

https://dspace.org/

Energy storage the unrepentant attentionseeker

Kenneth Ozoemena

https://hdl.handle.net/10539/39433
SCOPUS

Downloaded from a DSpace repository. https://dspace.org/

Routledge
Taylor & Francis Group

African Journal of Research in Mathematics, Science and Technology Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rmse20

Coding Additive Word Problem-solving to see Shifts Around an Intervention

Herman Tshesane & Hamsa Venkat

To cite this article: Herman Tshesane & Hamsa Venkat (2023) Coding Additive Word Problem-solving to see Shifts Around an Intervention, African Journal of Research in Mathematics, Science and Technology Education, 27:3, 324-334, DOI: 10.1080/18117295.2023.2226548

To link to this article: https://doi.org/10.1080/18117295.2023.2226548

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 19 Jul 2023.

Submit your article to this journal

Article views: 31

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rmse20

African Journal of Research in Mathematics, Science and Technology Education, 2023
Vol. 27, No. 3, 324–334, https://doi.org/10.1080/18117295.2023.2226548
© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Check for updates

Coding Additive Word Problem-solving to see Shifts Around an Intervention

Herman Tshesane [a*], and Hamsa Venkat[a,b]

[a] *Wits School of Education, University of the Witwatersrand*
[b] *CASTeL, Dublin City University*
Corresponding author. Email: Herman.Tshesane@wits.ac.za

Evidence in South Africa of poor performance on word problems points to particular challenges among learners relating to difficulties with, on the one hand, making sense of mathematical word problems and, on the other, moving beyond inefficient counting based calculation methods. While many studies in mathematics education in South Africa have tended to focus on one or the other issue, in this paper our focus is on a coding model that brings both interpretation of word problems and calculation into view simultaneously. This coding model was devised to capture and summarise learner work on solving additive relations word problems across pre-, post- and delayed post-tests around an intervention focused on improving additive word problem-solving. The coding model allowed for initial quantitative summarising of learner responses in ways that facilitated the tracking of shifts in learners' interpretations of the problem, their calculations and answers, while also pointing to salient avenues for more in-depth qualitative analysis. We argue through the analysis that the coding scheme afforded a holistic view of the profiles of learners' engagement across the three tests, thus opening a path away from the tendency towards the single-focus of analyses of word problems seen in many prior studies.

Keywords: *Additive word problems; models & modelling; problem-solving; calculation; coordinating models; intervention*

1. Introduction

In this paper, our focus is on a coding scheme that was devised to capture and provide an initial quantitative summary of key patterns and shifts in learners' work with additive word problems prior to, post and delayed-post an intervention lesson sequence. The phenomenon of interest in the intervention study was additive word problem-solving. The study was designed by the first author with the objective of addressing two problems that have been widely reported in mathematics in South Africa. First, there is widespread evidence of poor performance on mathematical word problems, with learner difficulties linked with challenges in making sense of the problem situation (Sepeng, 2014). Second, and identified recurrently in primary mathematics education, there are graphic images of learners using inefficient counting-based methods to work out the answer to calculations well into the middle grades (Schollar, 2008).

While both of these problems have an extensive South African evidence base, they have tended to be written about rather separately. Sepeng's (2014) study attended to the ability of learners to make sense of word problems set in the real world, and noted learners' tendency to relegate reality from their solution processes when solving these word problems. More recent South African research has also focused on learner sense-making of word problems with a focus on the extent of willingness to create a diagrammatic representation, and noting that this aspect was related to successful problem-solving (Mudaly &

Narriadoo, 2023). This emphasis on 'externalising' understandings of word problems is also seen in the international literature, for example, in Björn et al.'s (2019) intervention using 'think-aloud' protocols to support young learners to verbalise their thinking. On the counting-based approaches side, a raft of studies have noted the prevalence of counting to work out the answer to number problems (e.g. Hoadley, 2012; Venkat et al., 2021) with studies of teaching indicating the ongoing advocacy of counting by Foundation Phase teachers across Grades 1–3 (Ensor et al., 2009). However, solving word problems requires both skills: interpreting the relationship between the quantities in the problem situation, and, once the operational relation has been established, efficiently calculating the missing value.

