14% of 468)	15 (23% of 66)	35 (7% of 468)	11 (31% of 35)
Totals	**468**	**134**	**468**	**173**	**468**	**229**

Table 16: Prevalence of answer-only responses/modelled responses and their feed-in to correct answers from pre- to post- to delayed post-test

The percentage of correct solutions from the total of each type of solution show the relative efficiency of the type pf solution for students. The level of success for learners using modelling is going up from 38% to 39% to 50%. The efficiency of non-model solutions is also improved from 25% to 23% to 31%.

Alternatively considered, the 89 correct answers among the 351 answer-only responses in the pre-test, constituted a 25% proportion of correct answers associated with answer-only responses. This was lower than the proportion of correct answers associated with modelled responses (45/117 = 38%), pointing to greater 'success' rates in the modelled response category. This gap had widened by the time of the delayed post-test, in spite of increased success rates in both categories: the proportion of correct answers associated with answer-only responses was 11/35 (31%), while the proportion of correct answers associated with modelled responses was 218/433 (50%).

With the number of answer-only responses reduced to 66 in the post-test, the contribution made by correct answer–only responses also declined to 15, which amounted to 9% of the 173 correct answers. In the delayed post-test, the contribution made by correct answer–only responses to the performance in the delayed post-test further declined to only 5% of the 229 correct answers. Contrariwise, the contributions made by modelled responses to the performance in each of the three tests increased from 45 (34%) in the pre-test, to 158 (91%) in the post-test, to 218 (95%) in the delayed post-test. From one test to the next therefore, modelled responses contributed substantially increasing proportions of the correct answers.

The dominance of answer-only responses in the pre-test reveals, not only an absence of models of situations in learners' responses, but also an absence of calculations of solutions. Given the increased evidence of learners interpreting problem situations into models amidst a declining prevalence and influence of answer-only responses, I was interested in finding out the extent to which learners were also willing to perform a written calculation, and the influence that responses reflecting a model of the situation AND a calculation of the solution had on the improvement in performance.

5.3 Contrasting modelled-responses-with-calculations and modelled-responses-without-calculations

The significance of disaggregating modelled-responses-with-calculations from modelled-responses-without-calculations can be seen in the distribution in table 17. Amidst the declining prevalence and influence of answer-only responses, this distribution shows increases in the prevalence of both modelled-responses-with-calculations (99 - pre-test; 369 - post-test; 353 - delayed post-test) and modelled-responses-without-calculations (18 - pre-test; 33 - post-test; 80 - delayed post-test). However, by the end, modelled-responses-without-calculations contributed only 23/229 (10%) of the correct answers, contrasting sharply with the contribution of 195/229 (85%) of the correct answers in the delayed post-test from modelled-responses-with-calculations. By disaggregating modelled-responses-with-calculations from modelled-responses-without-calculations, not only was I able to observe that, in general, modelled-responses-with-calculations accounted for the overwhelming majority of correct answers in the post-tests, but I was now able to actually quantify the extent of the tilt in balance. As seen in table 17, the tilt in balance was away from 75% answer-only responses in the pre-test to 75% modelled-responses-with-calculations in the delayed post-test.

Type of response	With/without calculation	Pre-test		Post-test		Delayed post-Test	
		Extent	Success	Extent	Success	Extent	Success
Modelled responses	With calculations	99 (21% of 468)	40 (40% of 99)	369 (79% of 468)	155 (42% of 369)	353 (75% of 468)	195 (55% of 353)
	Without calculations	18 (4% of 468)	5 (28% of 18)	33 (7% of 468)	3 (9% of 33)	80 (17% of 468)	23 (29% of 80)
Answer-only responses		351 (75% of 468)	89 (25% of 351)	66 (14% of 468)	15 (23% of 66)	35 (8% of 468)	11 (31% of 35)
Totals		468	134	468	173	468	229

Table 17: The distribution of modelled and answer-only responses across the 468 responses in each of the three tests

As in the last section, increases in the extent of modelled-responses-with-calculations were outstripped by shifts in the extent of success associated with these responses: while 75% of the 468 responses were in the modelled-responses-with-calculations category in the delayed post-test, 85% of all the correct answers (195/229) came from this category. The growing proportion of the success of this type of solutions (40-42-55) shows the growing mastery students develop with modelling and calculation. For the modelling without calculation category of responses, the number of cases were going down but the efficiency remains the same (28-9-29).

The finding of an increased willingness not only to interpret problem situations into models but also to perform written calculations provided the grounds for analyses of shifts in learner competence in terms of co-ordination between the interpretation phase and the calculation phase. With each response coded following the 3-digit coding referred to above, differences in the patterns of occurrence of these codes provided insights into their differential influence on the improvement in performance. In the initial 3-digit coding analysis that follows, I attend to the differences in the patterns of occurrence of, and influence on performance by, these codes at an overview level.

5.4 Three-bit code analysis of shifts in interpretation, calculation and evaluation

With 75% or more of the responses reflecting both a model of the situation and a calculation in the post- and delayed post-test, a great many responses could be analysed and coded for appropriateness or inappropriateness of interpretation and of calculation, alongside correctness of answer. Once all responses across all three tests were coded, counting how often each of the resulting 3-digit codes occurred provided a useful shorthand summary of the patterns of occurrence of both dominant and recessive codes across the pre-, post- and delayed post-tests. I attend, in the section that follows, to the differential contributions made to the improvement in performance, with particular focus on codes associated with large numbers of responses, across the three test sittings.

5.4.1 Distribution of 3-digit codes across the three tests

Given the increases in proportions of correct answers associated with modelled-responses-with-calculations from the pre-test to the delayed post-test, I drilled deeper into this category by using the 3-digit codes. In table 18, I summarise the patterned differences in the distribution of 3-digit codes in each the three tests, and the differential contributions made by each code to the improvement in performance reported in Chapter 4.

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	3-digit code	Pre-test	Post-test	Delayed Post-test
Modelled responses	With calculations	answer correct	111	34	130	188
			001	5	6	0
			101	1	19	7
		answer incorrect	000	35	105	51
			100	22	108	102
			110	2	1	5
	Without calculations	answer incorrect	0*0	5	12	33
			1*0	8	18	24
		answer correct	1*1	4	2	22
			0*1	1	1	1
Answer Only responses		answer correct	**1	89	15	11
		answer incorrect	**0	262	51	24
Proportions of correct answers				**134/468**	**173/468**	**229/468**

Table 18: The distribution of 3-digit codes across the 468 responses on each of the three tests

Among the responses with correct answers, the modelled-responses-without-calculations, codes reflecting appropriately coordinated structural and numerical reasoning were the 1*1 and the **1 codes. Whilst, on the one hand, the proportion of correct answers associated with the 1*1 code initially dropped from 4/134 (3%) in the pre-test, to 2/173 (1%) in the post-test, it subsequently increased to 22/229 (10%) in the delayed post-test. On the other hand, despite the proportion of correct answers associated with the **1 code dropping from 89/134 (66%) in the pre-test, to 15/173 (9%) in the post-test, to 11/229 (5%) in the delayed post-test, they still contributed one in every twenty correct answers in the end. This suggested some improvements in appropriate mental calculations in the 1*1 group, and mental work with interpretation and calculation in the **1group.

Table 18 shows that there were also codes with large numbers of responses reflecting inappropriately coordinated structural and numerical reasoning. Firstly, there were codes reflecting responses in which learners were able to set up appropriate models but were unable to see these through into calculations that produced correct answers. These were 100 codes amongst modelled-responses-with-calculations and 1*0 codes amongst modelled-responses-without-calculations. As recorded in table 18, both of these codes increased in prevalence from pre- to post- to delayed post-test: the 100 code went from 22 to 108 to 102 incidences and the 1*0 code went from 8 to 18 to 24 incidences. Whilst these responses were associated with incorrect answers, the five-and-a-half-fold increase in the prevalence of each of these codes was comprised of learners moving from the substantial prevalence of the **0 in the pre-test and indicated improvements in interpretations of the different problem situations.

Secondly, there were also codes reflecting responses in which learners were unable to set up an appropriate model, nor an appropriate calculation that could produce the correct answer. These were 000 codes for modelled-responses-with-calculations and 0*0 codes for modelled-responses-without-calculations. Whilst increases in the prevalence of 000 and 0*0 codes between the pre- and delayed post-test were not to the same degree as the increases in the 100 and 1*0 codes, the still high incidences of 000 and 0*0 codes suggests that many learners had ongoing difficulties with setting up models that reflected correct interpretations of situations.

Overall, following increases in responses reflecting the 111, 101, 100, 1*0 and 1*1 codes (all codes with an appropriate interpretation of the problem situation into a model), the findings following the first overview analysis indicated that increases in the number of correct answers from pre- to post- to delayed post-test were underpinned by improved interpretations of the different problem situations into models (mostly into bar models). With the majority of correct answers associated with the 111 code, the findings also indicated improved calculations of the solutions (with the use of the number line), suggesting that the improvement in performance was related primarily to an increase in appropriate coordination of structural and numerical reasoning on the part of the learners. 'Appropriate coordination of structural and numerical reasoning' refers to appropriately configured models and calculations that connected with each other at the level of the additive relationship presented in the problem. The extent to which the learners were coordinating structural and numerical reasoning was the focus of the second overview analysis, which was also conducted on the basis of the 3-digit coding. Here, responses reflecting appropriate interpretations and appropriate calculations – in particular, responses reflecting the 111 and 110 codes – were read as reflecting appropriately co-ordinated structural and numerical reasoning. The occurrence of the 110 code revealed, however, that, despite appropriate interpretations of the situations and appropriate calculations of the solutions, the appropriate coordination of structural and numerical reasoning did not guarantee correct answers.

Further, absence of evidence of a correctly configured model and/or calculation (as seen in answer-only responses, for instance) did not necessarily mean an incorrect answer. In the absence of evidence of appropriate interpretation and/or calculations, an inference of appropriately co-ordinated mental structural and numerical reasoning was applied to responses reflecting the 1*1 and the **1 codes. Conversely, explicit evidence of inappropriate interpretation and/or calculation with incorrect answers, seen in responses given the 0*0 and the **0 codes, were read as reflecting inappropriately coordinated structural and numerical reasoning. Additionally, responses reflecting incorrect interpretations and follow-through calculations based on the inappropriately configured model were also read as reflecting inappropriately coordinated structural and numerical reasoning. Increases in responses reflecting the 100 and 000 codes in the post-tests, suggested – in the second overview analysis – that despite increases in appropriate coordination of structural and numerical reasoning seen in the predominance of 111 codes in the delayed post-test, widespread evidence of inappropriate coordination of structural and numerical reasoning - which curtailed improvements in performance - remained.

In summarizing, large increases in the proportion of responses reflecting the 111 and 1*1 codes, and a large residual of **1 codes shows that, ultimately, the improvement in performance could be attributed to appropriate coordination of structural and numerical reasoning. At the same time, the cumulative incidences of 000, 0*0, 100 and 1*0 codes remained relatively high – even at the end – revealing continued inappropriate coordination of structural and numerical reasoning. In particular, the incidences of 000 and 0*0 codes point to ongoing challenges with the setting-up of models, and incidences of 100 and 1*0 codes point to ongoing challenges with appropriate calculations of the solutions, despite incidences of learner responses with appropriate mathematical expressions in standard form. Whilst increases in incidences of modelled responses reflecting the 101, 100, 1*0 and 1*1 codes indicated increases in the number of unsuccessful attempts at a (written or mental) calculation, the move over time was from learners who did not interpret situations into models at all (**0 codes), to learners who were, not only willing to interpret situations (mostly into bar models), but could now produce appropriately configured models, even for some of the most difficult additive word problem situations. The cumulative effect of having 7, 102, 24 and 22 responses reflecting the 101, 100, 1*0 and 1*1 codes respectively – alongside 188 responses reflecting the 111 code and 5 reflecting the 110 code – meant that 348/468 (74%) of the responses in the delayed post-test featured appropriately configured models and/or appropriate mathematical expressions in standard form. Compared to the 71/468 (15%) of the responses in the pre-test, this move over time indicates major improvements in the learners' ability to correctly interpret problem situations.

In the section that follows, I make use of the 3-digit codes and their spread across the three sub-categories of HIP, MIP and LIP devised in Chapter 4, with learner responses in each sub-category analysed in light of the findings reported on above in order to further delineate the shifts. Part of my interest is in looking at whether there were different patterns in the interpretation and calculation phases across the three categories to start with, and whether there were changes from pre-test to delayed post-test in the patterns of interpretation and calculation that linked with the changes in the evaluation phase seen in the last chapter.

5.5 A reminder of the three categories of performance: shifts from pre- to delayed post-test

As outlined in the Evaluation chapter, learners' performance on the 12 individual items in the pre-test was spread across three categories of high, mid and low initial performance as follows, with Nesher et al.'s (1982) levels of difficulty (L1, 2 or 3) included with the items:

- 2 high initial performance (HIP) items for which above 50% of the 39 learners got the correct answer (Q1: L1 & Q11: L1)

- 5 mid initial performance (MIP) items for which between 25% and 50% of the 39 learners got the correct answer (Q 2: L1; Q4: L2; Q10: L3; Q5: L3 & Q3: L1)

- ✓ 5 low initial performance (LIP) items for which less than 25% of the 39 learners got the correct answer (Q2: L1; Q8: L3; Q7: L2; Q6: L3 & Q9: L3)

Following intervention lessons,

- ✓ all 5 low initial performance (LIP) items (Q7: L2; Q12: L3; Q6: L3; Q9: L3 & Q8: L3) had shifted into the mid performance band,

- ✓ 3 of the 5 mid initial performance (MIP) items (Q5: L3; Q4: L3 & Q10: L3) remained in the mid performance band, and together with the 5 low initial performance (LIP) items brought to 8 the number of items in the mid final performance band (MFP),

- ✓ learner performance on Q2: L1 & Q3: L1 shifted these items from the mid initial performance (MIP) to the high final performance (HFP) band. Together with the two HIP items that remained in the high performance band (Qs 1: L1 & Q11: L1), there were 4 HFP items following the delayed post-test.

As noted in Chapter 4, as a result of these shifts, there were no longer any items that 25% of the learners or less got right in the delayed post-test. Thus, performance in the final test was compressed into the high and mid facility categories.

The analysis that follows is focused on the patterns of occurrence of the 3-digit codes across the LIP, MIP and HIP categories from pre- to delayed post-test, and whether these patterns were localised within particular categories.

5.5.1 Three-bit codes across the low, mid and high initial performance categories

As reported in chapter 4, the improvement in performance was such that learners provided 229 correct answers in the delayed post-test, 95 more than in the pre-test. Table 19 indicated 59 more correct answers spread across the five items in the LIP category, equivalent to an average of +11.8 correct responses per item. In the MIP category, there were 28 more correct answers also spread across five items, which translated to an average of +5.6 correct responses per item. The average increase in correct answers per item in the HIP category was +4 correct responses per item based on 8 more correct responses across 2 items. The improvement in performance was, therefore, largest in the LIP category and provides the detail behind the closing of the performance gap between the 12 items in the tests reported on in chapter 4.

When the overall results of decline in **0 codes (from 262 in the pre-test to 24 in the delayed post-test) and increase in 111 codes (from 34 in the pre-test to 188 in the delayed post-test) were applied to the sub-categories of LIP, MIP and HIP, the 3-digit code analysis revealed that the extent to which **0 codes dropped and 111 codes increased varied from category to category.

The first row of table 19 shows that the drop in **0 codes was distributed as -140 across the five items in the LIP category, -78 across the five items in the MIP category, and -20 across the two items in the HIP category. These translate to an average drop of 28, 15.6 and 10 less **0 coded responses per item in the LIP, MIP and HIP categories respectively. The decline in **0 codes was, therefore, most marked in the LIP category, followed by the MIP, then the HIP categories.

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	3-digit code	LIP (Qs 7, 6 9, 12 & 8)		MIP (Qs 5, 4, 10, 3 & 2)		HIP (Qs 11 & 1)	
				Pre-test	Delayed Post-test	Pre-test	Delayed Post-test	Pre-test	Delayed Post-test
Modelled responses	With calculations	answer correct	111	3	59	19	86	12	43
			001	2	0	2	0	1	0
			101	0	2	1	2	0	3
		answer incorrect	000	15	31	18	19	2	1
			100	8	47	13	44	1	11
			110	0	2	2	3	0	0
	Without calculations	answer incorrect	0*0	4	22	1	10	0	1
			1*0	1	7	5	13	2	4
		answer correct	1*1	1	10	2	6	1	6
			0*1	0	1	1	0	0	0
Answer Only responses		answer correct	**1	9	2	44	3	36	6
		answer incorrect	**0	152	12	87	9	23	3
Proportions of correct answers				**15/195**	**74/195**	**69/195**	**97/195**	**50/78**	**58/78**

Table 19: A summary of the distribution of 3-digit codes across the LIP, MIP, HIP categories

The last row of table 19 shows that the increase of 154 in 111 codes (from 34 in the pre-test to 188 in the delayed post-test) was distributed as 56 more 111 codes across the five items in the LIP category, 67 more 111 codes across the five items in the MIP category, and 31 more 111 codes across the two items in the HIP category. These translate to an average increase of 11.2, 13.4 and 15.5 more 111 codes per item in the LIP, MIP and HIP categories respectively. The extent of increase of 111 codes was, therefore, highest in the HIP category, followed by the MIP, then the LIP categories.

The 3-digit code overview analysis also revealed that the increase in 000 codes was localised in the LIP category (from 15 to 31 occurrences from pre- to delayed post-test). Similarly, with only one 0*0 code in the HIP category in the delayed post-test, increases in the 0*0 code were localised in the LIP and MIP categories. In contrast, increases in responses with the 100 code spanned all three initial performance categories. This localization of increases in 000 and 0*0 codes in the LIP category suggested that, despite the largest improvement in performance associated with the LIP category, LIP items still remained beyond the ability of some learners to correctly interpret by the end of the study. Nevertheless, average increases of 7.8, 6.2 and 5 more 100 codes per item across the LIP, MIP and HIP categories respectively, suggested that the increase in incidences of 100 codes per item – and hence, in the number of learners who could now correctly interpreted the word problem situations – was largest in the LIP category.

5.5.1.1 Three-digit codes from pre- to delayed post-test: LIP items

Staying with the most salient features of the shifts – the steep drop in **0 codes coinciding with a sharp increase in 111 codes – recorded in table 20 is the distribution of 3-digit codes in the pre- and the delayed post-test across the five LIP items. These items[7] have been ordered in table 20 from the most improved item (Q7) to the least improved item (Q8). With 35 and 32 responses, the two items associated with the lowest initial performance of only 1 correct answer each were Qs 6 & 9. These two items were also associated with the highest number of incorrect answer-only responses (**0) in the pre-test. Qs 6 & 9 were also the second and third most improved items after Q7, and amongst the most difficult problems on Nesher et al.'s (1982) 'scale of difficulty'. With increases to 15 and 14 correct answers in the delayed post-test, the 35 and 32 incorrect answer-only responses to Qs 6 & 9 respectively in the pre-test were reduced to only 3 & 0 in the delayed post-test.

[7] Across all sub-categories, the items have been ordered from the most improved to the least improved item.