Polotskaia's (2017) Relational Paradigm (RP) includes attention to both 'interpretation' (representing the problem situation in terms of its relational structure in a numerical formulation) and 'calculation' (using the numerical formulation to calculate the answer). With this theory underpinning our coding model, we were further able to explore the co-ordinations between these two phases in learner working, co-ordinations that are viewed as necessary for successful problem-solving (Giroux & Ste-Marie, 2001). Following an introduction of the RP theory that frames our coding scheme's capture of the 'interpretation–calculation–answer' phases of working, we briefly describe the intervention in which this coding was employed as an initial approach, how it was applied to learner responses on the three test administrations, and provide a summary of the outcomes.

In this paper, our focus is on the coding scheme used to summarise learner test responses around a theoretically informed intervention that was devised to attend to both the interpretation and calculation phases. This approach to coding allowed us to see patterned shifts in learner working over time by summarising a relatively large dataset (39 learners each responding to 12 items in each of the three tests, so $468 \times 3 = 1404$ responses in total) in ways that pointed to the phenomena of interest to be studied in more qualitative depth. The quantitative analyses were also useful for summarising shifts in outcomes in the context of the critique that too many education interventions fail to include rigorous assessment of their impact (Besharati et al., 2021). Explicating the quantitative coding is also important to supporting replication and expansion studies in the cases where outcomes show promise. This leads us into a discussion of the following issues:

- What does this coding methodology make visible in terms of:
  - differences in learner responses across the pre-, post- and delayed post-tests, and through this, any shifts in overall outcomes; and
  - coordination of models across the Interpretation and Calculation Phases of problem-solving?
- What pointers for qualitative exploration does the coding methodology point to?

We conclude the paper with a consideration of why we consider this kind of coding approach useful in the South African context, and particularly so around intervention studies.

2. Theory: The Relational Paradigm

A number of theories attend to stages of word problem-solving. Amongst the most well known is the Dutch Realistic Mathematics Education (RME) theory. Key features of RME are (a) a focus on 'emergent modelling': modelling that the theory argues children engage in spontaneously when provided with real/realistic situations that are familiar, and (b) attention to progressive mathematisation—ways of progressively engaging with mathematical concepts that move towards more formal and conventional mathematical forms. Studies in the RME tradition frequently study children's emergent models within the 'horizontal mathematizing' phase and then children's further work extending, refining and using their models for calculating answers in the 'vertical mathematizing' phase. In the domain of additive tasks, an example of progressive mathematisation 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).

A key problem noted in a previous South African intervention was a culture in which children did *not* spontaneously engage in emergent modelling even when provided with problems set in familiar contexts and languages (Takane, 2021), with the writing on learner dispositions from Mellony Graven's team at Rhodes (e.g. Graven, 2014) noting that sense-making and the need for agency and independence were largely absent from children's responses to what was involved in being good at

mathematics. The lack of fit between RME's learner-driven working and the passive, teacher-led culture on the ground led Tshesane's (2021) study towards more direct attention to introducing and discussing additive relations models for word problem-solving. Guidance on this intentional teaching approach was seen in Polotskaia's (2017) Relational Paradigm.

The problem-solving cycle in Polotskaia's Relational Paradigm can be compressed into three phases:

- *interpretation* of the problem situation, which involves making sense of the problem situation and creating a model to express this sense, and deriving the number sentence in standard form with the appropriate arithmetical operation from the model created;
- *calculation* of the unknown quantity, which proceeds by transforming the number sentence into a calculation method aimed at producing the numerical solution; and
- *evaluation* of the numerical solution, which involves making sense of the calculation results in relation to the initial sense made of the problem situation.

In empirical studies, Polotskaia and her colleagues (Polotskaia et al., 2015, 2016; Polotskaia & Savard, 2018) place emphasis on supporting learners with *interpretation* through the introduction and discussion of models that can capture structural relations. Their chosen model for making visible the relationship between quantities in additive situations involved part–part–whole line segments or bar diagrams. These studies did not deal with the calculation phase noting that for their Canadian learners, difficulties were predominantly related to setting up the appropriate model in word-problem situations rather than to calculating the answer once an appropriate model was in place.