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	3-digit code	Q7 8 + __ = 77		Q6 69 + __ = 77		Q9 53 = 19 + __		Q12 77 = 69 + __		Q8 40 = 17 + __	
				Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-
Modelled responses	With calculations	answers correct	111	*0*	*13*	*0*	*13*	*1*	*14*	*1*	*11*	*1*	*8*
			100	*1*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*1*	*0*
			101	*0*	*0*	*0*	*2*	*0*	*0*	*0*	*0*	*0*	*0*
		answers incorrect	000	*7*	*5*	*3*	*7*	*1*	*10*	*1*	*4*	*3*	*5*
			100	*2*	*9*	*2*	*12*	*1*	*5*	*1*	*11*	*2*	*10*
			110	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*2*
	Without calculations	answers incorrect	0*0	*0*	*3*	*1*	*4*	*1*	*6*	*1*	*3*	*1*	*6*
			1*0	*1*	*1*	*0*	*1*	*0*	*1*	*0*	*1*	*0*	*3*
		answers correct	1*1	*0*	*6*	*0*	*0*	*0*	*0*	*0*	*1*	*1*	*3*
			0*1	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*1*	*0*	*0*
Answer Only responses		answers correct	**1	*2*	*0*	*1*	*0*	*0*	*0*	*5*	*2*	*1*	*0*
		answers incorrect	**0	*26*	*2*	*32*	*0*	*35*	*3*	*30*	*5*	*29*	*2*
# correct answers				3	19	1	15	1	14	6	15	4	11
# total responses				39	39	39	39	39	39	39	39	39	39

Table 20: Distribution of 3-digit codes across the five LIP items in the pre-test and in the delayed post-test

More importantly, the increases in the number of learners getting Qs 6 & 9 right was seen in 13 more 111 codes each amongst these two Compare problems in the delayed post-test. While Q7 also registered 13 more 111 codes in the delayed post-test, what separated Q7 from the other items in the LIP category was a large increase of 6 more 1*1 codes. As a result, Q7 was the most improved item in the LIP category.

As pointed out in the overview analysis, the increase in the number of incidences of 000 codes was localized amongst LIP items. As seen in table 20, the five LIP items supplied more than three-fifths (5 +7 + 10 + 4 + 5 = 31) of the 51 responses reflecting the 000 code across all three categories. With 10 responses reflecting 000 codes, Q9 had the most 000 codes in the delayed post-test, followed by Q6 with 7. In contrast to these two items, as well as to every other item in this category, Q7 also includes a drop in 000 codes. With 0*0 codes also localised amongst LIP (and MIP) items, Q9 (and Q8) had the most 0*0 codes with 6, followed by Q6 with 4 in the delayed post-test. In essence, what this analysis shows is that the salient shifts in the distribution of 3-digit codes in the LIP category could be best represented by the shifts in the distribution of codes associated with the two Compare problems in the test, Qs 6 & 9.

Save for Q9, common across all items in this category are bigger increases in 111 and 100 codes relative to 000 codes. Given this result, it came as no surprise that, for learners who finished as mid- and high-attainers, the shift in the LIP category tended to be from **0 to 111 codes, and for learners who finished as low-attainers, it tended to be from **0 to 100 codes. The prevalence of 100 and 111 codes in the delayed post-test reflects shifts towards the inclusion of appropriate interpretations of problem situations into bar models. Consequently, following increased willingness by learners to interpret problem situations into bar models, the typical shift was from answer-only responses reflecting no model to responses with appropriately configured bar models in the LIP category.

To illustrate these typical moves, Mapesa started out as a low-attainer with 1/12 items correctly answered in the pre-test, and she finished as a high-attainer with 11/12 items correct in the delayed post-test. Nthabi started out as a mid attainer with 7/12 items correctly answered in the pre-test, and she finished as a high-attainer with 11/12 items correct in the delayed post-test. For instance, on Q8 (see table 21), Nthabi’s shift was from a 000 code response in the pre-test where she used tallies, to a 110 code response in the delayed post-test where she used the number line. In the pre-test, she sought to annotate four rows of ten tallies each to represent the 40 people at the park, and to strike out the 17 tallies that represented adults. Instead, because she had mistakenly annotated 11 instead of 10 tallies in the first row, there were 24 left after she had 17 struck out. In the final analysis, given that the model does not correctly reflect the whole and one of the two parts involved in the additive relationship, the interpretation and the calculation were deemed inappropriate (though perhaps more of a ‘slip’ in this case than an error), hence the 0s in the leading two positions. The rationale, as outlined in the methodology chapter, was that in the Relational Paradigm, priority is given to analysing and discerning the additive relationship, which, if misinterpreted into a model, renders the calculation inappropriate. Since Nthabi’s model is not a true reflection of the problem situation, the calculation likewise was not a true reflection of the appropriate solution process, hence the incorrect answer of 24. Over and above incorrect answer-only responses, the use of tallies in the pre-test proved to be error-prone.

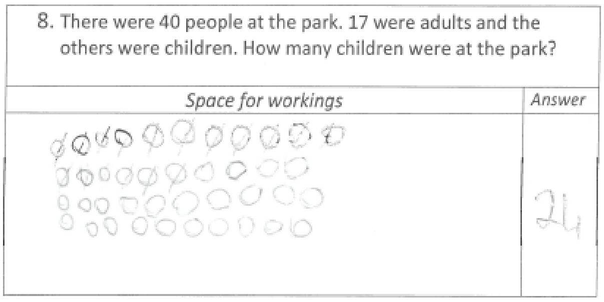
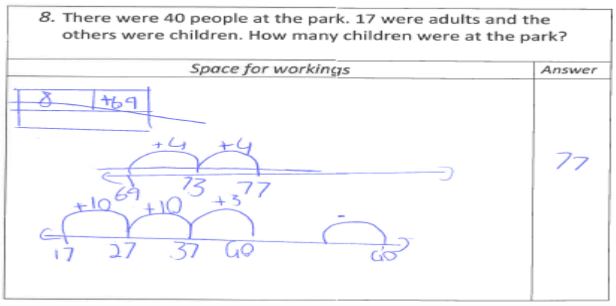
<table>
<tr><th>Nthabi’s response to the fifth LIP item (Q8) in the pre-test</th><th>Nthabi’s response to the fifth LIP item (Q8) in the delayed post-test</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Table 21: Nthabi’s responses to Q8 in the pre-test and the delayed post-test

In the delayed post-test, Nthabi appears to have read off the answer from the initially erroneous number line which she had subsequently discarded. Her working indicated an appropriate working with forward jumps from 17 to 40. In the typical N10 strategy and its variants (which she used in her responses to Qs 6, 9 & 12) the answer is always in the final landing position. What Nthabi used in Q8 is the connecting arc (CA) strategy that was not taught during intervention lessons. When using this strategy ‘adding on’ from one addend until one lands on the other addend, instead of reading off the answer from the final landing position, one has to be careful to read off the answer as the sum of the jumps. With only seven incidences of 110 code responses across all three test sittings, Nthabi’s response in the delayed post-test was an exception to the rule of correct answers resulting from appropriately coordinated structural and numerical reasoning (111 codes). As it turned out, had it not been for the incorrect choice of answer in Q8, Nthabi would have joined the other two top achievers on 12/12 in the delayed post-test.

5.5.1.2 Three-bit codes from pre- to delayed post-test: MIP items

The five items presented in table 22 are items in the MIP category. These items could be broken down into two sets: three items that ended up in the MFP band (Qs 5, 4 & 10) in which the pattern of change in 3-digit codes mirrored the changes seen in the LIP category, and two items that ended up in the HFP band (Qs 3 & 2) where higher increases were underpinned by a complete absence of 000 codes in the delayed post-test. Whilst the mean gains per item were generally lower across the three MIP items that ended up in the MFP band (Qs 5, 4 & 10) than across the five LIP items (Qs 7, 6, 9, 12 & 8), the distribution of 3-digit codes between the two categories bore a striking resemblance. Firstly, save for Q8 which moved from 1 to 8 incidences and Q10 which moved from 4 to 10 incidences, the increases in incidences of 111 codes were from single digits into teen numbers across all LIP items. Like Q9 which ended up with the most number of 000 codes amongst LIP items, Q10 also ended up with the same number of 000 codes amongst MIP items that stayed in the mid performance band in the delayed post-test. Also, like Q9 (and Q8) which had the most 0*0 codes with 6 in the delayed post-test amongst LIP items, Q10 also had the most 0*0 codes with 6 in the delayed post-test amongst Qs 5, 4 & 10. So not only did Qs 8 & 10 have larger increases in 000 codes than 111 codes, they were also associated with the most number of 0*0 codes in the end. This explains their placement at the bottom of the ranking-order post intervention, and further elucidates the localisation of 0*0 and 000 codes amongst LIP items and MIP items that ended up in the MFP band (Qs 5, 4 & 10).

Unlike the MIP items that remained at MFP level in the delayed post-test (Qs 5, 4 & 10), both MIP items that ended up as HFP items (Qs 3 & 2) registered no 000 codes in the delayed post-test. Additionally, there were larger increases in 111 codes between the two MIP items that ended up as HFP items (Qs 3 & 2) than there were amongst the three MIP items that remained at MFP level in the delayed post-test (Qs 5, 4 & 10).

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	3-digit code	Q3 52 + 49 = __		Q2 34 + 19 = __		Q5 __ - 19 = 34		Q4 40 - __ = 32		Q10 18 + __ = 40	
				Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-	Pre-	Dpost-
Modelled responses	With calculations	answers correct	111	*3*	*22*	*5*	*22*	*1*	*17*	*6*	*15*	*4*	*10*
			100	*0*	*0*	*0*	*0*	*1*	*0*	*0*	*0*	*1*	*0*
			101	*0*	*0*	*0*	*1*	*0*	*0*	*1*	*1*	*0*	*0*
		answers incorrect	000	*7*	*0*	*3*	*0*	*2*	*7*	*4*	*2*	*2*	*10*
			100	*3*	*11*	*5*	*10*	*2*	*9*	*3*	*9*	*0*	*5*
			110	*0*	*0*	*0*	*0*	*0*	*0*	*2*	*3*	*0*	*0*
	Without calculations	answers incorrect	0*0	*0*	*1*	*0*	*0*	*0*	*0*	*0*	*3*	*1*	*6*
			1*0	*2*	*3*	*1*	*3*	*1*	*4*	*1*	*3*	*0*	*0*
		answers correct	1*1	*0*	*1*	*1*	*2*	*0*	*1*	*1*	*1*	*0*	*1*
			0*1	*0*	*0*	*0*	*0*	*0*	*0*	*1*	*0*	*0*	*0*
Answer Only responses		answers correct	**1	*8*	*0*	*10*	*0*	*11*	*0*	*7*	*1*	*8*	*2*
		answers incorrect	**0	*16*	*1*	*14*	*1*	*21*	*1*	*13*	*1*	*23*	*5*
# correct answers				11	23	16	25	13	18	16	18	13	13
# total responses				39	39	39	39	39	39	39	39	39	39

Table 22: Item by item distribution of codes for MIP items that ended up as HFP items (Qs 3 & 2) and MFP items that ended up as MFP items (Qs 5, 4 & 10)

Whilst Mapesa's response to Qs10 and 4 reflects mental processing of the numerical solutions, in her responses to Qs 2, 3 and 4 (captured in table 23) there was clear evidence that the different situations in these questions were correctly interpreted into BDs, NSs and/or ENLs with correct answers in the delayed post-test.

Mapesa's responses to Qs 10, 2, 3 and 4	
Pre-test	Delayed post-test
10. 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? *Space for workings* | *Answer* 58	*10.* 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? *Space for workings* | *Answer* 22
2. Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether? *Space for workings* | *Answer* 43	*2.* Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether? *Space for workings* | *Answer* 34 | 19 53 34 + 19 = 53 +10 +9 34 44 53 53
3. 52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether? *Space for workings* | *Answer* 91	*3.* 52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether? *Space for workings* | *Answer* 52 | 49 101 52+49=101 +20 +20 +9 52 72 92 101 101
4. Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? *Space for workings* | *Answer* 12	*4.* Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? *Space for workings* | *Answer* 32 | 8 ~~8~~ 40 8

Table 23: Mapesa's responses to Qs 10, 2, 3 & 4 in the pre-test and delayed post-test

Nthabi’s responses to three of the five MIP items (Qs 2, 3 &4) reflect correct answer-only responses in the pre-test, as seen in table 24. In the delayed post-test, Nthabi’s responses reflected appropriately configured BDs alongside appropriately configured ENLs.

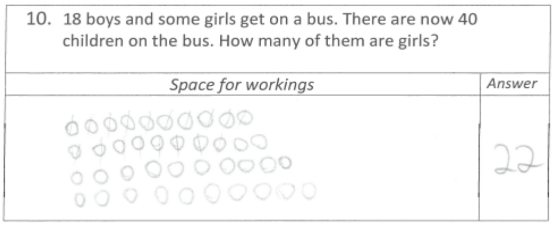
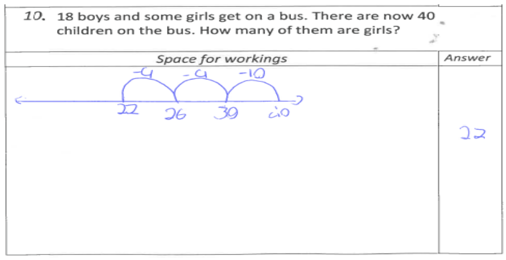
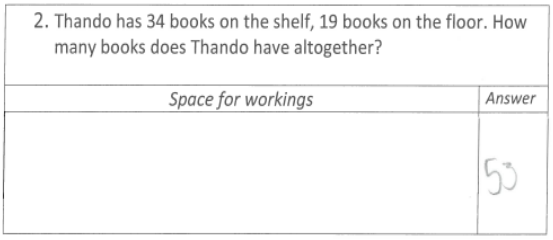
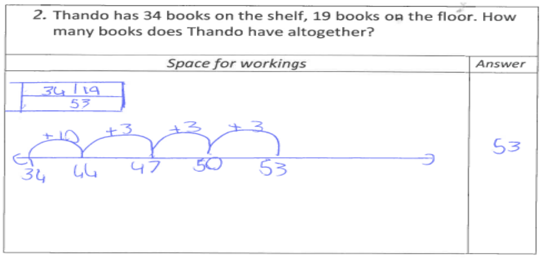
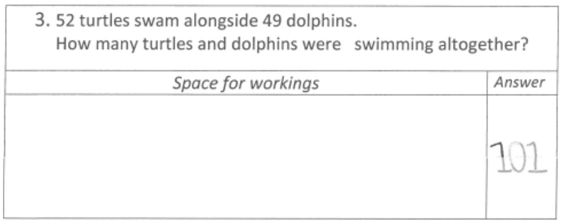
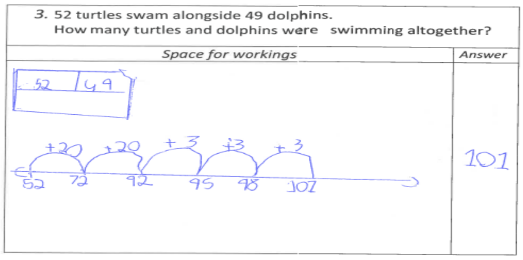
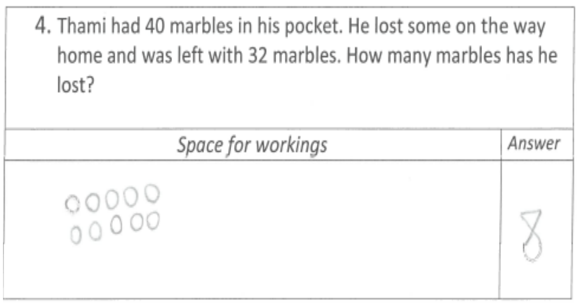
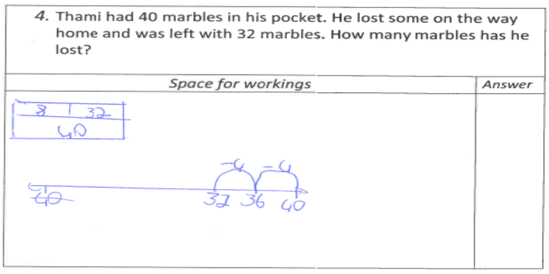
Nthabi’s responses to Qs 10, 2, 3 and 4	
Pre-test	Delayed post-test
 10. 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? Space for workings | Answer 22	 10. 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? Space for workings | Answer -4 -4 -10 22 26 30 40 22
 2. Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether? Space for workings | Answer 53	 2. Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether? Space for workings | Answer 34 | 19 53 +10 +3 +3 +3 34 44 47 50 53 53
 3. 52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether? Space for workings | Answer 101	 3. 52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether? Space for workings | Answer 52 | 49 +20 +20 +3 +3 +3 52 72 92 95 98 101 101
 4. Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? Space for workings | Answer 8	 4. Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? Space for workings | Answer 8 | 32 40 -4 -4 40 32 36 40

Table 24: Nthabi’s responses to Qs 10, 2, 3 and 4 in the pre-test and delayed post-test

5.5.1.3 Three-digit codes from pre- to delayed post-test: HIP items

With the improvement in performance associated with the two HIP items localised to Q11, Q1 maintaining its top ranking status (with 29 correct answers) following a switch from 20 correct answer-only responses (**1) to 25 responses reflecting 111 codes. With the improvement in performance on Q1 largely following a similar switch from 16 correct answer-only responses (**1) to 18 responses reflecting 111 codes, Q11 ended up on par with Q1 largely due to a retention of 6 correct answer-only responses (**1). The patterns overall were similar to the ones observed among MIP items that ended up as HFP items (Qs 3 & 2), with 111 changes appearing similarly large, 000 codes remaining at 1 or dropping to 0, and 100 codes increasing from 0 and 1 to 3 and 8 for Qs 11 & 1 respectively.

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	3-digit code	Q11 __ = 7 + 19		Q1 17 + 23 = __	
				Pre-	Dpost-	Pre-	Dpost-
Modelled responses	With calculations	answers correct	111	4	18	8	25
			100	0	0	1	0
			101	0	1	0	2
		answers incorrect	000	1	1	1	0
			100	0	3	1	8
			110	0	0	0	0
	Without calculations	answers incorrect	0*0	0	1	0	0
			1*0	0	2	2	2
		answers correct	1*1	1	4	0	2
			0*1	0	0	0	0
Answer Only responses		answers correct	**1	16	6	20	0
		answers incorrect	**0	17	3	6	0
# correct answers				21	29	29	29
# total responses				39	39	39	39

Table 25: Item by item distribution of codes for MIP items that ended up as HFP items in the delayed post-test

Also, with 4 and 2 responses reflecting 1*1 codes between Qs 11 & 1 respectively, the two HIP items recorded twice as many 1*1 codes as the three MIP items that ended up as MFP (Qs 5, 4 &10) and the two MIP items that ended up as HFP items (Qs 3 & 2). Although Q1 showed a strong shift towards modelling responses, the need to include a model did not help more learners to produce correct answers. It seems, rather, to have displaced learners' own – possibly mental –

methods. In tables 26 and 27, these typical shifts are illustrated through ‘telling’ excerpts from Nthabi’s and Mapesa’s responses to the HIP items (Qs 11 & 1), in the pre- and the delayed post-test.

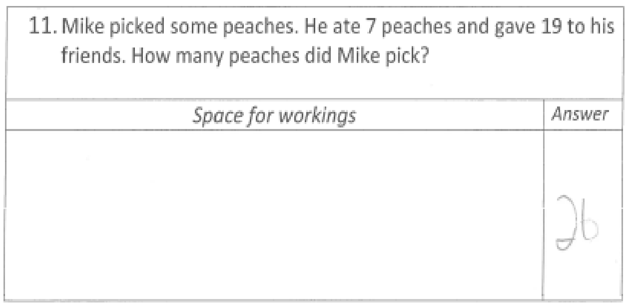
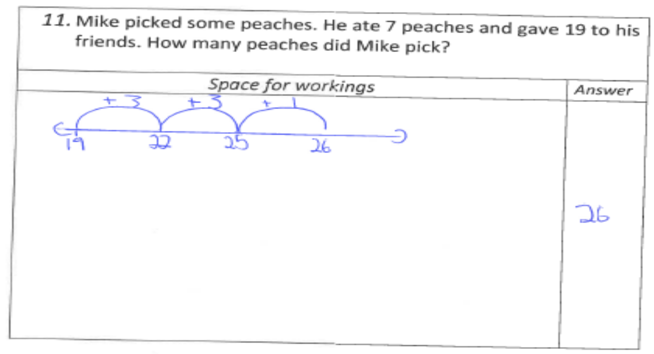
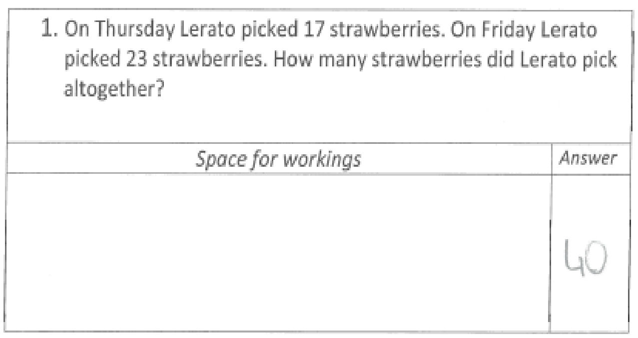
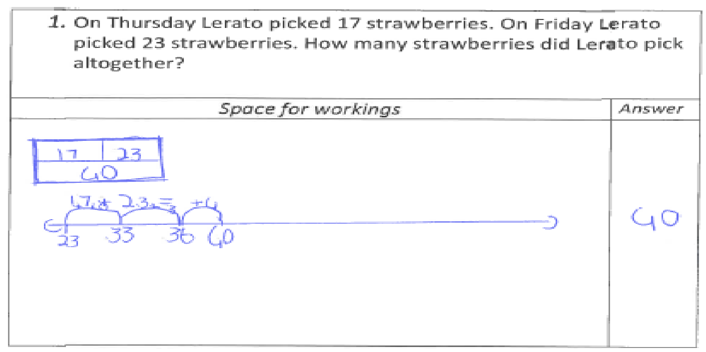
Nthabi’s responses to Qs 11 & 1 in the pre-test	**Nthabi’s responses to Qs 11 & 1 in the delayed post-test**
 11. Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick? *Space for workings* *Answer* 26	 11. Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick? *Space for workings* *Answer* +3 +3 +1 19 22 25 26 26
 1. On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether? *Space for workings* *Answer* 40	 1. On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether? *Space for workings* *Answer* 17 23 40 17 + 23 = +4 23 33 36 40 40

Table 26: Nthabi’s responses to the HIP items in the pre-test and the delayed post-test

Whilst the typical shifts in 3-digit codes for low attainers who ended up as high-attainers (like Mapesa) was from **0 codes to 111 codes, the typical shift for mid-attainers that ended up as high-attainers (like Nthabi) was from **1 to 111 codes. Nthabi’s response to Q1 in the delayed post-test shows a problem situation that was appropriately interpreted into a bar model, whilst for Q11 the correct interpretation of the situation can be read-off from the ENL. With answers correct both in the pre-test and delayed post-test, the shifts in Nthabi’s responses consist mainly in respecting the requirement for modelling. Where Mapesa’s responses in the pre-test are concerned, the correct answer-only response was provided for Q1 and an incorrect answer of 0 reflects for Q11. Her response in the delayed post-test show problem situations correctly interpreted into BDs. Whilst the response to Q11 only reflects an appropriately configured BD and the correct answer, the response to Q1 also reflects a mathematical expression with the correct operation in standard form, as well as an ENL reflecting the correct calculation of the numerical solution.

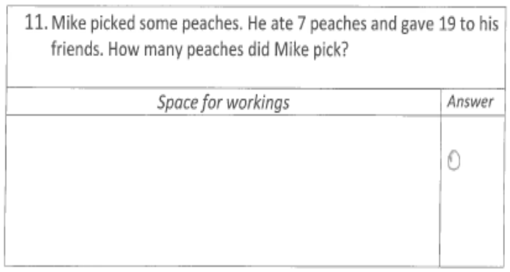
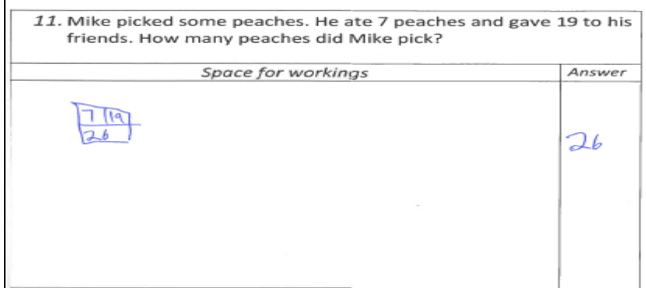
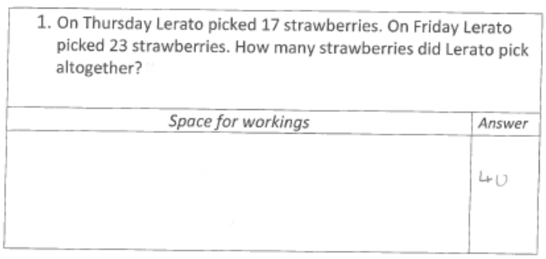
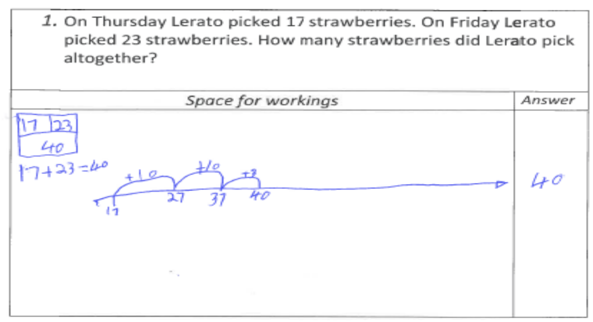
Mapesa's responses to Qs 11 & 1 in the pre-test	Mapesa's responses to Qs 11 & 1 in the delayed post-test
 11. Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick? Space for workings / Answer 0	 11. Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick? Space for workings / Answer 7 19 / 26 26
 1. On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether? Space for workings / Answer 40	 1. On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether? Space for workings / Answer 17 23 / 40 17+23=40 +10 +10 +3 17 27 37 40 40

Table 27: Mapesa's responses to the HIP items in the pre-test and delayed post-test

In the sections that follow, I provide a summary of the findings detailed above, as well as a brief concluding commentary regarding the efficacy of the intervention implemented in this study in light of the findings.

5.6 A summary of findings and concluding comments

I presented, in this chapter, analyses of learner responses across the three test sittings at an overview level where I coded for the extent of appropriate response to each question in terms of interpretation, calculation and evaluation phases, producing three digit codes (*: no evidence, 0: incorrect, 1: correct) for every single learner response. Differences in the patterns of occurrence of these codes across the three test sittings provided overview and category specific insights into shifts in learner competence within the interpretation, calculation and evaluation phases.

The overview analyses showed that, following an increase in learners' willingness to interpret problem situations into bar models, the general trend was away from answer-only responses reflecting no model towards modelled responses with appropriately coordinated structural and numerical reasoning across the attainment range. The illustrations also show a move away from less structured models (in tallies) towards the use of the bar model that allowed a focus on the structural relationship between the quantities in the problem. Further, where responses included a model in the pre-test, there was usually only one model. By the delayed post-test, and following the intervention lesson focus, the shift in the use of models was away from the exclusive use of a single model per response towards the coordinated use of the bar model for supporting structural reasoning and the number line for supporting numerical reasoning. This shift formed the basis of increases in appropriate interpretations of problem situations and appropriate calculations.

These shifts also highlighted contrasting fortunes for modelled-responses-without-calculations which showed twice the success associated with answer-only responses in the delayed post-test. This absence of a written calculation was read as a reliance on mental calculation, a reliance in turn suggesting further evidence of the contingency of working within an increased number range, 8 weeks beyond the last intervention lesson, by which time, a lot of derived number facts had become known number facts for the learners who got the problems concerned right. Overall then, given the error-prone nature of mental processing, even when situations were interpreted appropriately, the proportion of modelled-responses-without-calculations associated with correct answers remained relatively the same before and after the intervention. And herein lies the key point: these analyses reveal an empirical hierarchy for success, which, from least successful to most successful seems to be:

- modelled responses with calculations and reflecting inappropriately coordinated structural and numerical reasoning;
- answer-only responses;
- modelled-responses-without-calculations;
- modelled responses with calculations and reflecting appropriately coordinated structural and numerical reasoning.

This empirical hierarchy for success highlights the importance of coordinating structural and numerical reasoning for successful additive word problem-solving, and suggests that a relational model-based approach designed and implemented within the Relational Paradigm is valuable in the South African context. While, in some ways, this trajectory is a predictable one, it is of interest in a South African context where – as the 3-digit coding analysis of the pre-test responses revealed, learner working with interpretation and calculation is far from natural, and where, within this, the majority of correct answers came from within the answer-only response category.

The introduction of the relational models during intervention lessons appears to have helped particularly in instances where learners did not have a prior model or calculation approach that they knew how to use. The patterns of distribution of the 3-digit codes across the three categories of LIP, MIP and HIP items indicated that the general trend across the three categories was also from incorrect answer-only responses to appropriately coordinate structural and numerical reasoning. Where the HIP category was concerned, however, existing approaches were displaced with models, but this use of model seemed not to extend on the improvement in performance in any major way. In general, though, improvements in the proportion of learner responses associated with correct answers was largely due to improved interpretation of the situation into a bar model and improved calculation of the solution with the use of the number line.

A closer look at the responses that have kept the proportions of correct answers in learner responses low revealed ongoing difficulties with appropriately interpreting problem situations (as seen in the large number of 000 codes indicating inappropriately set-up models of the situations) and/or appropriately calculating the solutions for many of the learners involved (as seen in the large number of 0*0 codes indicating modelled-responses-without-calculations). Whilst the large numbers of 000 and 0*0 codes was due to still widespread evidence of inappropriate coordination of structural and numerical reasoning which curtailed the improvement in performance, the surge in 100 codes shows widespread evidence of improved interpretation into bar models. These analyses provided insights, not only into the importance of interpreting the structural relationships between the quantities in additive situations into models (Polotskaia, 2017), but perhaps more importantly, of the requirement for coordination of structural and the numerical reasoning (Giroux & Ste-Marie, 2001) for successful problem-solving.

In the chapter that follows, I analyse the calculation phase responses in more depth, with a particular focus on differences in the extent of sophistication of calculation strategies across the three test sittings.

CHAPTER 6

CALCULATION PHASE ANALYSIS OF CALCULATION STRATEGIES

6.1 Introducing the calculation phase analysis

In the previous chapter, the analysis showed improvements in performance that were underpinned by a major shift away from learner responses dominated by incorrect answers only responses (coded **0) in the pre-test to a predominance of modelled-responses-with-calculations reflecting correct answers on the basis of appropriately coordinated structural and numerical reasoning (coded 111) in the post- and the delayed post-test.

Model	**Strategy**	**Nature of strategy**	**Description of the calculation strategy**
Bar diagram (BD)	Mental	Mental	Mental
Number Sentence (NS)			
Empty Number Line (ENL)	NN (number number)		
	Connecting arc (CA)	Strategies that use 10 and its multiples/partitions as benchmarks.	Adds-on the difference instead of taking away the second operand, or subtracts the missing added. For example 51 – 49 will be treated as 49+? = 51 (and the answer is 2)
	N10C		Adds-on (or takes away) a value that is a multiple of ten closest to the second operand and then **compensates** for the amount by which the jump was exceeded. For example, 51 – 49 will be treated as 51-50+1 = (and the answer is 2)
	A10		Stands for **adding/subtracting through 10** and is a variation of N10 in which the second operand is decomposed in such a way as to enable landing on a multiple of ten. For example, 51 – 49 will be treated as 51-1-40-8 = (and the answer is 2)
	N10		‘N’ represents the initial number that is kept whole while ‘10’ represents the series of tens and then units that are then added or subtracted. For example, 51 – 49 will be treated as 51-40-9 = (and the answer is 2)
Column Model	Column addition or subtraction algorithm		Add (or subtract) the units followed by adding (or subtracting) the tens. For example, 51 – 49 will be treated as 11-9 after ‘borrowing a ten’ = (and the answer is 2)
Tallies	Counting	Counting	Count each and every tally mark or count-on from first quantity

Table 28: Hierarchy of sophistication in calculation strategies

Table 28 provides a brief recap of the range of calculation strategies discussed in Chapter 2, starting with the most sophisticated and ending with the most rudimentary. Each calculation strategy is mapped to the model that supports it. With the number line promoted in this intervention study to facilitate a shift from a reliance on rudimentary counting strategies (such as count-all and count-on) to the use of strategies that harness the base ten structure of numbers, the focus of the analysis in this chapter was on the extent to which the shift to modelled-responses-with-calculations was underpinned by shifts towards the use of ten-based calculation strategies. With the increases in the extent of modelled-responses-with-calculations outstripped by shifts in the extent of success associated with these responses, these findings provided the grounds for an analysis into patterned differences in the nature of the calculation strategies used by the learners.

With the analytical framework in table 28, I outline in this chapter, the broad shifts in learners' use of models and calculation strategies from pre- to post- to delayed post-test.

6.2. The patterns of model-use and the distribution of calculation strategies

Across the three tests, learners' solutions reflected differential use of tallies (TL), column models (CM), number sentences (NS), bar diagrams (BD) and empty number lines (ENL). In table 29, I summarize the number of answer-only responses – that is, responses that reflected no model (NM) – as 351 in the pre-test, 66 in the post-test and 35 in the delayed post-test. It also details the increasing number and nature of visible models, and the evidence of calculation strategies seen in the midst of these models.

Answer-only responses were associated with lower success rates (89/351 = 25%) than modelled-responses-without-calculations and modelled-responses-with-calculations in the pre-test. Also, as already alluded to in the literature review chapter, modelled-responses-without-calculations revealed only the interpretation of the situation and not the calculation strategy. The 18 incidences of modelled-responses-without-calculations reflected number sentences and a marginally higher success rate (5 correct answers to the 18 responses (27%). With 34/79 = 43% correct responses (all involving tally representations), the modelled-responses-with-calculations category had a higher success rate than the other two categories in the pre-test.

Type of response	With/without calculation	3-digit codes	Models	Strategies	Pre-test		Post-test		Delayed post-test	
					Extent	**Success**	**Extent**	**Success**	**Extent**	**Success**
Modelled responses	**With calculations**	111, 001, 101, 000, 100 & 110	Empty Number Line (ENL)	N10			329	136	307	162
				A10			12	7	23	21
				NN			14	5	12	9
				N10C			11	5	9	2
				CA			3	2	1	0
			Tallies (TL)	Counting	79	34			1	1
			Column Model (CM)	Column algorithm	20	6				
	Without calculations	0*0, 1*0, 1*1 & 0*1	Bar Diagram (BD) and/or Number Sentence (NS)	Not Visible (NV)	18	5	33	3	80	23
Answer-only responses		**0 & **1	No model (NM)		351	89	66	15	35	11
# responses					**468**	**134**	**468**	**173**	**468**	**229**

Table 29: Distribution of models and calculation strategies across the three tests

With calculation strategies not visible for 369 (= 351+18) of the 468 responses, there were also exactly 99 modelled-responses-with-calculations in the pre-test. Of these 99 responses, 20 showed incidences of column addition/subtraction and 79 instances of tallies. Not only did 79/99 (80%) of modelled-responses-with-calculations reflect the use of tallies, the proportions of learners who successfully used tallies was higher at 34/79 (43%) when compared to the 6/20 (30%) associated with the column model. In general then, the pre-test was characterized by the dominance of no model responses (351/468) in extent of use, and by tally-based solutions (34/79) in success in use.

Following the introduction of the BD during intervention lessons, modelled-responses-without-calculations included BD and/or NS in the post- and delayed post-tests, increasing from the 18 in the pre-test, to 33 in the post-test, to 80 in the delayed post-test. Overall, however, the success associated with modelled-responses-without-calculations, remained similar between pre- and delayed post-test.

Given that the BD was introduced alongside the ENL during intervention lessons, the post-test saw the emergence of calculations supported by the use of the ENL and the disappearance of the CM model amongst modelled-responses-with-calculations. All three ten-based calculation strategies that were promoted along with the ENL model were visible amongst learners' modelled-responses-with-calculations. These were the N10 with 329 incidences, followed by the A10 with 12 and the N10C with 11. Thus, the post-test was characterized by shifts into the N10 calculation strategy, and to a much lesser extent, into the A10 and N10C calculation strategies. Across all the strategies seen in the post-test, the dominant strategies were the N10 in extent of use (329/468 = 70%) and the A10 in success in use (7/12 = 58%) – although the raw number of successful responses involving the N10 calculation strategy (136) in the post-test was much higher.

There were two other ways of working that were not promoted during intervention lessons and yet were seen in learners' responses: the NN and the CA. While the 3 incidences of the CA – 2 of which had the correct answer – reflected a calculation strategy that exploits the inverse relationship between addition and subtraction as recognised by Beishuizen (1997), NN is probably not even a legitimate calculation strategy, let alone an efficient one. The reason for my doubts is that this way of working does not involve any decomposing or recomposing of the quantities involved. Instead, it involved a single jump from one quantity, with the other quantity annotated on the jump, hence the acronym NN for Number-Number which refers to the two quantities in the problem situation. The absence of some kind of decomposing of at least one of the quantities suggests answers that were worked out in some other way and then inscribed into the number line model. There were 14 incidences of this ‘strategy’, 5 of which returned the correct answer; the lowest success rate across strategies associated with modelled-responses-with-calculations in the post-test.

From post-test to delayed post-test, an even higher proportion of correct responses were associated with the use of the ENL underpinned by i) a reduction in the extent of use of the N10 strategy (from 329 to 307 incidences) and an increase in the success associated with its use (from 136/329 to 162/307), and ii) an increase in the extent (from 12 to 23 incidences) and in the success (from 7/12 to 21/23) in the use of the A10 strategy. In general then, the delayed post-test was characterized by the dominance of the N10 strategy (307/468) in extent of use, and by the A10 strategy (21/23) in success in use.

Whilst there were cumulative declines in the success in use of N10C and the CA strategies from one correct in every two attempts (7/14) to one in five (2/10), these declines were cushioned by the success associated with the use of the NN strategy which increased from less than one correct in every three attempts (5/14) in the post-test to three in every four (9/12) in the delayed post-test, also likely underpinned by some mental calculation. The combined increases in the overall success associated with the N10, the A10 as well as the NN strategies point to a further narrowing of the gap between correct and incorrect answers moving into the delayed post-test.

In summary then, the analysis reported in the current chapter revealed that the substantial shift into 111 codes (reported on in chapter 5) was located primarily in shifts into the N10 calculation strategy, and to a lesser extent into the A10 calculation strategy, with proportions of successful use also rising from pre- to post- to delayed post-test. There were also increases in modelled-responses-without-calculations, with success proportions remaining similar between pre- and delayed post-test, but with rises in raw numbers likely underpinned by some mental calculation in terms of success. These rises represent the upside of the drops in answer-only responses, tally-based responses, column-based and number-sentence-based responses.

In the section that follows, the broad shifts in the nature of calculation strategies reported on above are disaggregated according to the patterns of performance on the 12 items in the test established in chapter 4. I scrutinize these shifts from the point of view of the hierarchy of sophistication in calculation strategies detailed in table 29.