In the interpretation phase, the structural relation in the additive relationship is the primary focus of the solver with attention to discerning the parts and the whole that constitute the additive relationship; hence, structural reasoning is 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, and using awareness of number relations to decide on an efficient calculation strategy. Thus, numerical reasoning is 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. To illustrate this, we use one of the items in the assessment around the focal intervention in Table 1.

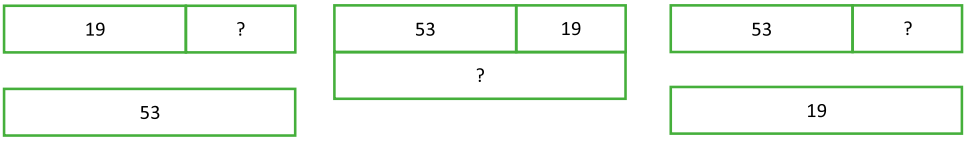
From this vantage point, success in solving additive relations word problems was viewed as a function of the extent to which the solver could interpret the additive relationship involved, express the

Table 1. One test item to exemplify the process of solving additive word problems promoted in the intervention study

Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles?

Interpretation Phase

Involves deciding the part–part–whole relation between the quantities in the problem. Is it:



And formulating the 'standard form' number sentence to calculate the answer—in this case:

53 – 19 =?

Calculation Phase

Involves calculating the answer to 53 – 19 efficiently through use of base 10 structure affordances. One option for this would be to work the problem mentally, or with the support of an empty number line as: 53 – 20 = 33 and then 33 + 1 = 34.

additive relationship in a number sentence written in standard form, and calculate the sum or difference implied by the additive relationship. In the teaching of the intervention (led by the first author) and in the subsequent analyses, RME's attention to using number line models for increasingly efficient calculation was incorporated. For the initial quantitative summary though, the aim was simply to consider the appropriateness of set up of calculation, and carrying out the calculation to produce the answer. The gains seen through this initial quantitative analysis led us into a subsequent qualitative analysis of the models that learners were using to support this set-up, and for efficient calculation. Additionally, while the evaluation of the numerical solution is an important phase of problem-solving in the Relational Paradigm, in this initial step of getting an overview sense of any changes in ways of working, we truncated the evaluation phase to simply looking for the correctness of the final answer that learners produced for each test item. This led to a coding based on the three phases of *interpretation–calculation–answer,* producing a three-digit code for each response.

3. The Additive Reasoning Intervention

The first author worked with a Grade 3 class in one of the 10 schools in the broader Wits Maths Connect-Primary project. The school was in a suburban area serving a historically disadvantaged population, with English as the language of learning and teaching. The focal class had 40 learners, with matched initial, post and delayed-post test data for 39 of these learners. Ethical clearances following informed consent from all participants were granted from the University and the provincial Department of Education. The pre-test was administered in August 2015 prior to the start of the 10 intervention lesson sequence with each lesson about an hour long and two lessons taught per week on average. The post-test was set in November 2015 immediately after the end of the intervention lesson sequence, and a delayed post-test was set in January of the following year, with learners then in Grade 4. Administering the delayed post-test after the long summer break was intended to understand the possibilities for longer-term impact, given the international evidence of learning losses being common after the long summer holidays (Cooper et al., 2006). 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. This timing was intentionally long given the broader evidence of prevalent use of tally counting in South Africa (Schollar, 2008) if that was what learners opted to use. Similarly, while bar diagram, number sentence and empty number line models were shared and discussed in the lessons, the coding took in whichever model/s the children chose to use in their responses.

The test that was used across the pre-, post- and delayed post-test sittings consisted of 12 additive word problem items drawn from Askew's (2004) work, selected to encompass the semantic categories of additive word problems described in the literature (change, combine and compare situations) and including variation in the position of the unknown (result unknown/part unknown/start unknown). The items were piloted with a different class to ensure their validity for use in the study. The first author's broader doctoral study analysis (Tshesane, 2021) includes attention to the difficulty hierarchies described in the literature: strong evidence that compare problems are harder for children to solve than problems from the other two categories, and change situations usually the easiest—see Carpenter et al. (1999); and that 'result unknown' problems have higher facility levels than 'part unknown' or 'start unknown' problems. In the doctoral study, these issues were explored in the qualitative analysis that followed the initial quantitative summary using the coding framework, and thus, we do not expand on this literature here. The items used in the tests with their semantic categories and positions of the unknown are listed in Table 2 for transparency.