6.3. Linking the patterns of strategy-use to the sub-categories of performance items.

As already pointed out in section 6.2, the typical shift (from pre-test to delayed post-test) in learners' profiles of engagement was away from their 12 responses reflecting mostly answers only, and to a lesser extent the use of tallies, to a predominance of modelled-responses-with-calculations reflecting the predominant use of the N10, and to a lesser extent the A10 calculation strategies. In fact, this was the case for 34 of the 39 learners, and the shifts in their profiles of engagement from pre- to delayed post-test were best exemplified by Nthabi's work, as recorded in table 30.

Nature of shift in calculation strategy	**Number of learners whose work reflects the given pertinent shift**	**Learners whose work is typical of the given pertinent shift**	**Initial profile of engagement: correct responses/total #attempts**	**Final profile of engagement: correct responses/total #attempts**
From mental responses to N10	34	Nthabi	5/9 mental responses; 2/3 tally-based responses	8/9 N10; 3/3 A10
From tally-based responses to N10	1	Momo	6/11 tally-based responses; 0/1 mental calculation	9/9 N10; 1/1 Counting; 2/2 mental
From tally-based responses to A10	1	Sindi	6/10 tally-based responses; 0/2 mental calculation	6/6 A10; 4/5 N10; 1/1 mental
From column-algorithm to A10	1	Khufu	6/8 column-based responses; 1/4 mental calculation	7/7 A10; 5/5 N10.
From column-algorithm to mental responses	1	Sbu	0/9 column-algorithm-based responses; 1/1 tally-based response; 1/1 mental calculation	3/10 mental; 2/2 N10.
From NS-based responses to N10	1	Kamo	1/9 mental calculation; 0/3 column-based responses	8/12 N10

Table 30: The pertinent shifts in learners' profiles of engagement from pre- to delayed post-test

Across tables 31 – 34, the typical shifts are illustrated through 'telling' excerpts from Nthabi's responses.

Nthabi's responses to four of the five LIP items	
Pre-test	**Delayed post-test**
7. James had 8 marbles. John gives him some more marbles. James now has 77 marbles. How many marbles did John give to James? Space for workings / Answer: 85	7. James had 8 marbles. John gives him some more marbles. James now has 77 marbles. How many marbles did John give to James? Space for workings / Answer: 69 77
6. 69 children travel by bus. 77 children travel by train. How many more children travel by train? Space for workings / Answer: 10	6. 69 children travel by bus. 77 children travel by train. How many more children travel by train? Space for workings / Answer: 8
9. Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles? Space for workings / Answer: 72	9. Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles? Space for workings / Answer: 34
12. There were 77 people on the school bus during the trip. 69 of them were children and the others were teachers. How many of them were teachers? Space for workings / Answer: 10	12. There were 77 people on the school bus during the trip. 69 of them were children and the others were teachers. How many of them were teachers? Space for workings / Answer: 8

Table 31: Nthabi's responses to Qs 7, 6, 9 & 12 in the pre-test and delayed post-test

In the pre-test, Nthabi provided correct answers to 5/9 answer-only responses and 2/3 tally-based responses. The items with incorrect answer-only responses were four of the five LIP items. Whilst in all four of her responses to Qs 7, 6, 9 & 12 in the delayed post-test the answers were in the final landing position, she performed backward jumps in Qs 6, 9 & 12, and forward jumps in Q7. Whilst, during intervention lessons, learners were taught to perform forward/backward jumps starting from one of the two quantities given in the word problem, Nthabi's working suggests that she annotated two forward jumps of 4 that land on 79, counted back 8 from 79, and subsequently annotated the result: 69. Although eventually I identified the strategy as N10, the engagement reflects a degree of flexibility in the use of the inverse relation between addition and subtraction not seen in the use of the N10 strategy. Thus, Nthabi's profile of engagement shifted from one with incorrect answer-only responses in the pre-test to one involving the use of the N10 strategy in Qs 6 & 12 and the use of the 'subtracting through ten' strategy (A10) in Q9 to produce correct answers in the delayed post-test. Noticeably, for Qs 7 & 8, Nthabi did not use the N10 strategy to add-on the single-digit number. Instead, her strategy was to decompose the addend (8) into two parts of 4 and to add them onto the subtotals accordingly.

On the last LIP item (Q8), Nthabi's shift was from an engagement that relied on the use of tallies in the pre-test to one involving the use of the N10 calculation strategy on a number line model to produce correct answers in the delayed post-test. In the pre-test, Nthabi appears to have miscounted the tallies that remained after she had struck out 17 of them. Given the error-prone nature of having to count out each and every tally mark, it is not surprising that her answer to Q8 in the pre-test was 24, one more than the correct answer of 23.

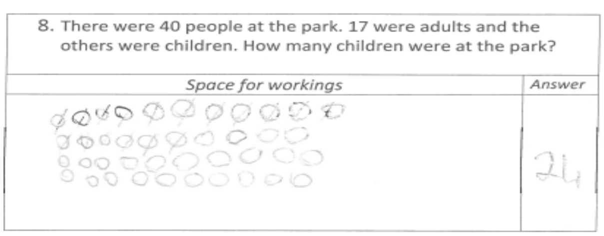
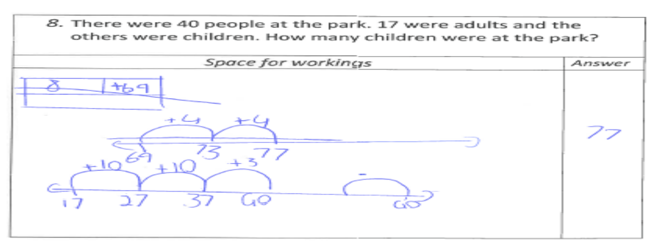
<table>
<tr><td colspan="2">Nthabi's response to the one LIP item that reflected the successful use of tallies in the pre-test (Q8)</td></tr>
<tr><td>Pre-test</td><td>Delayed post-test</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Table 32: Nthabi's responses to Q 8 in the pre-test and the delayed post-test

In the delayed post-test Nthabi appears to have read off the incorrect answer of 77 from the initially erroneous number line which she had subsequently discarded. Her working on the appropriate number line, however, indicated three forward jumps from 17 to 40. This 'adding onto' the smaller of the two quantities until you land on the other contrasted with her responses to Qs 7, 6, 9 & 12 where the answers were in the final landing position, instead of in the sum of the jumps.

The 2/3 tally-based responses with correct answers in Nthabi's profile of engagement in the pre-test were responses to two of the five MIP items in Qs 4 & 10. Nthabi approached Q10 in a manner similar to how she approached Q8 – annotating 40 tallies, striking out 18 of them and counting out the remaining 22 – suggesting a count-all strategy. For Q4, Nthabi annotated 8 tallies and wrote 8 as the answer, suggesting a counting-on strategy.

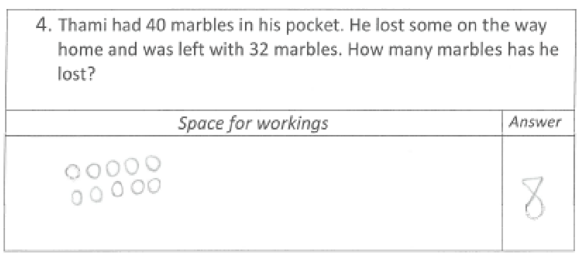
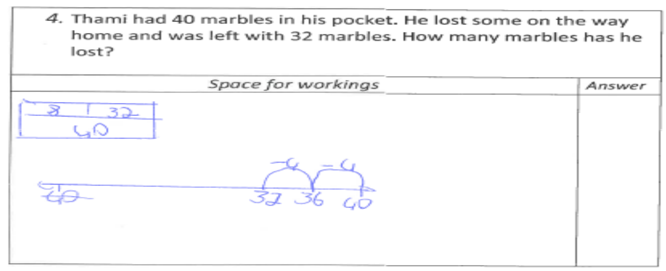
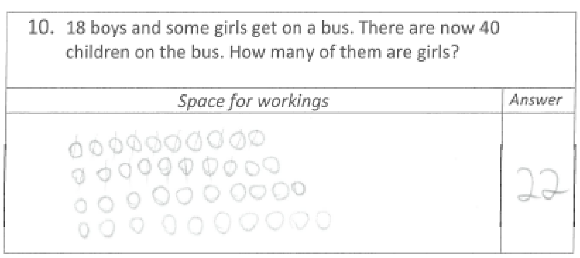
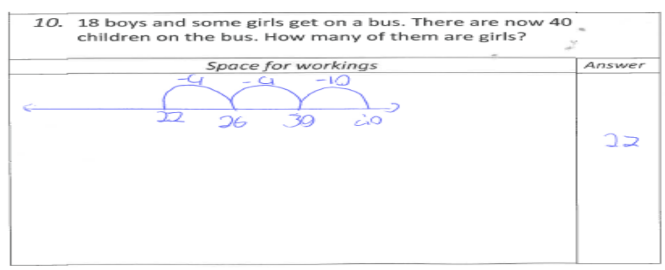
Nthabi's response to the two MIP item that reflected the successful use of tallies in the pre-test	
Pre-test	Delayed post-test
 4. Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? Space for workings Answer 8	 4. Thami had 40 marbles in his pocket. He lost some on the way home and was left with 32 marbles. How many marbles has he lost? Space for workings Answer 8 32 40 -4 -4 40 32 36 40
 10. 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? Space for workings Answer 22	 10. 18 boys and some girls get on a bus. There are now 40 children on the bus. How many of them are girls? Space for workings Answer -4 -4 -10 22 26 30 40 22

Table 33: Shifts in Nthabi's responses to the two MIP items that reflected the successful use of tallies in the pre-test

In the delayed post-test, Nthabi annotated backward jumps without landing on a multiple of ten, with answers as the sum of the jumps for Q4, and on the final landing position for Q10.

As already alluded to in the previous chapter, although Q1 showed a strong shift towards modelled responses, the inclusion of the number line model in the delayed post-test did not, however, help more learners to produce correct answers. It seems, rather, to have replaced learners' own –

possibly mental – methods, with ten-based calculation strategies, as seen in Nthabi's responses to Qs 11 & 1, both in the pre-test and in the delayed post-test.

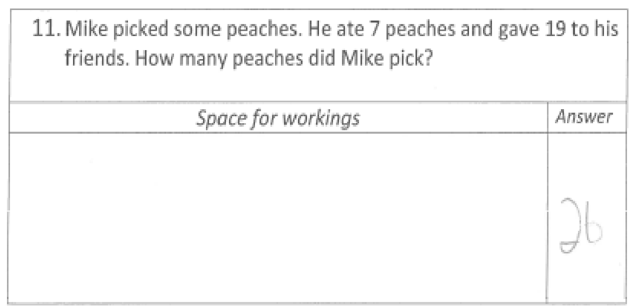
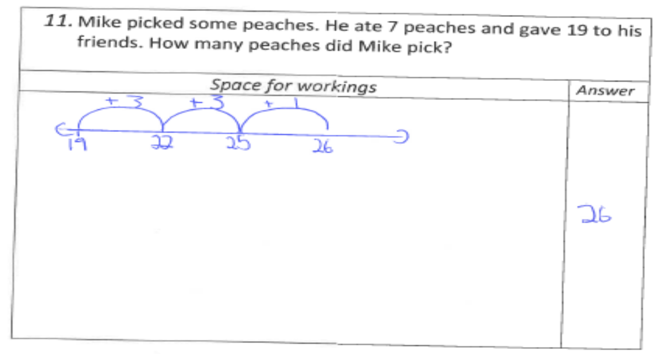
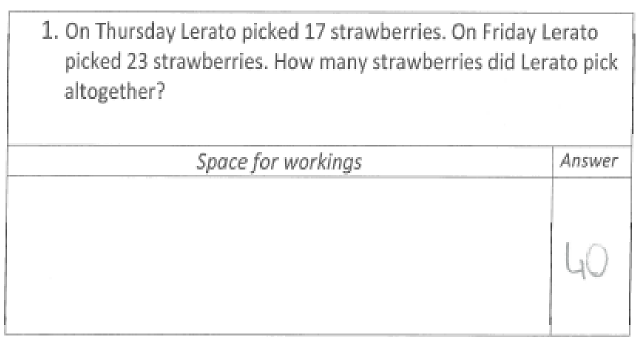
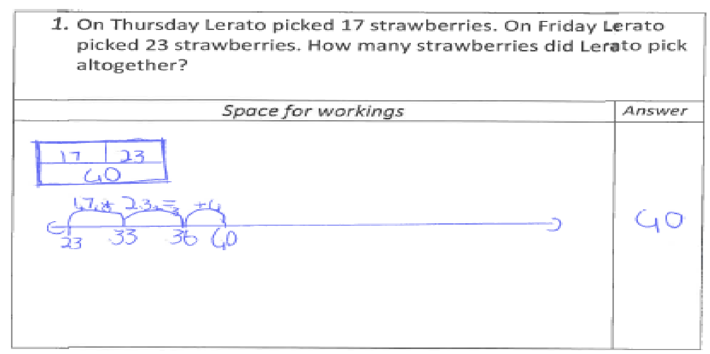
<table>
<tr><th colspan="2">Nthabi's responses to the two HIP items</th></tr>
<tr><th>Pre-test</th><th>Delayed post-test</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Table 34: Shifts in Nthabi's responses to the two HIP items in the pre- and the delayed post-test

In a pre-test in which calculation strategies were not visible for 369/468 (79%) of all responses in the pre-test, the shifts in profiles of engagements seen in the work of Momo, Sindi, Khufu, Sbu and Kamo represented important exceptions to this general shift, and possibly, deviations from the norm on the South African terrain.

More specifically, with 6/10 and 6/11 successful tally-based attempts in their pre-test profiles of engagement, the extent to which Sindi (10/12) and Momo (11/12) relied on tally-based counting fits with Hoadley's (2007) and Schollar's (2008) finding of a dominance of concrete methods such as tally counting in learners' working with number in the early grades. In the delayed post-test, whilst Sindi's profiles of engagement were dominated by the A10 strategy in extent of use (6/12) and success in use (6/6), Momo's profiles of engagement was dominated by the N10 strategy in extent of use (9/12) and in success in use (9/9) strategies.

Also unique were the overall shifts in Khufu and Sbuda's profiles of engagement both of which were initially dominated by column models. In particular, all 6/20 correct answers that reflected the use of the column addition/subtraction algorithm in the pre-test were provided by the same learner: Khufu. These six correct answers were in response to two LIP items (Qs 9 & 12), two MIP items (Qs 3 & 4), and two HIP item (Qs 1 & 11). On the one hand, having started out with a profile of engagement dominated by column-based responses in extent of use (8/12) in the pre-test, Khufu's profile of engagement was transformed to one dominated by A10 strategies in extent of use (7/12) in the delayed post-test. On the other hand, whilst Sbu also started out with a profile of engagement dominated by column-based responses in extent of use (9/12) in the pre-test, Sbu's profile of engagement was dominated by responses with no visible strategy in extent of use (10/12) in the delayed post-test. Lastly, Kamo's shifts in profiles of engagement represented the fifth exception to the rule exemplified by the shifts in Nthabi's profiles of engagement. Having started out with a profile of engagement dominated by NS-based responses (9/12) in the pre-test, Kamo's profile of engagement showed a prevalence of N10 strategies (10/12) in the delayed post-test.

In the sections that follow, I analyse how these patterns of shift played out across each of the sub-categories of LIP, MIP and HIP, particularly as they relate to the localization of the shifts primarily into the N10, and to a lesser extent into the A10 calculation strategy. To this end, the hierarchy of sophistication in calculation strategies detailed in table 28 served as the analytical framework.

6.3.1 Shifts in profiles of engagement from pre- to delayed post-test: LIP items

As already pointed out in the overview analysis, the pre-test was characterised by answer-only, tally-based, column-based and number-sentence-based responses. From pre-test to delayed post-test there were drops in the numbers and proportions of answer-only, tally-based, column-based and number-sentence-based responses, and increases in responses involving ten-based strategies on a number line. Table 35, in which the items have been ordered from the most improved item (Q7) to the least improved item (Q8), reveals the extent to which the overall shift away from mostly incorrect responses with no visible calculation strategy towards the use of ten-based strategies played out across the five LIP items (Qs 7, 6, 9, 12 & 8). The table further reveals the extent to which this shift towards the use of ten-based strategies was localized primarily into the N10 calculation strategy.

Following big shifts towards the use of ten-based jump strategies on a number line, the proportions of correct answers associated with the use of the N10 strategy were higher per item in the delayed post-test than the proportions of correct answers associated with answer-only responses per item in the pre-test. Despite these increases, however, the success levels associated with the use of the N10 strategy remained relatively low across the five LIP items - always less than 50% correct, and around 30% correct on some items. Also, one of the four attempts at using the A10 strategy for Q6 was unsuccessful, accounting for one of only two such unsuccessful attempts in the delayed post-test.

Also, tally-based counting was the most frequently appearing strategy across modelled-responses-with-calculations in learners' responses to the LIP items in the pre-test. Qs 7 & 8 were associated with relatively high incidences of tally-based counting of all five LIP items in the pre-test respectively. However, with only 1 and 2 correct answers respectively, the success associated with the use of tallies was low.

Mo dels	***Strategies***	**Q7 Correct responses/total attempts**		**Q6 Correct responses/total attempts**		**Q9 Correct responses/total attempts**		**Q12 Correct responses/total attempts**		**Q8 Correct responses/total attempts**	
		Pre	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**
ENL	N10		11/25		9/28		10/22		9/22		7/22
	A10		2/2		1/1		3/4		2/2		
	NN (mental)		1/1		3/4		1/1				0/1
	N10C		0/1						0/1		
	CA				0/1						
TL	Counting Strategies	1/10		0/3		0/1		0/1		2/6	1/1
CM	Column Algorithm			0/2		1/2		1/2		0/1	
BD and /or NS	Mental	0/1	5/8	0/1	2/5	0/1	0/9	0/1	2/7	1/2	3/12
NM	Mental	2/28	0/2	1/33		0/35	0/3	5/35	2/7	1/30	0/3
	Proportions of correct answers	**3/39**	**19/39**	**1/39**	**15/39**	**1/39**	**14/39**	**6/39**	**15/39**	**4/39**	**11/39**

Table 35: Distribution of calculation strategies across the five LIP items

Moving into the delayed post-test, the prevalence of tallies dropped substantially, but perhaps not into correct use of N10, as seen in the responses to Qs 8 & 9 by Sindi and Khufu, whose shifts were atypical. Sindi provided 1 of the 2 correct answers to Q8 with the use of tally-based unit counting in the pre-test by annotating 40 tallies, and striking out 17 of them, and Khufu provided the two correct answers associated with the use of the column addition/subtraction algorithm amongst LIP items. As seen in table 35, one of these items was the only correct answer to Q9, which, along with Q6 were the only two items with exactly one learner providing the correct answer in the pre-test. Whilst Sindi used tallies, Khufu set up the column model to subtract the 19 from the 53, and used the column subtraction algorithm to calculate the correct answer of 34.