Each intervention lesson was split into two main sections. In the initial section, the first author focused explicitly on introducing and discussing combinations of 10 and 20 in bar diagram and empty number line formats. We entitled this initial section as 'connecting models', and it was used—in out-of-context 'bald' number engagements—to demonstrate, discuss and practice translating between a bar diagram, a number sentence in standard form and an empty number line. These skills were then applied in the context of whole class and individual working on selected word problems in the second section of each lesson, with emphasis on interpreting each contextualised

Table 2. Items used across the three test sittings

Question no.	Word problem	Category/unknown
1	On Thursday Lerato picked 17 strawberries. On Friday Lerato picked 23 strawberries. How many strawberries did Lerato pick altogether?	Change increase Result unknown
2	Thando has 34 books on the shelf, 19 books on the floor. How many books does Thando have altogether?	Combine whole unknown
3	52 turtles swam alongside 49 dolphins. How many turtles and dolphins were swimming altogether?	Combine whole unknown
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
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
6	69 children travel by bus. 77 children travel by train. How many more children travel by bus?	Compare Difference unknown
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
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
9	Chris collected 53 stickers and Charles collected 19 stickers. How many more stickers did Chris collect than Charles?	Compare Difference unknown
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
11	Mike picked some peaches. He ate 7 peaches and gave 19 to his friends. How many peaches did Mike pick?	Combine Whole unknown
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

word problem into a bar diagram, formulating the number sentence in standard form, and then calculating the answer, i.e. using the model sequence engaged with in the first part of the lesson.

Given the prior research noting that (a) word problem tasks have been useful in supporting sense-making, but also that (b) expecting learners to engage in the sense-making required for interpretation naturally and independently was unlikely in the South African context, this lesson model, with its explicit attention to introducing and discussing models that would support interpretation and calculation-phase working, was used across the 10-lesson sequence.

4. Methodology: Coding Learner Responses Using the Three-digit Coding Scheme

Given our interest in learners' success with, and co-ordination of, structural and numerical reasoning to produce answers**,** we analysed learners' solutions with the truncated interpretation–calculation–answer cycle as the basis. Profiling learner responses according to this cycle provided a macro-level vantage point on the success of the intervention approach based on explicit teaching of models and their use in mediating the solution process.

An important phenomenon identified in TIMSS analyses and noted as more marked in South Africa than in other countries (Bowie et al., 2022) is that of learners writing guessed answers rather than leaving answer spaces blank regardless of whether they know how to produce the answer. In order to enable a holistic analysis of learners' engagement with each item, every learner response across the three test administrations was studied and coded for the Interpretation, Calculation and Evaluation Phases in the following way: for the interpretation and calculation phases, three coding options were possible: *, no evidence; 0, incorrect; and 1, correct. For the Evaluation phase, while three options were theoretically possible (no answer, incorrect answer, correct answer), the phenomenon of

Table 3. Key to coding for appropriateness of interpretation, calculation and evaluation

Digit	Interpretation	Calculation	Evaluation
*	No evidence of a model reflecting an interpretation of the problem situation	No evidence of a calculation	No answer in evidence
0	Incorrect interpretation of the problem situation seen in inappropriately configured model/s	Evidence of an inappropriate calculation	Incorrect answer
1	Correct interpretation of the problem situation seen in appropriately configured model/s	Evidence of an appropriate calculation	Correct answer

guessed answers flagged above likely contributed to our needing only two coding options to deal with the empirical base: 0, incorrect; and 1, correct. The coding scheme is detailed in Table 3.

This three-digit coding system allowed us to summarise the processes of working across the interpretation and calculation phases, and the extent of success/failure related to these two phases for every single item across the three test administrations.

Each three-digit code combination describes a particular form of engagement with an additive relations problem, with attention on whether 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, an incorrect calculation attempt and an incorrect answer to an item was allocated a '100' code; and an incorrect answer with no working shown was allocated a '**0' code. A 1*1 coding reflected a response with an appropriate model, no evidence of a calculation but with a correct answer.