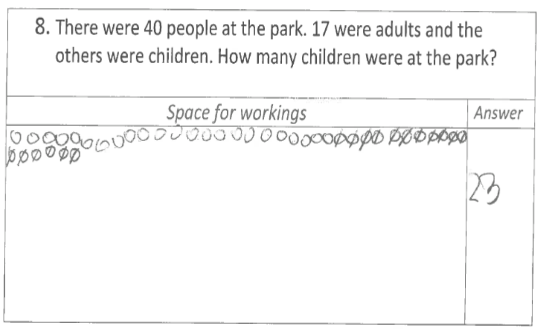
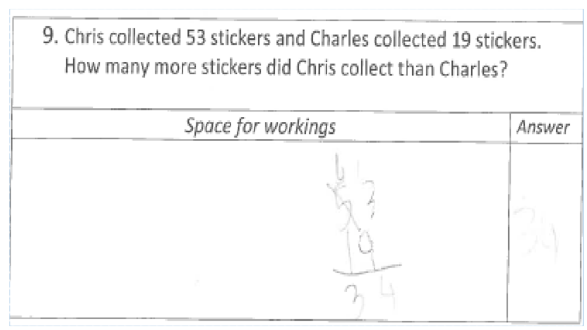
Sindi's response to Q8 in the pre-test	Only response to Q9 with correct answer in the pre-test
 8. There were 40 people at the park. 17 were adults and the others were children. How many children were at the park? Space for workings | Answer 23	 9. Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles? Space for workings | Answer 34

Table 36: Sindi's response to Q8 and Khufu's response to Q9 in the pre-test

With the cumulative use of the NN strategy seen amongst the five LIP items amounting to 7 of the total number of 12 attempts in the delayed post-test, the use of the NN strategy – which suggests mental processing of calculations – was localized in the LIP category of items.

6.3.2 Shifts in profiles of engagement from pre- to delayed post-test: MIP items

Amidst increasing incidences of responses reflecting the use of the N10 strategy, a key feature that distinguished LIP items from MIP items is a slightly higher extent of the use of the N10 strategy, coupled with above 50% success in all but one of the items in this set: Q10. In fact, with 21 of 34 attempts for Q3 and 19 of 26 attempts for Q2 returning correct answers, more than 60% of the responses associated with the use of the N10 strategy were correct. Also, not only did both items finish in the delayed post-test with success proportions tilted in favour of the use of the N10 strategy, all five attempts at using the A10 strategies on Qs 3 & 2 were successful, as recorded in table 37.

With incidences of responses reflecting no visible strategy declining from pre-test to delayed post-test and success proportions remaining similar between pre- and delayed post-test, Q10 was also the only MIP item with a residual of more than one answer-only response in the delayed post-test. Considering the interplay in the patterns of shifts between answer-only responses and responses reflecting the use of the N10 strategy, this result came as no surprise. Also, as was the case with the five LIP items, all but one (Q5) of the attempts at solving the five MIP items using the A10 strategies were successful. Overall then, the only two unsuccessful attempts at using the A10 strategy in the delayed post-test were shared one each between LIP items and MIP items.

An interesting statistic was that, in the pre-test, there were almost as many incidences of tally-based counting between the two MIP items that ended up in the HFP band (Qs 3 & 2) as there were for the five LIP items. Learners' success in tackling the two MIP items that ended up in the HFP band was, however, more than twice that associated with the five LIP items.

Models	***Strategies***	**Q3 Correct responses/total attempts**		**Q2 Correct responses/total attempts**		**Q5 Correct responses/total attempts**		**Q4 Correct responses/total attempts**		**Q10 Correct responses/total attempts**	
		Pre	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**	**Pre**	**DPost**
ENL	N10		21/34		19/26		14/28		15/27		10/24
	A10		1/1		4/4		3/4				
	NN								1/1		0/1
	N10C				1/3		0/1		0/2		
	CA										
TL	Counting Strategies	2/10		5/10		2/5		6/14		5/6	
CM	Column Algorithm	1/3		0/3		0/1		1/1		0/1	
BD and/ or NS	Strategy not visible	0/2	1/3	1/2	1/6	0/1	1/5	2/4	2/8	0/1	1/7
NM	Strategy not visible	8/24	0/1	10/24		11/32	0/1	7/20	0/1	8/31	2/7
	Proportions of correct answers	**11/39**	**23/39**	**16/39**	**25/39**	**13/39**	**18/39**	**16/39**	**18/39**	**13/39**	**13/39**

Table 37: The distribution of calculation strategies across the five MIP items

Overall then, the patterns of strategy use across all five MIP items were similar to the patterns of strategy use across the five LIP items. With each set of items showing nine attempts of the use of the A10 strategies, all but one were successful. Whilst the patterns of strategy-use revealed a similar interplay in shifts between answer-only responses and responses reflecting the use of the N10 strategy, LIP items had less than half of all attempts returning the correct answer. For MIP items, however, the cumulative effect was that more than half of all attempts to use the N10 calculation strategy returned the correct answer.

6.3.3 Shifts in profiles of engagement from pre- to delayed post-test: HIP items

A key feature that distinguished HIP items from LIP and MIP items was that, save for the N10 strategy, all attempts at using the ten-based calculation strategies promoted during intervention lessons were successful, as captured in table 38. Also, the proportion of correct answers associated with the use of the N10 strategy was higher for HIP items (Qs11 & 1) than it was for LIP and MIP items in the delayed post-test. Overall, there were higher levels of use of the range of strategies promoted during intervention lessons, and greater success with their use.

Models	***Strategies***	**Q11 Correct responses/total attempts**		**Q1 Correct responses/total attempts**	
		Pre	**DPost**	**Pre**	**DPost**
ENL	N10		14/19		23/30
	A10		2/2		3/3
	NN		2/2		1/1
	N10C		1/1		
	CA				
TL	Counting Strategies	3/3		8/10	
CM	Column Algorithm	1/2		1/2	
BD and/or NS	Strategy not visible	1/1	4/7	0/1	2/5
NM	Strategy not visible	16/33	6/8	20/26	
	Proportions of correct answers	**21/39**	**29/39**	**29/39**	**29/39**

Table 38: The distribution of calculation strategies across the two HIP items

With the performance on Q1 unchanged from pre-test to delayed post-test, the improvement in performance on HIP items was localized in Q11. Amidst global declines in incidences of responses reflecting no visible strategy from pre-test to delayed post-test, Q11 was the only HIP item with a residual of responses for which the strategy was not visible in the double digits in the delayed post-test, AND a success proportion that was higher in the delayed post-test than in the pre-test. Otherwise, the shifts in learners' choice of calculation strategy were similar for Q11 as for Q1.

6.4 A summary of findings and concluding comments

In the previous chapter, the analysis revealed responses in the pre-test dominated by incorrect answers only responses (reflecting **0 codes). The dominance of answer-only responses in the pre-test reveals, not only an absence of models of situations in learners' problem-solving toolbox, but also limitations in their ability to write out their calculations. That analysis also revealed responses with calculations reflecting greater use of tally-based over column-based and number-sentence-based (mental) calculation strategies. This finding effectively fits with Hoadley's (2007) and Schollar's (2008) observation of a dominance of concrete methods such as tally counting in learners' working with number in the early grades.

Amidst declines in incidences of responses reflecting answer-only responses, tally-based responses, column-based and number-sentence-based responses, the analysis also showed a major shift towards a predominance of modelled-responses-with-calculations reflecting shifts towards using bar models for interpretation and number line for calculation producing a correct answer in the post- and the delayed post-test. In the first part of this chapter, the analysis showed that this major shift into 111 codes was made possible by the introduction of the number line during intervention lessons. More importantly, that this major shift to 111 codes was located primarily in shift into the N10 calculation strategy, and to a lesser extent into the A10 calculation strategy, with proportions of successful use also rising from post- to delayed post-test.

With the profiles of engagement of the majority of learners transformed from mostly incorrect answer-only responses to modelled-responses-with-calculations, there were also increases in the raw numbers of modelled-responses-without-calculations, with success proportion remaining similar between pre- and del post. These increases in the in raw numbers of modelled-responses-without-calculations were likely underpinned by some mental calculation, and thus represented the upside of drops in answer-only responses, tally-based responses, column model-based calculation and number-sentence-based mental calculations.

In the second part of this chapter, I extended the analysis into an investigation into the extent to which these broad shifts in profiles of engagement played out in the three categories of LIP, MIP and HIP items. In this part of the analysis, the key patterns of shifts in calculation related to the interplay shifts between answer-only responses in the pre-test, and responses reflecting the use of the N10 strategy in the delayed post-test. The overarching pattern was that of an inverse relationship between declining incorrect answer-only responses and increasing proportions of correct answers associated with the use of the N10 strategy: the bigger the drop in incorrect answer-only responses across the three categories of LIP, MIP and HIP items, the higher the proportions of correct answers associated with the use of the N10 strategy.

This analysis also showed differences in the proportions of correct answers associated with the A10 strategy. Although the 23 incidences of the use of this strategy were spread across the three categories of LIP, MIP and HIP, it was the items that ended up in the HFP band (Qs 11, 1, 2 & 3) which yielded correct answers in every one of the responses in which the A10 strategy was used. In general, with easier items, we saw the emergence of a broader range of structure-based calculation strategies, with increasing levels of success across their use.

CHAPTER 7

SYNOPSIS OF THESIS AND CONCLUSIONS

7.1 Synopsis of literature-based theoretical framing of the study

Since the publication of the seminal work by Nesher et al. (1982), there has been a shift away from explaining the challenges that learners experience with solving additive word problems purely in cognitive terms, towards accounting for these challenges in terms of the linguistic constraints in the text of a word problem as well. One manifestation of this shift is the broad practice, catalyzed by this seminal work, of identifying additive relations word problems according to the semantics of the word problem statement.

Within traditional research, developmental levels are used to account for the different levels of difficulty associated with the different categories and types of word problems on the basis of two knowledge structures – often expressed in ordinary everyday language and arithmetic symbolic language – that are identified in this paradigm as complementary component parts and deemed as necessary for solving additive word problems. Nesher et al. (1982) hypothesized four developmental levels, and hence four levels of difficulty where additive relations word problems are concerned.

These two knowledge structures relate to learners' knowledge of the world and its objects (often described in ordinary everyday language) and learners' knowledge of logico-mathematical structures (often expressed in symbolic arithmetic language). Regarding these two knowledge structure, researchers agree that 'the logico-mathematical growth of the child cannot, of course, be understood as divorced from his experience with physical objects' (Nesher et al., 1982, p. 382). As a result, the pedagogical implication of this shift in research attention was that, instead of attending solely to developing learners' arithmetic knowledge, the expectation was now for instruction to also attend to supporting the sense that learners make of the problem situation by supporting learners' development of the knowledge that, semantically, each problem can be categorised as one of a generic class of problems. In Nesher et al.'s (1982) words: 'in fact, it is the understanding of [ordinary everyday language] as formulated in a text of a word-problem, which enables the child to choose the correct [symbolic arithmetic] operation' (p. 382). Thus, at the heart of the approach they propose for teaching is to work to establish connections between the two knowledge structures.

Around the same time that Nesher and colleagues hypothesized developmental levels on the basis of a semantic classification of additive word problems: change, combine and compare, Vergnaud (1982) proposed a classification based on the nature of the quantities involved: measure, transformation and static relation. Instead of explaining differences in learners' performance on additive relations word problems in terms of developmental levels, Vergnaud (1982) hypothesized that there are two calculi that are brought to bear during problem-solving: a *relational calculus* that is concerned with the structuring of the relationships between the quantities in the problem situation and transforming this structure into a mathematical expression (with an arithmetic operation) in standard form, and a numerical calculus to facilitate the efficient calculation of the numerical solution. In this regard, Giroux and Ste-Marie (2001) found that the difficulties that learners experience when solving additive word problems emanate from an inappropriate coordination of the operations of thought required to handle the relationships involved in the problem situation (the relational calculus) and the arithmetic operations necessary to calculate the solution (numerical calculus). Working from this dual calculi base, Polotskaia (2017) classified all approaches that have the dual development of a *flexible relational understanding of problem structures* and a *numerical sequential understanding of operations* as a goal, into what she and her colleagues termed the Relational Paradigm.

Based on evidence pointing to particular challenges among South African learners relating to difficulties with making sense of mathematical word problems and moving beyond counting based calculations, this study was designed and implemented with a dual focus on developing learners' sense-making (awareness of structure) and efficient calculation necessary for solving additive word problems. As outlined in chapter 2, the position taken in this study is that the relational calculus is primarily brought to bear whenever the need to interpret (or make sense of) the mathematical structure of the problem at hand arises, and that it manifests as structural reasoning. I therefore, defined structural reasoning as reasoning on the structural relationships involved in the problem situation in attempts to make sense of it and to transform it into a mathematical expression in standard form, and numerical reasoning as reasoning on the relationships between numbers in the process of calculating the solution to the problems.

Drawing from Hegarty et al. (1995) who found the differentiating feature of successful problem-solving to be the interpretation into a model of the situation described in the problem, the implication for designing intervention lessons for a study conducted within the Relational Paradigm were, not only to specifically encourage learners to interpret problem situations into models and base their solution plan on this model, but more generally, to disrupt learners' inclination to 'immediately lock themselves into manipulating numbers' (Freiman et al., 2017, p. 836) by keeping them focused on the mathematical structure of the problem. In particular, the defining feature the Relational Paradigm is distinguishing between the analyses of the word problem, which involves the construction of a model of the situation described in the problem together with the deriving a NS, and the execution of the operation chosen based on the model of the situation described in the problem. Making this distinction explicit allows for deliberate attention to be paid to each of the two distinct phases separately. Motivated by Polotskaia (2015) attributing success in solving word problems to the construction of a model of the situation described in the problem and basing the solution plan on this model, and given my interest in the dual development of learners' structural and numerical reasoning, I designed and implemented this study within the Relational Paradigm.

In valuing the balanced development of structural and numerical reasoning, I adapted the Equilibrated Development Approach (EDA) for this study from the version developed and promoted by Polotskaia and colleagues within the Relational Paradigm. One key adaptation related to my interest in the development of learners' structural reasoning for which the bar diagram (BD) was promoted in this study as a relational model capable of expressing the part-part-whole relations in any additive relationship. This contrasted with Polotskaia's (2017) approach involving the use of line segment diagrams named Arrange-All (A-A) diagrams as models that can facilitate thinking about and configuring the additive relationships whilst representing quantities rather than numbers Given the context of a South African curriculum that introduces learners to number through counting, and then swiftly proceeds to expose the learners to symbolic representations of number, it made sense to promote models that incorporate numbers symbolically. Thus, in this study, an indicator of a correct interpretation of the additive relationship in the word problem was an appropriately configured BD, or any other model of the situation that incorporates numbers symbolically.

Another key adaptation related to my interest in the development of learners' numerical reasoning was to incorporate the RME number structuring approach with a view to transforming learners tally-based counting into efficient calculation using ten based calculation strategies. For this reason, the empty number line (ENL) was promoted in this study as a relational model that embodies the sequential structure of number, and can support efficient calculation of the solution. As a result, an indicator of a correct calculation was a correct numerical answer. Moreover, an indicator of sophistication during calculation was a part-whole construction of number a compressed jump of a multiple of 10 instead of multiple jumps by 10 or counting in ones (Ellemor-Collins & Wright, 2009).

In light of the aforementioned issues, difficulties associated with the solving of additive relations word problems are reconceptualised and reformulated in the Relational Paradigm into either a failure to grasp the additive relationship in the word problem and interpret it into a part-part-whole model, or a failure to execute the chosen operation. With difficulties associated with the solving of additive relations word problems reconceptualised and reformulated in this way, the overarching claim made within the Relational Paradigm is that working relationally allows learners to move towards being as effective in solving 'difficult' problems as in solving 'easy' problems. I also analysed the outcomes from the current study in relation to this hypothesis, and the findings were interesting.

7.2 Synopsis of the findings of the study

As outlined in chapter 3, over and above the analysis conducted to determine whether or not there was improvement in the learners performance on the items in the test, I conducted three micro analyses relating to the interpretation, calculation and evaluation phases as promoted during intervention lessons, and one macro analysis relating to the extent to which leaners' responses reflected appropriate coordination of structural and numerical reasoning. Whilst the phases of problem-solving were described in the order of the problem-solving cycle (interpretation, calculation and evaluation) in the literature and theory chapter, I began my analysis with a primary focus on the evaluation phase. This is how I provide a summary of the findings as well.

7.2.1 Evaluation phase analysis

The first part of the evaluation phase analysis related to overall performance as seen purely from the point of view of the proportion of correct answers provided by the learners in the pre-test. These numbers (percentages) provided the grounds for the second part of the initial analysis in which the focus was on the shifts in the 39 learners' performance on the 12 test items from pre- to delayed post-test. These analyses demonstrated increases in the proportion of correct answers from 29% in the pre-test, to 37% in the post-test, to 49% in the delayed post-test respectively. With the largest gains associated with the most 'difficult' problems, the findings point to how, following the intervention, the gap between learner performance on 'easy' and 'difficult' items narrowed. In particular, the two Compare problems in the test, which, in requiring learners to make sense of the relation between the two quantities involved, were theoretically the most difficult items in the test according to Nesher et al.'s (1982) 'scale of difficulty'. With only one learner getting each of them right in the pre-test, they were also the most difficult problems in the pre-test from an empirical point of view. Following increases of 33% or more in correct answers, there were many more learners providing correct answers to the Compare problems than to some combine result unknown problems in the delayed post-test. Additionally, not only was the gap between the high and the mid performance bands reduced, but overall, the difference in correct answers between any two consecutive items in the final performance ranking order was not more than 4.

This finding of a narrowing of the performance gap between 'easy' and 'difficult' items is key as it points to the salience of Polotskaia's (2017) hypothesis that working relationally allows learners to move towards being as effective in solving 'difficult' problems as in solving 'easy' problems. Interesting alongside this finding, following a juxtaposing of learners' empirical performance (as seen in the number of learners who got each item correct) with each item's difficulty level (as determined by research), the second analysis revealed how working relationally brought learners' empirical performance on these items into perfect alignment with the 'levels of difficulty' hypothesised by Nesher et al. (1982). Taken together, particularly in a short intervention on a Grade 3 class, these findings suggest that, whilst working relationally certainly does allow learners to move some way towards being as effective in solving ‘difficult’ problems as in solving ‘easy’ problems, it does so without necessarily disrupting the hierarchy of difficulty hypothesised by Nesher et al. (1982).

7.2.2 Interpretation phase analysis

In the preliminary analysis of the sense that learners made of the word problems in the interpretation phase, the findings related not only to an increased willingness to interpret problem situations into models. This finding provided the grounds for analyses of shifts in learner competence across ‘easy’ and ‘difficult’ items. The overview findings indicated that the increases in the number of correct answers seen in the evaluation phase analysis were underpinned by improved interpretations of the different problem situations into models (mostly into bar models).

The extent to which the learners were coordinating structural and numerical reasoning was the focus of the macro analysis (and second overview analysis), which was also conducted on the basis of the 3-digit coding. Here, responses reflecting appropriate interpretations and appropriate calculations – in particular, responses reflecting the 111 and 110 codes – were read as reflecting appropriately co-ordinated structural and numerical reasoning. The occurrence of the 110 code revealed, however, that, despite appropriate interpretations of the situations and correct calculations of the solutions, the appropriate coordination of structural and numerical reasoning did not guarantee correct answers. Despite increases in appropriate coordination of structural and numerical reasoning (as seen in the predominance of 111 codes in the delayed post-test), widespread evidence of inappropriate coordination of structural and numerical reasoning remained (as seen in increases in responses reflecting the 100 and 000 codes in the post-tests).

These two key findings from the two overview analyses led into an analysis of learner responses in the interpretation, calculation and evaluation phases across the three categories of high, mid and low initial performance items (HIP, MIP and LIP) identified in the last chapter. This was undertaken with an interest in whether there were different patterns of occurrence of 3-digit codes in each sub-category which linked with the increases in proportions of correct-answer responses presented in the evaluation phase analysis chapter over time. In particular, I was interested in shifts in learners' reasoning strategies, modelling and calculation and in the extent to which these shifts were spread across all three categories, or localised within particular categories.

These patterns of distribution of the 3-digit codes across the three categories of LIP, MIP and HIP revealed that, whilst the improvement in appropriate interpretation and coordination of structural and numerical reasoning could be seen across all three item sub-categories, it was in the items in the LIP sub-category where inappropriate coordination of structural and numerical reasoning remained most widespread. Thus, despite the improvements in performance (reported in 7.2.1) that saw the items in the LIP category move as a group into the mid performance band in the delayed post-test following an almost twenty fold increase in 111 codes and an almost six fold increase in 100 codes (due to teaching effort), these items remained beyond the reach of some learners by the end of the study (due to persistent challenges with calculation).