While theoretically, a 'three-digit' code with three options for each coding segment can produce 27 coding options, the empirical data contained 12 different ways of working across the 468 responses across the three test sittings. Of these 12 different ways of working, 2 were 'answer-only' responses (**1 where the correct answer was provided and **0 for an incorrect answer), so they could not be analysed for interpretation and calculation. All 12 empirical three-digit codes are detailed in Table 4.

The other 10 ways of working demonstrated by the learners related to modelled responses. Quantifying the prevalence of each of the three-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. The three-digit codes that emerged from coding learner responses across all three test sittings could be 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). Moreover,

Table 4. Codes emerging from learners' responses across all three test sittings

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	Three-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

'modelled' responses included a calculation in some instances and did not in others, and, in both cases, either yielded correct or incorrect answers. Modelled responses were associated with correct answers in half of the 10 modelled responses codes—111, 001, 101, 1*1, and **1 in Table 4. Modelled responses with incorrect answers are represented in the other five modelled responses codes, indicated in black text in Table 4.

Differences in the patterns of occurrence of these codes across the three test sittings provided us with an initial sense of the outcomes of the intervention, in ways that went beyond looking only at the extent of correct answers. The coding additionally provided an overview of the extent of appropriate interpretation of the word problem into a model, and of appropriate calculation. The three-digit coding also allowed for consideration of how the interpretation and calculation processes are connected, which was useful given Giroux and Ste-Marie (2001) emphasising the need, not just for models to be used, but also to be appropriately coordinated. To this end, each modelled response was analysed for coordination of structural reasoning in the interpretation phase and numerical reasoning in the calculation phase. In general, a response including a model reflecting an inappropriate interpretation (0 in the first digit) and/or an inappropriate calculation (0 in the second digit) was deemed to reflect inappropriate coordination of structural and numerical reasoning. Examples of inappropriate coordination codes are the '100' code (a correct interpretation of the problem situation into a model, followed by an incorrect calculation attempt), the '101' code (a correct interpretation of the problem situation into a model, followed by an incorrect calculation and the correct answer) and the '000'/'001' codes (inappropriate interpretations of the problem situation). Thus, only learner responses showing a correct interpretation of the problem situation into a model, a correct calculation and either an incorrect (110 code) or a correct answer (111 code) were regarded as responses with appropriate coordination of the two aspects of reasoning. While this approach can be seen as setting a high bar for coordination, we chose this approach in the face of strong evidence that competence with setting up and using models of situations is—in and of itself—a hallmark of mathematical working, as seen for example in Freudenthal's (1973, 1986) writing on mathematising. This holds regardless of whether a final correct answer is eventually produced. In the absence of evidence of appropriate interpretation and/or calculations, an inference of appropriately coordinated mental structural and numerical reasoning was applied to responses reflecting the 1*1 and the **1 codes as well. These analytical tools provided lenses for characterising shifts across the range of additive relations problem classes and types that learners encountered in the pre-, post and delayed post-test.

This approach produced 'three-digit' codes for every single learner response. It facilitated a developmental analysis of the extent to which learners were able to transform the additive relationship in a word problem into an appropriate model, an appropriate calculation, and a correct answer—with each of these parts made visible separately, but analysable in conjunction too through their coordination. By extension, this analysis provided the initial bases upon which this intervention could be adjudged to have been successful.

5. Outcomes Using the Coding Framework

The significance of disaggregating modelled-responses-with-calculations from modelled-responses-without-calculations can be seen in the distribution in Table 5. Answer-only responses predominated in the pre-test (351/468 responses, or 75% of all responses). Amidst the sharp decline of answer-only responses in the post- and delayed-post tests, 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.

Disaggregating modelled-responses-with-calculations from modelled-responses-without-calculations indicated that increases in the extent of use of models and calculations were outstripped by shifts in the extent of success associated with these responses: while 353 (75%) of the 468 responses

Table 5. The distribution of modelled and answer-only responses across the 468 responses in each of the three tests

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

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 success rate of modelled-responses-with-calculations from pre-test to post-test to delayed post-test (40–42–55%), resting on large increases in the base numbers, revealed growing learner willingness, not only to interpret problem situations into models, but also to perform written calculations.

This finding led into our analysis into the extent of coordination of problem interpretation and calculation within responses. With each response coded using the three-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. It also set the scene for analyses of shifts in learner competence in terms of coordination between the interpretation phase and the calculation phase. In the three-digit coding analysis that follows, we attend to what this coding methodology made visible in terms of:

- differences in learner responses across the pre-, post- and delayed post-tests, and through this, any shifts in overall outcomes; and
- coordination of models across the Interpretation and Calculation Phases of problem-solving.

5.1. *Three-digit Code Analysis of Shifts in Interpretation, Calculation and Evaluation*

We attend, in the next section, to the differential contributions made to the improvement in performance, with particular focus on codes with large numbers of responses across the three test sittings.

5.2. *Distribution of Three-digit Codes Across the Three Tests*

Given the increases in proportions of correct answers associated with modelled-responses-with-calculations from the pre-test through the post-test to the delayed post-test, we drilled deeper into this category by using the three-digit codes. In Table 6, we summarise the distribution of three-digit codes in each of the three tests, and the differential contributions made by each code. The important shifts to highlight in Table 6 relate to big increases in 111 codes, smaller increases in 1*1 codes, and a big drop in **0 and **1 codes underpinned by learner responses moving into the modelled response category.

Table 6 also shows increasing prevalence of 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: 100 and 1*0 codes. The 100 code went from 22 to 108 to 102 incidences and the 1*0 code went from 8 to 18 to 24 incidences across pre-, post- and delayed post-tests. Whilst these responses were associated with incorrect answers, the five-and-a-half-fold increase in the prevalence of each of these codes comprised learners moving from the substantial prevalence of the **0 in the pre-test and indicated improvements in interpretation of the different problem situations.

Secondly, there were also increases in the prevalence of 000 (modelled-responses-with-calculations) and 0*0 (modelled-responses-without-calculations) codes between the pre- and delayed post-test though these were not to the same degree as the increases in the 100 and 1*0 codes.

Table 6. The distribution of three-digit codes across the 468 responses on each of the three tests

Interpretation phase analysis	Calculation phase analysis	Evaluation phase analysis	Three-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

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, the increases in the 111, 101, 100, 1*0 and 1*1 codes reflected improvements in appropriate interpretation of the problem situation into a model. In the qualitative analysis, we saw that the increases were mostly underpinned by growing use of bar models in the latter tests. With the majority of correct answers associated with the 111 code (188 of the 229 correct answers in the delayed post-test), the three-digit coding also revealed improved calculation of solutions, with qualitative analysis indicating greater use of the number line in the latter tests. Both of these qualitative indicators, though not a part of this paper's analysis, suggested that the gains could be linked to the intervention where the bar and number line models had been presented and discussed.

The large increase in 111 codes pointed to improvements in performance linked to an increase in appropriate coordination of learners' structural and numerical reasoning. The extent of coordinating structural and numerical reasoning was the focus of the second overview analysis, conducted using the three-digit 111 and 110 codes. As noted already, the occurrence of the 110 code indicated that despite appropriate interpretations of situations and appropriate calculation approaches, these coordinations of structural and numerical reasoning could still involve errors. However, the numbers in this latter category were small across all three test administrations (2, 1 and 5 responses respectively).

We noted earlier that we interpreted responses reflecting the 1*1 and the **1 codes as appropriately coordinated mental structural and numerical reasoning. There were increases in 1*1 occurrences (4 to 2 to 22 across the three tests), suggesting small improvements in mental working in the group. The large decrease in the **1 category (89 to 15 to 11 across the three tests) was underpinned predominantly by greater willingness to work with models.

The increases in responses reflecting the 100 (from 22 to 108 to 102) and 000 (35 to 105 to 51) codes in the post-tests were also underpinned largely by moves from the initial **0 category, and thus, a move to attempting to interpret the situation in some way, though these attempts were not always appropriate. Thus, despite increases in appropriate coordination of structural and numerical reasoning seen in the predominance of 111 codes in the delayed post-test, relatively widespread evidence of inappropriate coordination of structural and numerical reasoning—which curtailed improvements in performance—remained.

6. Discussion

As seen in the large increases in the proportion of responses reflecting the 111 and 1*1 codes, and a large residual of **1 codes, the three-digit coding methodology used in this study made clear the fact

that, overall, the improvement in performance is attributable to increased 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 learners producing 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, these moves over time were from learners who initially 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 'compare' 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, sometimes with appropriate mathematical expressions in standard form. This compared with 71/468 (15%) of the responses with these codes in the pre-test. This move over time—tracked with the distribution of the three-digit codes—indicates major improvements in learners' ability to correctly interpret problem situations.