7.2.3 Calculation phase analysis

With the majority of correct answers associated with the 111 code, the findings also indicated improved calculations of the solutions (with the use of the number line), suggesting that the improvement in performance was related primarily to an increase in appropriate coordination of structural and numerical reasoning on the part of the learners.in this regard, an overarching finding of this study related to the dominance of answer-only responses in the pre-test (75% of all pre-test responses), which revealed, not only an absence of models of situations in learners' problem-solving repertoire, but also limitations in learners' ability to produce written calculations. Where the pre-test was concerned, the calculation phase analysis revealed responses with calculations reflecting greater use of tally-based over column-based and number-sentence-based calculation strategies. This finding effectively corroborated Hoadley's (2007) and Schollar's (2008) observation of a dominance of concrete methods such as tally counting in learners' working with number in the early grades.

Amidst declines in incidences of responses reflecting answer-only responses, tally-based responses, column-based and number-sentence-based responses, the analysis also showed a major shift towards a predominance of modelled-responses-with-calculations reflecting 111 codes in the post- and the delayed post-test. The calculation phase analysis showed that this major shift into 111 codes was made possible by the introduction of the number line during intervention lessons. Further, this major shift to 111 codes was located primarily in shifts into the N10 calculation strategy, and to a lesser extent into the A10 calculation strategy, with proportions of successful use also rising from post- to delayed post-test. This finding is particularly important in a South African context where inefficient counting-in-ones has been described as so prevalent.

With the profiles of engagement with the word problems transformed from mostly incorrect answer-only responses to modelled-responses-with-calculations for the majority of learners, there were also increases in the raw numbers of modelled-responses-without-calculations, with success proportions remaining similar between the pre- and delayed post-tests. The increases in the raw numbers of modelled-responses-without-calculations were likely underpinned by some mental calculation, and thus represented a further upside of drops in answer-only responses and tally-based responses.

With patterns of shifts into the use of more efficient calculation strategies across the three categories of LIP, MIP and HIP localized primarily into the N10, and to a lesser extent into the A10 calculation strategy, two more important outcomes were associated with the emergence of strategies that were not taught to learners during intervention lessons: the NN and the CA. Whilst the use of the NN involves a single jump linking the two given quantities, with the sum/difference annotated on the jump, and hence was likely underpinned by some mental calculation, the CA exploits the inverse relationship between addition and subtraction as recognised by Beishuizen (1997). Their emergence in the post- and delayed post-test suggest that there are extended benefits of working relationally, both within and across the interpretation and calculation phases of the problem-solving cycle promoted in this study.

7.3 Research implications

Despite the short-term nature of the intervention, these results suggest that, in the context of a relational approach that promotes the use of relational models in order to support structural and numerical reasoning, improvements in performance can be achieved in relatively short time frames. This study has shown that these improvements in performance were underpinned by improvements in interpretation, calculation and the coordination of structural and numerical reasoning brought to bear during problem-solving as supported by relational models. This has implications both for teaching and learning additive relations word problems, as well as for future research.

7.3.1 Implications for learning and teaching additive relations word problems

The outcomes of this study corroborate the finding by Venkat et al. (2021) that, in a South African context of low performance and low resources, an approach based on 'key models and connections' for improving sense-making can improve results substantially. The contribution made by this study is that incorporating the RME number structuring approach with a view to transforming learners tally-based counting into efficient calculation using ten based calculation strategies can do the same for improving calculation. This suggests that, given the SA context and the need to focus both on interpretation and calculation, improving learners' performance on additive word problems can be achieved with an approach that views problem-solving as a cycle with a clear distinction between the interpretation and the calculation phases. The reason for the success of the intervention can be attributed to 'relational-directed teaching practices' (Polotskaia, 2017) that explicitly distinguished between the interpretation and the calculation phases, thus enabling the teacher of additive relations problem-solving to pay independent and deliberate attention to supporting each phase with appropriate relational models.

The results show that, like the A-A model promoted by Polotskaia and colleagues, the bar diagram and the number line can serve as meta-cognitive tools, the appropriate use of which can be taught and learned. Additionally, compressing the problem-solving cycle into three phases allows for the protocol (to first interpret the mathematical structure of the word problem into a BD, which is then translated into a NS, and finally into an ENL), to be explicated to learners in ways that serve as a meta-cognitive tool combination process that they can apply across all additive word-problem.

7.3.2 Implications for further research

What this study has demonstrated is the veracity of Davydov's (1982) conviction that young learners, regardless of their 'already acquired' knowledge, are capable of grasping the additive relationship. Considering that the learning trajectory envisioned by Davydov (1982) involved learners first playing with physical objects such as strings, and describing relationships between quantities without referring to numbers, it is particularly telling that the successes of this intervention were realised in a context of ‘a trajectory in which counting functions as the fundamental action, with counting actions becoming reified and truncated over time around mathematical structures and properties’ (Venkat et al., 2021, p. 3).

The results also suggest that the improvements in performance can be broadly attributed to increases in appropriate coordination of structural and numerical reasoning seen in the predominance of 111 codes. However, the lingering evidence of inappropriate coordination of structural and numerical reasoning suggests the need for further study of ways in which the inappropriate coordination of structural and numerical reasoning can be further reduced.

7.4 Research limitations

Despite the success of the intervention, some limitations of the study need noting. Firstly, the study was based on observations made in one class in one suburban school. There is a strong possibility, given the widespread inequality in South Africa, that results from a class in different setting, perhaps in a rural or more disadvantaged location, would yield different outcomes. Secondly, in between intervention lessons, the class continued to be taught by its designated teacher who sometimes sat in on intervention lessons. Judging by the bigger improvement from post- to delayed post-test relative to the improvement from pre- to post-test, there is a possibility that this class may have benefited from some teaching related to the intervention approach that happened beyond the intervention lessons rolled out. Finally, the exploratory nature of this intervention meant that a parallel control group was not the priority in the way that it might have been in a more experimental study design. However, the finding of a tilt in balance away from a predominance of answer-only responses in the pre-test to a prevalence of the responses involving the use of the BD and the ENL, models with which learners were not familiar prior to their participation in intervention lessons, suggests that the improvements in performance were strongly linked to the intervention.

Given the limitations, it would be inappropriate to claim any generalizability of the results of this study. Nevertheless, the results show that intervening with relational-directed teaching practices in an instructional design formulated within the Relational Paradigm can improve the sense that learners make of additive word problem situation. Further, when the design of the intervention is adapted to also attend to learners’ calculation, learners’ reliance on tally-based counting strategies can be transformed to the type of calculation that utilizes the structure of number, with ten as a benchmark. More importantly, that the development of learners' structural and numerical reasoning can be initiated in the early grades of their mathematical instruction in school.

7.5 Autobiographical reflections

We learn from ancient manuscripts that are continuously being unearthed in all sorts of unexpected places that the basis of the science of Khemet is the knowledge that:

the human brain is composed of two parts with diametrically opposed, yet complementary functions, known as the right and the left hemispheres of the brain. Whilst the right hemisphere is in charge of all functions of unification and integration, the left hemisphere of the brain is in charge of all functions of separation and differentiation. It is also said that, whilst the mode of thinking of the right hemisphere is synthetical, holistic, congregative, etc., the mode of thinking of the left hemisphere is analytical, serial, segregative etc. In other words, whilst the right hemisphere notices similarities between things, and their relation to each other and the whole, thus unifying them, the left hemisphere notices differences between things, separating wholes into parts, and enabling us to deal with all sequential phenomena (wholes presented pieces at a time)' (Amen, 1977, p. 9).

I am sure that the extended quote above will resonate with you, Dr. Polotskaia, given the reference that you and your colleagues are continuously making to neuroscience in your research. It was this reference that caught my attention as I was grappling with how to design my study in such a way as to cater for the inherent duality that our ancestors declared as the foundation of our being. The knowledge and understanding expressed in this quote is where it all started for me.

Thank you for affirming, for me at least, that the discerning of the additive relationship and the calculation of the numerical solution are indeed two critical and complementary processes that integrate word problem-solving. Processes which, in instruction, can and should be attended to, both separately and as one. It is only with this equilibrated attention that we will be able to mould future generations into integrated individuals, i.e. men and women of integrity.

I thank you

LIST OF REFERENCES

Amen, R. U. N. (1977). Metu Neter Vol. 1, Bronx, New York, Khamit Co: Pub.

Anghileri, J. (2006). Scaffolding practices that enhance mathematics learning. *Journal of Mathematics Teacher Education, 9*(1), 33-52.

Askew, M. (2004). *BEAM's Big Book of Word Problems for Years 3 and 4*: Nelson Thornes.

Askew, M. (2019). Mediating primary mathematics: Measuring the extent of teaching for connections and generality in the context of whole number arithmetic. *ZDM, 51*(1), 213-226.

Askew, M., Venkat, H., Abdulhamid, L., Mathews, C., Morrison, S., Ramdhany, V., & Tshesane, H. (2019). *Teaching for structure and generality: Assessing changes in teachers mediating primary mathematics.* Paper presented at the 43rd Conference of the International Group for the Psychology of Mathematics Education, 07-12 July 2019, Pretoria, South Africa.

Bai, X., & Ola, A. (2017). A tool for performing item analysis to enhance teaching and learning experiences. *Issues in Information Systems, 18*(1), 128-136.

Barnes, D. (2010). Why talk is important. *English Teaching: Practice and Critique, 9*(2), 7-10.

Baroody, A. J. (2006). Mastering the basic number combinations. *Teaching children mathematics, 23*, 22-31.

Beckmann, S. (2004). Solving algebra and other story problems with simple diagrams: A method demonstrated in grade 4-6 texts used in Singapore. *The mathematics educator, 14*(1).

Beishuizen, M. (1993). Mental strategies and materials or models for addition and subtraction up to 100 in Dutch second grades. *Journal for Research in Mathematics Education*, 294-323.

Beishuizen, M. (1997). Development of mathematical strategies and procedures up to 100. *The role of contexts and models in the development of mathematical strategies and procedures*, 127-162.

Beishuizen, M., Van Putten, C., & Van Mulken, F. (1997). Mental arithmetic and strategy use with indirect number problems up to one hundred. *Learning and Instruction, 7*(1), 87-106.

Blöte, A. W., Klein, A. S., & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction, 10*(3), 221-247.

Blum, W., & Leiss, D. (2005). *„Filling Up "-the problem of independence-preserving teacher interventions in lessons with demanding modelling tasks.* Paper presented at the CERME 4–Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education.

Cai, J. (2004). Developing algebraic thinking in the earlier grades: A case study of the Chinese elementary school curriculum. *The mathematics educator, 8*(1), 107-130.

Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of problem solving: A study of kindergarten children's problem-solving processes. *Journal for Research in Mathematics Education*, 428-441.

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementary School Journal, 97*(1), 3-20.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). Children's mathematics. *Cognitively Guided*.

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education*, 3-20.

Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, 179-202.

Carpenter, T. P., Moser, J. M., & Bebout, H. C. (1988). Representation of addition and subtraction word problems. *Journal for Research in Mathematics Education*, 345-357.

Carpenter, T. P., Moser, J. M., & Hiebert, J. (1981). Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems. *Journal for Research in Mathematics Education, 12*(1), 27-39.

Chan, E. C. M. (2009). Mathematical modelling as problem solving for children in the Singapore mathematics classrooms. *Journal of Science and Mathematics Education in Southeast Asia, 32*(1), 36-61.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational researcher, 32*(1), 9-13.

Cohen, L., Manion, L., & Morrison, K. (2013). Validity and reliability *Research methods in education* (pp. 203-240): Routledge.

Davydov, V. V. (1999). What is real learning activity. *Learning activity and development*, 123-138.

Davydov, V. V. (2008). *Problems of developmental instruction: A theoretical and experimental psychological study*: Nova Science Publishers, Incorporated.

Davydov, V. V., & Markova, A. l. K. (1982). A concept of educational activity for schoolchildren. *Soviet psychology, 21*(2), 50-76.

DBE, S. A. D. o. B. E. (2012). *Report on the Annual National Assessments 2012: Grades 1 to 6 & 9*. Retrieved from Department for Basic Education:

De Corte, E., Verschaffel, L., & Greer, B. (2000). *Connecting mathematics problem solving to the real world.* Paper presented at the Proceedings of the International Conference on Mathematics Education into the 21st Century: Mathematics for living.

Ellemor-Collins, D., & Wright, R. B. (2009). Structuring numbers 1 to 20: Developing facile addition and subtraction. *Mathematics Education Research Journal, 21*(2), 50-75.

Ellemor-Collins, D., & Wright, R. J. (2007). Assessing pupil knowledge of the sequential structure of number. *Educational and Child Psychology, 24*(2), 54-63.

Ensor, P., Hoadley, U., Jacklin, H., Kuhne, C., Schmitt, E., Lombard, A., & Van den Heuvel-Panhuizen, M. (2009). Specialising pedagogic text and time in Foundation Phase numeracy classrooms. *Journal of Education, 47*(2009), 5-30.

Fellus, O., & Biton, Y. (2017). One is Not Born a Mathematician: In Conversation with Vasily Davydov. *International Journal for Mathematics Teaching & Learning, 18*(2).

Fischer, J.-P., Vilette, B., Joffredo-Lebrun, S., Morellato, M., Le Normand, C., Scheibling-Seve, C., & Richard, J.-F. (2019). Should we continue to teach standard written algorithms for the arithmetical operations? The example of subtraction. *Educational Studies in Mathematics, 101*(1), 105-121.

Fleisch, B. (2008). *Primary education in crisis: Why South African schoolchildren underachieve in reading and mathematics*: Juta and Company Ltd.

Fleisch, B., Schöer, V., Roberts, G., & Thornton, A. (2016). System-wide improvement of early-grade mathematics: New evidence from the Gauteng Primary Language and Mathematics Strategy. *International Journal of Educational Development, 49*, 157-174.

Freiman, V., Polotskaia, E., & Savard, A. (2017). Using a computer-based learning task to promote work on mathematical relationships in the context of word problems in early grades. *ZDM, 49*(6), 835-849.

Freudenthal, H. (1973). *Mathematik als pädagogische Aufgabe* (Vol. 2): Klett Stuttgart.

Freudenthal, H. (1986). *Didactical phenomenology of mathematical structures* (Vol. 1): Springer Science & Business Media.

Galbraith, P. (2012). Models of modelling: Genres, purposes or perspectives. *Journal of Mathematical Modelling and application, 1*(5), 3-16.

Giroux, J., & Ste-Marie, A. (2001). The solution of compare problems among first-grade students. *European journal of psychology of education, 16*(2), 141-161.

Gravemeijer, K. (1997). Solving word problems: A case of modelling? *Learning and Instruction, 7*(4), 389-397.
Gravemeijer, K. (2002). *Emergent modeling as the basis for an instructional sequence on data analysis.* Paper presented at the Proceedings of the 6th International Conference on Teaching Statistics.
Gravemeijer, K. (2007). Emergent modelling as a precursor to mathematical modelling *Modelling and applications in mathematics education* (pp. 137-144): Springer.
Gravemeijer, K., Bruin-Muurling, G., Kraemer, J.-M., & van Stiphout, I. (2016). Shortcomings of mathematics education reform in the Netherlands: a paradigm case? *Mathematical thinking and learning, 18*(1), 25-44.
Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics, 39*(1-3), 111-129.
Gravemeijer, K., & Stephan, M. (2002). Emergent models as an instructional design heuristic *Symbolizing, modeling and tool use in mathematics education* (pp. 145-169): Springer.
Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: a mathematician on didactics and curriculum theory. *Journal of Curriculum Studies, 32*(6), 777-796.
Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A" proceptual" view of simple arithmetic. *Journal for Research in Mathematics Education*, 116-140.
Greeno, J. (1980). *Development of processes for understanding problems.* Paper presented at the Heidelberg Conference, Germany.
Greer, B. (1993). The mathematical modeling perspective on wor (l) d problems. *The Journal of Mathematical Behavior*.
Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*(1), 76.
Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*(1), 18.
Hoadley, U. (2007). The reproduction of social class inequalities through mathematics pedagogies in South African primary schools. *Journal of Curriculum Studies, 39*(6), 679-706.
Hoadley, U. (2012). What do we know about teaching and learning in South African primary schools? *Education as Change, 16*(2), 187-202.
Hoven, J., & Garelick, B. (2007). Singapore math: Simple or complex? *Educational Leadership, 65*(3), 28.
Huttenlocher, J., Eisenberg, K., & Strauss, S. (1968). Comprehension: Relation between perceived actor and logical subject. *Journal of Verbal Learning & Verbal Behavior*.
Kilpatrick, J., & Swafford, J. (2002). *Helping children learn mathematics*: National Academy Press.
Klein, A. S., Beishuizen, M., & Treffers, A. (1998). The empty number line in Dutch second grades: Realistic versus gradual program design. *Journal for Research in Mathematics Education*, 443-464.
Lesh, R., Carmona, G., Hjalmarson, M., & Mason, G. (2006). *Working group models and modeling.* Paper presented at the Psychology of Mathematics Education.
Lesh, R., & Caylor, B. (2007). Introduction to the special issue: Modeling as application versus modeling as a way to create mathematics. *International Journal of computers for mathematical Learning, 12*(3), 173-194.
Lesh, R., & Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical thinking and learning, 5*(2-3), 157-189.
Lesh, R., & Lehrer, R. (2003). Models and modeling perspectives on the development of students and teachers. *Mathematical thinking and learning, 5*(2-3), 109-129.
Levitin, K. (1982). One is not born a personality. *Moscow: Progress*, 5-24.
Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of educational psychology, 81*(4), 521.

Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of educational psychology, 79*(4), 363.

Mason, J. (2018). Structuring structural awareness: a commentary on Chapter 13 *Building the foundation: whole numbers in the primary grades* (pp. 325-340): Springer, Cham.

Mayer, R. E., & Hegarty, M. (2012). The process of understanding mathematical problems *The nature of mathematical thinking* (pp. 45-70): Routledge.

Mayer, R. E., Lewis, A. B., & Hegarty, M. (1992). Mathematical misunderstandings: Qualitative reasoning about quantitative problems *Advances in psychology* (Vol. 91, pp. 137-153): Elsevier.

Mukhopadhyay, S., & Greer, B. (2001). Modeling with purpose: Mathematics as a critical tool. *Sociocultural research on mathematics education: An international perspective*, 295-311.

Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical thinking and learning, 10*(4), 374-406.

Nesher, P., Greeno, J. G., & Riley, M. (1982). The development of semantic categories for addition and subtraction. *Educational Studies in Mathematics, 13*(4), 373-394.

Nesher, P., & Katriel, T. (1978). *Two cognitive modes in arithmetic word problem solving.* Paper presented at the second annual meeting of the International Group for the Psychology of Mathematics Education, Osnabruck, West Germany.

Ng, S. F., & Lee, K. (2005). How primary five pupils use the model method to solve word problems. *The mathematics educator, 9*(1), 60-83.

Pape, S. J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary educational psychology, 28*(3), 396-421.

Peled, I., & Balacheff, N. (2011). Beyond realistic considerations: modeling conceptions and controls in task examples with simple word problems. *ZDM, 43*(2), 307-315.

Piaget, J. (1965). The stages of the intellectual development of the child. *Educational psychology in context: Readings for future teachers, 63*(4), 98-106.

Piaget, J. (1973). Psychology and epistemology.

Polotskaia, E. (2015). *How elementary students learn to mathematically analyze word problems: The case of addition and subtraction.* McGill University (Canada).

Polotskaia, E. (2017). How the Relational Paradigm Can Transform the Teaching and Learning of Mathematics: Experiment in Quebec. *International Journal for Mathematics Teaching & Learning, 18*(2).

Polotskaia, E., & Savard, A. (2018). Using the Relational Paradigm: effects on pupils' reasoning in solving additive word problems. *Research in Mathematics Education, 20*(1), 70-90.

Polotskaia, E., Savard, A., & Freiman, V. (2015). Duality of mathematical thinking when making sense of simple word problems: theoretical essay. *EURASIA Journal of Mathematics, Science and Technology Education, 11*(2), 251-261.

Polotskaia, E., Savard, A., & Freiman, V. (2016). Investigating a case of hidden misinterpretations of an additive word problem: structural substitution. *European journal of psychology of education, 31*(2), 135-153.

Polya, G. (1945). How to Solve It Princeton Univ. *Press, Princeton, NJ*.

Rabardel, P., & Bourmaud, G. (2003). From computer to instrument system: a developmental perspective. *Interacting with Computers, 15*(5), 665-691.

Rasmussen, C., Zandieh, M., King, K., & Teppo, A. (2005). Advancing mathematical activity: A practice-oriented view of advanced mathematical thinking. *Mathematical thinking and learning, 7*(1), 51-73.

Riley, M. (1979). *The development of children's ability to solve arithmetic word problems.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Riley, M. (1984). Development of children's problem-solving ability in arithmetic.

Savard, A., & Polotskaia, E. (2017). Who's wrong? Tasks fostering understanding of mathematical relationships in word problems in elementary students. *ZDM, 49*(6), 823-833.

Savinainen, A., & Viiri, J. (2008). The force concept inventory as a measure of students conceptual coherence. *International Journal of Science and Mathematics Education, 6*(4), 719-740.

Schollar, E. (2008). Final Report: The primary mathematics research project 2004-2007–Towards evidence-based educational development in South Africa. *Johannesburg: Eric Schollar & Associates*.

Sepeng, P., & Sigola, S. (2013). Making sense of errors made by learners in mathematical word problem solving. *Mediterranean Journal of Social Sciences, 4*(13), 325.

Spaull, N. (2013). South Africa's education crisis: The quality of education in South Africa 1994-2011. *Johannesburg: Centre for Development and Enterprise*, 1-65.

Steffe, L. P., & Cobb, P. (2012). *Construction of arithmetical meanings and strategies*: Springer Science & Business Media.

Tabak, I. (2004). Reconstructing context: Negotiating the tension between exogenous and endogenous educational design. *Educational psychologist, 39*(4), 225-233.

Tabulawa, R. (1997). Pedagogical classroom practice and the social context: The case of Botswana. *International journal of educational development, 17*(2), 189-204.

Tabulawa, R. (2013). *Teaching and learning in context: Why pedagogical reforms fail in Sub-Saharan Africa*: African Books Collective.

Taylor, N. (2021). The dream of Sisyphus: Mathematics education in South Africa. *South African Journal of Childhood Education, 11*(1), 1-12.

Thompson, I. (1999). Mental calculation strategies for addition and subtraction. Part 1. *Mathematics in school, 28*(5), 2-4.

Treffers, A. (1987). Integrated column arithmetic according to progressive schematisation. *Educational Studies in Mathematics, 18*(2), 125-145.

Treffers, A. (1991). Didactical background of a mathematics program for primary education. *Realistic mathematics education in primary school*, 21-56.

Treffers, A. (2001). Grade 1 (and 2): Calculation up to 20. *Children Learn Mathematics: A Learning-Teaching Trajectory with Intermediate Attainment Targets for Calculation with Whole Numbers in Primary School. The Netherlads: Freudenthal Institute, Utrecht University*.

Treffers, A., & Goffree, F. (1985). *Rational analysis of realistic mathematics education–the Wiskobas program.* Paper presented at the Proceedings of the ninth International Conference for the Psychology of Mathematics Education.

Treffers, A., & Moor, E. d. (1990). Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 2: Basisvaardigheden en cijferen: Uitgeverij Zwijsen.

Van den Heuvel-Panhuizen, M. (1998). Realistic Mathematics Education as work in progress. *Theory into practice in Mathematics Education. Kristiansand, Norway: Faculty of Mathematics and Sciences*.

Van den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54*(1), 9-35.

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education *Encyclopedia of mathematics education* (pp. 521-525): Springer.

Venkat, H. (2011). *Wits Maths Connect–Primary Project.* Paper presented at the Discussion document for Community of Practice Forum–Johannesburg.

Venkat, H. (2013). Using temporal range to theorize early number teaching in South Africa. *For the learning of mathematics, 33*(2), 31-37.

Venkat, H., Askew, M., & Morrison, S. (2021). Shape-shifting Davydov's ideas for early number learning in South Africa. *Educational Studies in Mathematics, 106*(3), 397-412.

Venkat, H., Ekdahl, A.-L., & Runesson, U. (2014). Connections and Simultaneity: Analysing South African G3 Part-Part-Whole Teaching. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Venkat, H., & Mathews, C. (2019). Improving multiplicative reasoning in a context of low performance. *ZDM, 51*(1), 95-108.

Venkat, H., & Naidoo, D. (2012). Analyzing coherence for conceptual learning in a Grade 2 numeracy lesson. *Education as Change, 16*(1), 21-33.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. *Addition and subtraction: A cognitive perspective*, 39-59.

Vergnaud, G. (2009). The theory of conceptual fields. *Human development, 52*(2), 83-94.

Vergnaud, G., & Durand, C. (1976). Structures additives et complexité psychogénétique. *Revue française de pédagogie*, 28-43.

Verschaffel, L. (1994). Using retelling data to study elementary school children's representations and solutions of compare problems. *Journal for Research in Mathematics Education, 25*(2), 141-165.

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction, 4*(4), 273-294.

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*(1), 85.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*: Swets & Zeitlinger Lisse.

Verschaffel, L., Van Dooren, W., Greer, B., & Mukhopadhyay, S. (2010). Reconceptualising word problems as exercises in mathematical modelling. *Journal für Mathematik-Didaktik, 31*(1), 9-29.

Webb, P., & Sepeng, P. (2012). Exploring mathematical discussion in word problem-solving. *Pythagoras, 33*(1), 1-8.

Weitz, M., & Venkat, H. (2013). Assessing early number learning: How useful is the Annual National Assessment in Numeracy? *Perspectives in Education, 31*(3), 49-65.

Wright, R. J., Martland, J., & Stafford, A. K. (2006). *Early numeracy: Assessment for teaching and intervention*: Sage.

Xin, Y. P. (2012). *Conceptual model-based problem solving: Teach students with learning difficulties to solve math problems*: Brill Sense.

Xin, Y. P. (2013). *Conceptual model-based problem solving: Teach students with learning difficulties to solve math problems*: Springer Science & Business Media.

Xin, Y. P., Wiles, B., & Lin, Y.-Y. (2008). Teaching Conceptual Model—Based Word Problem Story Grammar to Enhance Mathematics Problem Solving. *The Journal of Special Education, 42*(3), 163-178.

APPENDICES

Appendix A: Detailed descriptions of intervention lessons

Lesson #1	**Date: 20 Oct**	**Theme: Disco fever**		
Relational models as objects	**Relational models as tools**			
Connecting models	*Solving intro context problems*	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
BD introduced with 10 as the whole and 5s as parts. **Translation from BD into NS** demonstrated. **ENL introduced**. Partitions of 10 used to familiarize learners with structure of BD, with translations into a NSs exemplified. **Translating from NS to ENL** exemplified with configurations of partitions of 10. **Family of number facts associated with the BD**	Problems done on the board: **Change Increase Result Unknown problem:** 75 children drank a glass of lemonade**.** Another 35 drank coke. How many children had something to drink? (**N10 strategy introduced**) **Change Increase Result Unknown problem:** There were 45 children dancing**.** Another 15 got up to dance to 'The Wonder Boys'. How many children were dancing? (**A10 strategy introduced**)	Established connections between the two **Change Increase Result Unknown problem**s done in the previous stage. Difference between **N10 and A10** made explicit.	Third *intro context problem* done by learners in pairs: **Change Increase Result Unknown problem:** There were 76 bags of crisps on the table. Sam unpacked another 37 bags. How many bags of crisps were out on the table?	Notion of a Change problem summarised. Idea of coherence between BD, NS and ENL made explicit. Exercise set 13A for homework

<table>
<tr><td colspan="5">Lesson #2   Theme: Sunflowers   Date: 21 Oct</td></tr>
<tr><td>Relational models as objects</td><td colspan="4">Relational models as tools</td></tr>
<tr><td>Connecting models</td><td>Solving intro context problems</td><td>Linking-up intro context problems</td><td>Follow-up problems</td><td>Wrap-up</td></tr>
<tr>
<td>Recap on BD and NL with 20 as the whole and 10s as parts.

Translation from BD into NS demonstrated.

Introduced partitions of 5 & 6, and used them for more familiarity with structure of BD. Introduced partitions of 20, and had learners completing the list, and translating them into a NS and ENL.</td>
<td>Problems done on the board:

Change Increase Result Unknown problem: Aysha planted sunflower seeds in her garden. One day she saw that 56 seeds had sprouted. A week later another 34 seeds had sprouted. How many seeds had sprouted altogether? (A10 strategy used)

Translating from NS to ENL exemplified with configurations of partitions of 10.</td>
<td>Problem-solving protocol established as beginning with a word problem in context interpreted into a BD, translated into a NS, and finally into a ENL</td>
<td>Second and third intro context problem done by learners in pairs, and then by the teacher on the board :

Change Decrease Result Unknown problem: Aysha was happy. 82 of her sunflowers had come up. She picked 9 of them to give to her sister. How many sunflowers were left in the garden? (A10 strategy used)

Change Decrease Result Unknown problem: Aysha was happy. There were 73 sunflowers still growing in the garden.
69 sunflowers wilted in the midday sun. How many sunflowers did not wilt?
(N10C strategy introduced)</td>
<td>Established connections between the two Change Decrease Result Unknown problems done in the previous stage.

Difference between N10, A10 & N10C made explicit.

Exercise set for homework</td>
</tr>
</table>

Lesson #3 **Theme: Sunflowers** (continued) **Date: 27 Oct**				
Relational models as objects	**Relational models as tools**			
Connecting models	*Solving intro context problems*	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Recap on the structure of the BD and how the parts come together to make the whole**.** Example with 12 and 13 as parts used to show 1. When parts are given, we add 2. Part + part = whole 3. when we add we jump to the right 4. ENL used to verify that the whole is 25 Another example with 12 as parts and 25 as a whole used to show 1. When the whole and a part are given, we subtract the part from the whole 2. Whole – part = part.	Problems done on the board: Exercise 14A **Change Increase Result Unknown problem:** Pierre Rabbit planted some lettuce seeds. One day he counted that 37 of the seeds had sprouted. A week later another 48 seeds had sprouted. How many lettuce seeds had sprouted altogether? **(N10: differentiating between jumping in units of 10 vs in multiples of 10)** **Change Decrease Result Unknown problem:** Pierre Rabbit counted that 93 of his lettuces were ready to be picked. During the night, Farmer Brown came and took 29 of the	Established connections between the **Change Increase Result Unknown** and the **Change Decrease Result Unknown problem** Established connections between jumping in **units of 10 (N10)** vs jumping **in multiples of**	Exercise 14A Q2: **Change Decrease Result Unknown problem:** There were 44 lettuces left growing in the field. Caterpillars nibbled 19 of them. How many lettuces did the caterpillars leave alone? (Consolidating jumping in **multiples of 10** whilst looking to **compensate).**	Notion of jumping **in multiples of 10** rather than jumping in **units of 10** encouraged. Exercise set 14B for homework

3. when we subtract we jump to the left 4. ENL used to verify that the missing part was 13.	lettuces. How many lettuces did Pierre Rabbit find the following morning? (**Contrasting N10 and N10C)**	**10 (N10)** vs jumping in **multiples of 10** whilst looking to **compensate (N10C).**	Q2: Write a Change problem to go with this calculation **62 – 45** =	

Lesson #4 Date: 28 Oct Theme: Country Capers			
Relational models as tools			
Solving intro context problems	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Problems done on the board: Exercise 14B **Change Increase Result Unknown problem:** Pierre Rabbit planted some lettuce seeds. One day he counted that 34 of the seeds had sprouted. A week later another 45 seeds had sprouted. How many lettuce seeds had sprouted altogether? (**N10: jumping in multiples of 10)** **Change Decrease Result Unknown problem:** Pierre Rabbit counted that 53 of his lettuces were ready to be picked. During the night, Farmer Brown came and took 31 of the lettuces. How many lettuces did Pierre Rabbit find the following morning? (**N10: jumping in multiples of 10)** Write a Change problem to go with this calculation **62 – 45** =	Established connections between the **Change Increase Result Unknown** and the **Change Decrease Result Unknown problem** Established connections between jumping in **units of 10 (N10)** vs jumping **in multiples of 10 (N10)** vs jumping in **multiples of 10** whilst looking to **compensate (N10C).**	1st *intro context problem* done on the board; 2nd and 3rd done by learners individually and then by the teacher on the board: Introduced **Compare problems** using differential heights of three leaners in the class. **Recap on the problem-solving protocol** as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL.** **Compare Problem #1**: Jolene and Tammy were taking part in the welly-tossing competition. Jolene threw the welly 59 metres. Tammy threw it 74 metres. How much further than Jolene did Tammy throw the welly? **(N10C)** **Compare Problem #2:** Dolly and Kenny were on an egg hunt.	Language of 'how much further' and 'how many more' unpacked as indicative of a Compare problem. Difference between **N10** & **N10C** made explicit following two learners' writing of their responses on the board. Exercise set 15A given as homework

		Dolly had found 164 hidden eggs. Kenny had found 49 eggs. How many more eggs had Dolly found than Kenny?	

<table>
<tr><td colspan="4">Lesson #5   Theme: Theme: Country Capers (continued)
Date: 03 Nov</td></tr>
<tr><td colspan="4" align="center">Relational models as tools</td></tr>
<tr><td>Solving intro context problems</td><td>Linking-up intro context problems</td><td>Follow-up problems</td><td>Wrap-up</td></tr>
<tr><td>2nd and 3rd Country Capers intro context problem done on the board:

Compare Problem #2: Dolly and Kenny were on an egg hunt. Dolly had found 164 hidden eggs. Kenny had found 49 eggs. How many more eggs had Dolly found than Kenny? (A10)

Problem-solving protocol as beginning with a word problem in context interpreted into a BD, translated into a NS, and finally into a ENL recapped.

Compare Problem #3: John-boy was trying to beat the domino tower record. There was a box of 575 dominoes. So far, Slim Jim had set the record with a tower 158 dominoes high. John-boy's tower was 36 dominoes tall. How many more dominoes</td><td>Language of 'how much taller' for heights: 'how much further' for distances; 'how much heavier' for weights; and 'how much older' for age unpacked as indicative of a Compare problem.

Learners' attention directed to the presence of additional irrelevant data which learners need to factor out. (the irrelevance of the 575 dominoes in the solving of the John-boy problem used as example)</td><td>Exercise 15A
Compare Problem #1: Lisa-Marie and Shelby were taking part in the welly-tossing competition. Lisa-Marie threw the welly 83 metres. Shelby threw it 69 metres. The field was 120 metres long. How much further than Shelby did Lisa-Marie throw the welly?

Compare Problem #2: There were 350 eggs hidden in the egg hunt.
Donny had found 70 eggs. Garth had found 41 eggs. How many fewer eggs had Garth found than Donny?</td><td>Language of 'how much further' and 'how many more' re-emphasized as indicative of a Compare problem.

Exercise set 15B for homework</td></tr>
</table>

did John-boy need to add to his tower to make it equal to Slim Jim's?? **(N10)**			

<table>
<tr><td colspan="5">Lesson #6 Theme: Travel Survey
Date: 04 Nov</td></tr>
<tr><td>Relational models as objects</td><td colspan="4">Relational models as tools</td></tr>
<tr><td>Connecting models</td><td>Solving intro context problems</td><td>Linking-up intro context problems</td><td>Follow-up problems</td><td>Wrap-up</td></tr>
<tr><td>Recap on the problem-solving protocol as beginning with a word problem in context interpreted into a BD, translated into a NS, and finally into an ENL

Using Exercise 15B, difference between semantics reflecting a Compare problem contrasted from the semantics of a Change problem.

Example with Change problem #

1. Part + part = whole
2. Add means jump to the right
3. NL used to verify that the whole is 51

Example with Compare problem #

1. The whole and a part are given,</td><td>Problems done on the board: Exercise 15B

All five compare word problems in exercise 15B are interpreted into a BD. No NS and ENL translations, thus emphasizing the interpretation phase of the problem-solving cycle.

Once again, learners' attention directed to the presence of additional irrelevant data which learners need to factor out.</td><td>Exercise really on how to tell from the semantics of the word problem if a problem is a Compare problem.

Language of 'how much further' and 'how many more' re-emphasized as indicative of a Compare problem.</td><td>Interpretation of the three Travel Survey intro context problem done on the board:

All three intro context compare word problems are interpreted without translating into an ENL, thus focusing exclusively on the interpretation phase of the problem-solving cycle.

2nd problem done by learners individually Compare Problem #2: In Year 6, 19 children have scooters and 36 children have bikes. How many more children have bikes than have scooters?</td><td>Notion of '6 more than 12' related to a adding on 6 to 12.

Notion of '11 fewer than 25' related to subtracting 11 from 25.

Reiterated that Compare problems invoke subtraction.

Exercise set 16A for homework</td></tr>
</table>

2. Subtract the part from the whole: Whole – part = part. 3. subtract means jump to the left				

Lesson #7	**Theme: Safari Park** **10 Nov**			**Date:**
Relational models as objects	**Relational models as tools**			
Connecting models	*Solving intro context problems*	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Recap on the problem-solving protocol as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL** Recap on the difference between semantics reflecting a **Compare** problem contrasted from the semantics of a **Change** problem. Language associated with each category revisited.	1st *intro context problem* done on the board: In Nekuru Safari Park there are 83 flamingoes. A loud bang startled 78 of them and they flew off. The other flamingoes are all deaf, so they did not fly off. How many deaf flamingoes are there? **(Change: A10)**	Exercise really on how to tell from the semantics of the word problem if a problem is a Compare problem. Language of 'how much further' and 'how many more' re-emphasized as indicative of a Compare problem.	2nd and 3rd done by learners individually and then by the teacher on the board: 2nd done by learners individually: There are 19 baby warthogs in the Safari Park. There are 65 adult warthogs. How many more of the warthogs are adults than babies? **(Compare: A10 & N10C)** 3rd also done by learners individually: **(Compare:** jumping in **units of 10 (N10)** vs jumping **in multiples of 10 (N10)**	All six word problems in exercise 17A are interpreted as compare or Change problem semantically. Reiterated that Compare problems invoke subtraction, and subtraction relates to 'backward jumps' on the ENL. Exercise set 17A for homework

Reiterated that Compare problems invoke subtraction Another practical example involving heights of two learners in the class used (110 – 85 = 25) Adding we jump to the right, subtracting we jump to the right on the ENL.				