7. Conclusions

The coding methodology using three-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. The overview of these profile of engagement revealed how the learners performed in each phase of the problem-solving cycle, and also made visible learners' coordination of the interpretation and calculation phases. The outcomes suggest that it is worth focusing classroom attention on the coordination of structural and the numerical reasoning (Giroux & Ste-Marie, 2001) for successful problem-solving, rather than dealing with interpretation and calculation in isolation of each other. They also point to the usefulness of expanding beyond the single-focus of analyses of word problems seen in many prior studies.

In the earlier doctoral study, the large increases in modelled responses seen through the three-digit coding led to further qualitative analysis of learners' choices for working with models in the interpretation and calculation phases. This points to some of the limitations of the coding scheme: it was a useful first-level analysis for seeing patterns across close to 1500 item responses, but it did not delve into the nature of models seen in learners' responses, nor to the extent of efficiency of their calculation strategies. Nevertheless, using the three-digit coding approach pointed to ways to zoom in on the stories of interest in the qualitative analysis which followed, and this is likely where the power of the approach lies. Crucially, the finding of the importance of coordination feeds back into our understandings of what needs emphasis in classroom working, and thus, the approach appears to have developmental, as well as research salience.

Acknowledgement

The financial assistance of the National Institute for the Humanities and Social Sciences, in collaboration with the South African Humanities Deans Association, as well as that of the Wits Maths Connect Primary Project towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Herman Tshesane http://orcid.org/0000-0002-4492-7241

References

Askew, M. (2004). *BEAM's big book of word problems Year 3 and 4 Set*. Nelson Thornes.

Besharati, N.A., Fleisch, B., & Tsotsotso, K. (2021). Interventions to improve learner achievement in South Africa: A systematic meta-analysis. In F. Maringe (Ed.) *Systematic Reviews of Research in Basic Education in South Africa. African Sun Media:* 27–67.

Björn, P.M., Äikäs, A., Hakkarainen, A., Kyttälä, M., & Fuchs, L.S. (2019). Accelerating mathematics word problem-solving performance and efficacy with think-aloud strategies. *South African Journal of Childhood Education*, *9*(1), a716. https://doi.org/10.4102/sajce.v9i1.716

Bowie, L., Venkat, H., Hannan, S., with Namome, C. (2022). *TIMSS 2019 South African Item Diagnostic Report: Grade 5 Mathematics*. Human Sciences Research Council.

Carpenter, T.P., Fennema, E., Franke, M.L., Levi, L., & Empson, S.B. (1999). Children's mathematics. *Cognitively Guided Instruction.* Heinemann. [Transl. from C. De Castro and M. Linares, *Las Matemáticas que hacen los niños*.]

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of Educational Research*, *66*(3), 227–268.

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.

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.

Freudenthal, H. (1973). *Mathematik als pädagogische Aufgabe* (vol. 2). Klett.

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

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.

Graven, M., & Heyd-Metzuyanim, E. (2014). Exploring the limitations and possibilities of researching mathematical dispositions of learners with low literacy levels. *Scientia in educatione*, *5*(1), 20–35.

Hoadley, U. (2012). What do we know about teaching and learning in South African primary schools? *Education as Change*, *16*(2), 187–202.

Mudaly, V., & Narriadoo, D. (2023). Solving word problems by visualising. *African Journal of Research in Mathematics, Science and Technology Education*, *27*(1), 47–59. doi: 10.1080/18117295.2023.2183612

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.

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

Sepeng, P. (2014). Use of common-sense knowledge, language and reality in mathematical word problem solving. *African Journal of Research in Mathematics, Science and Technology Education*, *18*(1), 14–24.

Takane, T. (2021). *Exploring mathematizing processes of South African Grade 2 learners for solving additive relation problems*. Unpublished doctoral dissertation. University of the Witwatersrand, Johannesburg.

Tshesane, H. (2021). *Developing Additive Relations Word Problem-solving with the use of Relational Models*. Unpublished doctoral dissertation. University of the Witwatersrand, Johannesburg.

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.