Lesson #8 **Date: 11 Nov**	**Theme: Safari Park (**continued**)**		
Relational models as tools			
Solving intro context problems	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Problem-solving **protocol** as beginning with a word problem in context interpreted into a BD, translated into a NS, and finally into a ENL **recapped**. All six problems in exercise set 17A done by the learners individually\in pairs: **Compare Problem #1: (N10 & A10)** **Compare Problem #2: (N10C)** **Compare Problem #3: (N10)**	Language of 'how much taller' for heights: 'how much further' for distances; 'how much heavier' for weights; and 'how much older' for age unpacked as indicative of a Compare problem.	Two problems from Exercise 17A done by learners in pairs and then by the teacher on the board: **Compare Problem #4:** On the plain, 90 zebras were keeping a look out for lions, while 53 zebras were grazing. How many fewer zebras were grazing than keeping a look out? **(N10)** **Compare Problem #5:** There were 124 male giraffes and 151 female giraffes. How many more female giraffes than male ones were there? **(N10)**	Language of 'how much further' and 'how many more' re-emphasized as indicative of a Compare problem. Exercise set 17B for homework

Lesson #9	Theme: Jurassic Park 18 Nov			Date:
Relational models as objects	**Relational models as tools**			
Connecting models	*Solving intro context problems*	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Recap on the problem-solving protocol as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into a E**NL** Recap on the difference between semantics reflecting a **Compare** problem contrasted from the semantics of a **Change** problem. Language associated with each category revisited. Family of number facts associated with a single combination of number triple (example used involved 20, 30, 50), modelled on a ENL **Combine problems introduced.**	All three *intro context problem* done on the board: **Change Increase Result Unknown problem**: 35 vultures were feeding off the body of a tyrannosaurus. Another 15 vultures joined them. How many vultures were feeding? **(A10)** **Combine problem**: 40 tyrannosauruses got locked in battle with another 70 tyrannosauruses**.** How many tyrannosauruses were in the battle? **(Combine: N10)** **Compare Problem:** 80 adult triceratops and 32 young triceratops were trampling through a forest. How many	Exercise really on how to tell from the semantics of the word problem if a problem is a Compare, Combine or Change problem.	Two problems from Exercise 18A done by learners in pairs and then by the teacher on the board: **Compare Problem #1:** 46 diplodocuses were grazing on the open plain. Another 29 diplodocuses joined them. How many diplodocuses were grazing? **(**Contrasted using **N10** vs **N10C**) **Compare Problem #2:** 46 female pterydactyls and 29 male pterydactyls	Reiterated that there are three different kinds of additive relations word problems: change, combine and compare. Also, that Compare problems invoke subtraction, and subtraction relates to 'backward jumps' on the ENL.

Another practical example involving comparing heights of two learners in the class used (152 – 98 = 54)	more adult triceratops were there? **(Compare: N10)**		were gliding through the skies. How many more female pterydactyls were there than male ones? **(N10)**	

Lesson #10 **Theme: Sunflowers (revisited)** **Date: 24 Nov**				
Relational models as objects	**Relational models as tools**			
Connecting models	*Solving intro context problems*	*Linking-up intro context problems*	*Follow-up problems*	*Wrap-up*
Recap on the problem-solving protocol as beginning with a word problem in context interpreted into a **BD**, translated into a **NS**, and finally into an E**NL** Recap on the difference between semantics reflecting a **Compare** problem contrasted from the semantics of a **Change** problem. **Combine problems recapped.** Language associated with each category revisited. Family of number facts associated with a single combination of number triple (5, 20, 25), two modelled on ENLs.	3rd *intro context problem* done on the board; **Change Decrease Result Unknown problem:** Aysha was happy. There were 73 sunflowers still growing in the garden**.** 69 sunflowers wilted in the midday sun. How many sunflowers did not wilt? **(A10 & N10C)**	Exercise really on how to telling from the semantics of the word problem if a problem it was a Change, Combine or a Compare problem. Questions like 'how much further' and 'how many more' re-	Problem #2 from Exercise 14A done by learners individually/in pairs and then by the teacher on the board: **Change Decrease Result Unknown problem:** Pierre Rabbit counted that 93 of his lettuces were ready to be picked. During the night, Farmer Brown came and took 29 of the lettuces. How many lettuces did Pierre Rabbit find the following morning? **(N10C)**	Reiterated that there are three different kinds of additive relations word problems: change, combine and compare. Also, that Compare problems invoke subtraction, and subtraction relates to 'backward jumps' on the ENL.

		emphasized as indicative of a Compare problem.		

Appendix B: Ethics approval

Wits School of Education

27 St Andrews Road, Parktown, Johannesburg, 2193 Private Bag 3, Wits 2050, South Africa. Tel: +27 11 717-3064 Fax: +27 11 717-3100 E-mail: enquiries@educ.wits.ac.za Website: www.wits.ac.za

25 February 2015

Student Number: 9108968T

Protocol Number: 2015ECE003D

Dear Herman Tshesane

Application for Ethics Clearance: Doctor of Philosophy

Thank you very much for your ethics application. The Ethics Committee in Education of the Faculty of Humanities, acting on behalf of the Senate has considered your application for ethics clearance for your proposal entitled:

Intervening to develop the inter-related number knowledge needed for facility with additive relation problems through a representational approach

The committee recently met and I am pleased to inform you that clearance was granted. However, there were a few small issues which the committee would appreciate you attending to before embarking on your research.

The following comments were made:

- The letter of information to the pupils who will participate in the study is too complex for them to understand. Re-write this information letter to make it more understandable to pupils at this level of schooling.
- The information letters should be checked for language, e.g., "doing a research" is incorrect please correct it and any other language incorrect language expressions.

Please use the above protocol number in all correspondence to the relevant research parties (schools, parents, learners etc.) and include it in your research report or project on the title page.

The Protocol Number above should be submitted to the Graduate Studies in Education Committee upon submission of your final research report.

All the best with your research project.
Yours sincerely,

MMasete

Wits School of Education
011 717-3416

Cc Supervisor: Prof Hamsa Venkatakrishnan

Appendix C: GDE approval for WMC-P project

Reference No: D2011/69

UMnyango WezeMfundo
Department of Education

Lefapha la Thuto
Departement van Onderwys

Enquiries: Diane Buntting (011) 843 6503

Date:	31 January 2011
Name of Researcher:	Professor H. Venkatakrishnan
Address of Researcher:	Room 2,
	WMC corridor,
	Marang Block;
	WITS School of Education
Telephone Number:	011 717 3742 / 082 099 1967
Fax Number:	0865088324
Email address:	hamsa.venkatakrishnan@wits.ac.za
Research Topic:	Improving the teaching and learning of primary school numeracy/mathematics
Number and type of schools:	TEN PRIMARY Schools
District/s/HO	Johannesburg East

Re: Approval in Respect of Request to Conduct Research

This letter serves to indicate that approval is hereby granted to the above-mentioned researcher to proceed with research in respect of the study indicated above. The onus rests with the researcher to negotiate appropriate and relevant time schedules with the school/s and/or offices involved to conduct the research. A separate copy of this letter must be presented to both the School (both Principal and SGB) and the District/Head Office Senior Manager confirming that permission has been granted for the research to be conducted.

Permission has been granted to proceed with the above study subject to the conditions listed below being met, and may be withdrawn should any of these conditions be flouted:

1. *The District/Head Office Senior Manager/s concerned must be presented with a copy of this letter that would indicate that the said researcher/s has/have been granted permission from the Gauteng Department of Education to conduct the research study.*
2. *The District/Head Office Senior Manager/s must be approached separately, and in writing, for permission to involve District/Head Office Officials in the project.*
3. *A copy of this letter must be forwarded to the school principal and the chairperson of the School Governing Body (SGB) that would indicate that the researcher/s have been granted permission from the Gauteng Department of Education to conduct the research study.*

Office of the Chief Director: Information and Knowledge Management
Room 501, 111 Commissioner Street, Johannesburg, 2000 P.O.Box 7710, Johannesburg, 2000
Tel: (011) 355-0809 Fax: (011) 355-0734

4. *A letter / document that outlines the purpose of the research and the anticipated outcomes of such research must be made available to the principals, SGBs and District/Head Office Senior Managers of the schools and districts/offices concerned, respectively.*
5. *The Researcher will make every effort obtain the goodwill and co-operation of all the GDE officials, principals, and chairpersons of the SGBs, teachers and learners involved. Persons who offer their co-operation will not receive additional remuneration from the Department while those that opt not to participate will not be penalised in any way.*
6. *Research may only be conducted after school hours so that the normal school programme is not interrupted. The Principal (if at a school) and/or Director (if at a district/head office) must be consulted about an appropriate time when the researcher/s may carry out their research at the sites that they manage.*
7. *Research may only commence from the second week of February and must be concluded before the beginning of the last quarter of the academic year.*
8. *Items 6 and 7 will not apply to any research effort being undertaken on behalf of the GDE. Such research will have been commissioned and be paid for by the Gauteng Department of Education.*
9. *It is the researcher's responsibility to obtain written parental consent of all learners that are expected to participate in the study.*
10. *The researcher is responsible for supplying and utilising his/her own research resources, such as stationery, photocopies, transport, faxes and telephones and should not depend on the goodwill of the institutions and/or the offices visited for supplying such resources.*
11. *The names of the GDE officials, schools, principals, parents, teachers and learners that participate in the study may not appear in the research report without the written consent of each of these individuals and/or organisations.*
12. *On completion of the study the researcher must supply the Director: Knowledge Management & Research with one Hard Cover bound and an electronic copy of the research.*
13. *The researcher may be expected to provide short presentations on the purpose, findings and recommendations of his/her research to both GDE officials and the schools concerned.*
14. *Should the researcher have been involved with research at a school and/or a district/head office level, the Director concerned must also be supplied with a brief summary of the purpose, findings and recommendations of the research study.*

The Gauteng Department of Education wishes you well in this important undertaking and looks forward to examining the findings of your research study.

Kind regards

31 January 2014

Shadrack Phele MIRMSA
[Member of the Institute of Risk Management South Africa]
CHIEF EDUCATION SPECIALIST: RESEARCH COORDINATION

The contents of this letter has been read and understood by the researcher.	
Signature of Researcher:	
Date:	

Appendix D – Participants information letters and consent forms

INFORMATION LETTER: LEARNERS

DATE: 29th July, 2015

Dear Learner

My name is Herman Tshesane. I am studying how a certain way of teaching may help us as teachers to teach better.

I would like to try out this way of teaching with you and your classmates. The lessons will be during normal class time and they will be taking place in your registered class. We will have about 10 lessons, with two lessons taught per week. I write to ask you to consider agreeing to be part of the group of learners who will participate. We plan to have the lessons sometime in the third term, in October. We will be learning how to solve word problems.

In order to check our progress, we will be writing a test immediately before the first lesson and immediately after the final lesson. The marks you get for the tests will **not** count towards your final mark.

Should you agree I will need your permission to video-tape the lessons, so kindly complete and sign the form at the back of this page and have your son/daughter return it to his/her class teacher and she will keep them safe for me to collect.

I promise not to use your son/daughter's name when I report on this work. Also, the recordings of the lessons and children's scripts will be kept safely away, where only members of the research team and I can access.

It is important for you to know that you are free to agree or disagree as it is not compulsory.

Thank you

____

SIGNATURE

NAME: Herman Tshesane

EMAIL : herman.tshesane@wits.ac.za or mohlogopela@gmail.com

TEL NUMBER: 011-7173412 (office) or 0824119164 (cell)

Learner's Consent Form

I agree to write the tests.
I agree to participate in the video-recorded lessons.
Name and Surname: ____ Sign:____ Date: ____

INFORMATION LETTER: Parent/Guardian/Carer,

DATE: 29th July, 2015

Dear Parents/Guardian,

My name is Herman Tshesane. I am studying how a certain way of teaching may help us as teachers to teach mathematics to children in a better way.

I would like to try out this way of teaching with your son/daughter and his/her classmates. The lessons will be during normal class time and they will be taking place in their registered class. We will have about 10 lessons, with two lessons taught per week. I write to ask you to consider agreeing to have your son/daughter be part of the group of learners who will participate. We plan to have the lessons sometime in the third term, in October. We will be learning how to solve word problems and hope that son/daughters will benefit from attending.

In order to check the progress of the leaners, we will be writing a test immediately before the first lesson and immediately after the final lesson. The marks they get for the tests will **not** count towards their final mark.

Should you agree I will need your permission to video-tape the lessons, so kindly complete and sign the form at the back of this page and have your son/daughter return it to his/her class teacher and I will collect it from her.

I undertake not to use your name, or the name of your child when reporting on this work. Also, the recordings of the lessons and your answers to the tests will be kept safely away, where only my professors and I can access.

It is important for you to know that you are free to agree or disagree to participate. The choice is yours.

Thank you

____

SIGNATURE

NAME: Herman Tshesane ***TEL NUMBER***: 011-7173412 (office) or 0824119164 (cell)

EMAIL : herman.tshesane@wits.ac.za or mohlogopela@gmail.com

Consent Forms – Parent/Guardian

I agree that my child can take part in the tests.
I agree to my child's participation in the video-recorded lessons.
Name: ____ Signature____ Date: ____

INFORMATION SHEET: TEACHER

DATE: 29th July, 2015

Dear Teacher,

My name is Herman Tshesane. I am studying how a certain way of teaching may help us as teachers to teach mathematics to children in a better way.

I would like to try out this way of teaching with your class. The lessons will be during normal class time and they will be taking place in your registered class. We will have about 10 lessons, with two lessons taught per week. I write to ask you to consider agreeing to using your class to conduct this teaching experiment. We plan to have the lessons sometime in the third term, in October, we will be solving word problems. I have no doubt that your class will benefit from the exercises.

In order to check the progress of the leaners, we will be writing a test immediately before the first lesson and immediately after the final lesson. The marks they get for the tests will **not** count towards their final mark.

Should you agree I will need your permission to video-tape the lessons, so kindly complete and sign the form at the back of this page and will collect it.

I promise not to use your name or the names of the learners in your class when I report on this work. Also, the recordings of the lessons and your answers to the tests will be kept safely away, where only my professors and I can access.

It is important for you to know that you are free to agree or disagree to participate. The choice is yours.

Thank you

____

NAME: Herman Tshesane

EMAIL : herman.tshesane@wits.ac.za or mohlogopela@gmail.com

TEL NUMBER: 011-7173266 (office) or 0824119164 (cell)

Teacher's Consent Form

Please fill in and return the reply slip below indicating your willingness to be interviewed and to have your interview audio-taped for my research project:

I hereby give my consent to availing my class for this teaching experiment.
I hereby give my consent to having my learners sit for the tests
I hereby give my consent to having the lessons video-recorded.
I hereby give my consent to being interviewed and to have the interview audio-taped
Print Name: ____   Sign: ____   Date: __________________________________

INFORMATION SHEET: PRINCIPAL

DATE: 29th July, 2015

Dear Principal,

My name is Herman Tshesane. I am studying how a certain way of teaching may help us as teachers to teach mathematics to children in a better way.

I would like to try out this way of teaching with your class. The lessons will be during normal class time and they will be taking place in your registered class. We will have about 10 lessons, with two lessons taught per week. I write to ask you to consider agreeing to having me teach one of your classes as part of a teaching experiment. We plan to have the lessons sometime in the third term, in October, we will be solving word problems. I have no doubt that the learners in the class will benefit from the exercises.

In order to check the progress of the leaners, we will be writing a test immediately before the first lesson and immediately after the final lesson. The marks they get for the tests will **not** count towards their final mark.

Should you agree I will need your permission to conduct the research at your school and to video-tape the lessons, so kindly complete and sign the form at the back of this page and I will collect it.

I undertake not to use your name, the school name, the name of the teacher whose class we will be teaching as part of this teaching experiment, nor the names of the learners in this class when we report on this work. Also, the recordings of the lessons and your answers to the tests will be kept safely away, where only my professors and I can access.

It is important for you to know that you are free to agree or disagree to participate. The choice is yours.

Thank you

____

NAME: Herman Tshesane

EMAIL : herman.tshesane@wits.ac.za or mohlogopela@gmail.com

TEL NUMBER: 011-7173412 (**office)** or 0824119164 (**cell)**

Principal's Consent Form

I hereby give my consent to availing one of my Grade 3 classes for this teaching experiment.
I hereby give my consent to having the learners in that class sit for the tests
I hereby give my consent to having the lessons video-recorded.
I hereby give my consent to being interviewed and to have the interview audio-taped
Print Name: ____ Sign: ____ Date: _________________________________

Turnitin Originality Report

Processed on: 01-Jul-2021 8:48 AM SAST
ID: 1614468576
Word Count: 64193
Submitted: 1

Herman Tshesane - PhD submission - 9108968T - RELOADED-1.pdf By Herman Tshesane

Similarity Index	Similarity by Source	
13%	Internet Sources:	11%
	Publications:	7%
	Student Papers:	2%

3% match (Internet from 14-Jul-2020)
http://wiredspace.wits.ac.za/bitstream/handle/10539/14924/MASTER%20MASTER_%20corrected_final.pdf?isAllowed=y&sequence=1

1% match (Internet from 06-Apr-2021)
http://wiredspace.wits.ac.za/bitstream/handle/10539/23036/2016%2005%2005%20Roberts%20PhD%20thesis%20FINAL.pdf?isAllowed=y&sequence=3

< 1% match (Internet from 14-Jul-2020)
http://wiredspace.wits.ac.za/bitstream/handle/10539/21744/Project%20Corrections%20September%205%20%20%20%282%29.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 14-Jul-2020)
http://wiredspace.wits.ac.za/jspui/bitstream/10539/13416/1/MSc%20Research%20Report%20-%20Final.pdf

< 1% match (Internet from 14-Jul-2020)
http://wiredspace.wits.ac.za/jspui/bitstream/10539/21024/1/Revised%20final%20report%202.pdf

< 1% match (Internet from 06-Apr-2021)
http://wiredspace.wits.ac.za/bitstream/handle/10539/9105/Mayonga-Thesis%202010-06-27%20Final%20Corrections%20post%20examination.pdf?isAllowed=y&sequence=1

< 1% match ()
Tshesane, Herman Makabeteng.. "Exploring Grade 4 learners' use of models and strategies for solving addition and subtraction problems", 2014

< 1% match ()
Abdulhamid, Lawan. "Primary mathematics in-service teaching development: elaborating 'in-the-moment'", 2016

< 1% match ()
Groenewald, Wendy. "The extent to which grade 7 educators in a full-service school apply differentiated instruction to promote inclusive teaching and learning", 2016

< 1% match ()
Alves, Ubiratã Kickhöfel, Luchini, Pedro Luis. "Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinean learners of English", 2017

< 1% match ()
Ogino, Masayoshi. "Modified Output in Response to Clarification Requests and Second Language", The University of Waikato, 2008

< 1% match ()
Robotti, E. "Dynamic representations for algebraic objects available in AlNuSet: how develop meanings of the notions involved in the equation solution.", 2014

< 1% match ()
Stott, Deborah Ann. "Learners' numeracy progression and the role of mediation in the context of two after school mathematics clubs", Faculty of Education, Education, 2015

< 1% match ()
Holmes, Haly Karen. "The prevalence and impact of oral lesions on the quality of life in persons with epidermolysis bullosa", 2010

< 1% match ()
Pithmajor, Aphane Koko. "Grade 9 mathematics learners' strategies in solving numbers pattern problems", 2019

< 1% match ()
De Villiers, Lidemari. "The concurrent development of mathematical modelling and engineering technician competencies of first-year engineering technician students", Stellenbosch : Stellenbosch University

< 1% match ()
Alexander, Michele. "Understanding pedagogic shifts from concrete to abstract conceptions of number.", 2015

< 1% match ()
Chikiwa, Samukeliso. "An investigation into the mathematics knowledge for teaching required to develop grade 2 learners' number sense through counting", Faculty of Education, Education, 2017

< 1% match ()
Visser, Christiaan Jacobus Du Toit. "A conceptual analysis of the quality and utilisation of school textbooks in classrooms and its relevance for teacher education", 2016

< 1% match ()