EXPLORING TRENDS IN PRE-SERVICE TEACHERS' ADDITIVE REASONING: A QUASI-LONGITUDINAL ANALYSIS ACROSS THREE COHORTS

S. Hansa*
https://orcid.org/0000-0002-2274-8883

L. Abdulhamid*
https://orcid.org/0000-0002-7822-9177

C. Mathews*
https://orcid.org/0000-0001-6666-0501

H. Tshesane*
https://orcid.org/0000-0002-4492-7241

A. A. Essien*
https://orcid.org/0000-0002-0040-8773

* Mathematics Education Division,
Wits School of Education
University of the Witwatersrand
Johannesburg, South Africa

ABSTRACT

The study reported in this article examines the trends in pre-service teachers' (PSTs) additive reasoning performance through a quasi-longitudinal analysis of three-year cohorts involving 498 PSTs. The findings indicate consistent improvement in additive reasoning performance across the cohorts. However, while PSTs excelled in simple missing addend tasks, they often struggled with tasks involving a deeper understanding of equivalence (such as equivalence tasks with two unknowns), as well as tasks requiring reasoning about structuring the numbers involved for efficient calculation and compensating for adding more or less than was intended. While these findings underscore the potential of targeted teaching interventions to shift PSTs performance across a variety of additive tasks, they also point to the need for more emphasis to be placed on developing PSTs reasoning about number structure and equivalence beyond simple missing addend problems. More generally, the findings highlight the need for continual assessment and tailored design instruction to address specific challenges in PSTs' mathematical competencies in additive reasoning.

Keywords: Additive reasoning; Mathematical equivalence; design research; mathematics performance; equality; teaching intervention; quasi-longitudinal

INTRODUCTION

Within the South African context, wide-spread concerns about performance in mathematics across different education levels have been well-documented (Fleisch 2008; Spaull and Kotze 2015; Reddy et al. 2016, 2022). Studies in primary education have noted that poor performance is exacerbated by the fact that many South African teachers do not have sufficient content knowledge to be able to teach mathematics effectively (Graven 2016; Taylor and Taylor, 2013; Bowie, Venkat and Askew 2019). Findings on mathematics teacher test from the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) 2007, which tested Grade 6 mathematics teachers from a nationally representative sample of primary schools, showed that 79 percent of Grade 6 mathematics teachers displayed content knowledge levels below the grade level at which they were teaching. Moreover, this problem has also come under the spotlight at a pre-service level, where teachers tend to come into their teaching degrees with weak mathematical knowledge from their schooling (Bowie et al. 2019).

In response to the teachers' knowledge gap, higher education institutions have been working on growing pre-service teachers' (PSTs') mathematics content knowledge by changing how pre-service teachers are being prepared for the content knowledge demands of teaching mathematics (See for example, Essien & Adler, 2016). The PrimTed project, for example, has been working to establish "a shared understanding of appropriate curricula and standards across mathematical knowledge and classroom practice for pre-service primary teachers" (Bowie et al. 2019, 287). One of the areas of focus and key to mathematical competency is additive reasoning, of which number structure and equivalence relations are central components. The ability for learners to know the additive structure (Vergnaud 2020) underpinning the additive relationship (Davydov 1998) that unites addition and subtraction is an essential starting point towards making sense of both elementary and higher mathematics. In particular, understanding the number structure that underpins all additive relations enables reasoning about the equal sign as a symbol for producing the answer or as a symbol denoting quantitative sameness. The latter understanding of the equal sign is critical for learners solving additive tasks (See Essien and Setati, 2006; Vermeulen and Meyer, 2017).

Furthermore, in response to findings from research regarding the need to strengthen the mathematics content knowledge of primary school teachers in South Africa (see, for example, Bowie et al. 2019; Venkat and Spaull 2015; Venkat and Adler 2012), in 2022, the Wits Maths Connect-Primary (WMC-P) project initiated a design research and development project on enhancing content knowledge in primary teacher education. As already mentioned in Essien *et al*. (2023), the aims of this project are 1) to gain an understanding of the gaps in the content knowledge of pre-service teachers (PSTs) on different foundational mathematics concepts (such as additive reasoning, multiplicative reasoning and geometric reasoning; and 2) based on this understanding, to propose design principles for the effective teaching of these foundational concepts in a way that attends to these content gaps. Drawing from this more extensive research and development project, the present article examines PSTs additive reasoning performance trends through a quasi-longitudinal analysis from 2022 to 2024. Therefore, this article aims to investigate shifts in PSTs' performance across cohorts in a design research approach aimed at developing design principles for the effective teaching of additive reasoning. In doing so, we analysed the performance of PSTs across the three years, from pre-test to post-test, in terms of their development of additive reasoning. We also identify items that PSTs performed consistently well and those that they performed consistently low across the three-year period. The aim is to further inform changes in (teaching) practice and emphasis in the subsequent research design interventions.
The research questions guiding this article are:

- What are the discernible trends or patterns (if any) in PSTs performance trajectories in additive reasoning across a three-year period (2022-2024)?
- What type of additive reasoning assessment items did PSTs perform consistently high or low on across the three-year period?

In answering the above questions, we also seek to inform a key question in our quest for design principles on how best to teach additive reasoning for understanding, viz "what does PSTs' performance tell us about areas of focus in further developing PSTs' understanding of additive reasoning?"

In the following sections, we begin with a literature review on the teaching and learning of additive relations, the central focus of this article. This is followed by a thorough discussion of the study's methodology. The findings are then presented and analysed, and finally, the study's implications are detailed.

ADDITIVE AND EQUIVALENCE RELATIONS IN TEACHER EDUCATION

The concept of the equal sign as a symbol of mathematical equivalence has been extensively studied (e.g. Simsek et al., 2021; Vermeulen and Meyer, 2017). Whilst much of these studies were conducted in the USA, in the last two decades, there has been a trickle of studies conducted in several countries such as England, China, New Zealand, Turkey, and South Africa (Capraro et al. 2010; Hunter 2007; Jones et al. 2012; Machaba 2017; Vermeulen and Meyer 2017, Essien 2009). Across studies on equivalent relations, researchers are in consensus that there are fundamentally two conceptions of the equal sign: the "relational conception*",* which embodies the two aspects of sameness and the interchangeability of two quantities and/or expressions (McNeil et al. 2019; Rittle-Johnson et al. 2011) and the "operational conception*"* of the equal sign which views the equal sign as a prompt to "do something" (Kieran 1981; Knuth et al. 2006). One critical hypothesis put forward by Donovan et al. (2022), which is of interest to this study, is that deliberately teaching students to understand the equal sign in a relational way as a mathematical object promises greater cognitive gains than if this understanding were to develop naturally.

For example:

"Fill in the number to make the number sentence true: 623 − 298 = 622 −___"

Figure 1: PSTs performance on the whole number and operations tasks (culled from Bowie et al. 2019, 291)

Cohorts	Mean	Std dev
University A		
1st Years	66	20
4th Years	68	18
University B		
1st Years	59	18
4th Years	56	19
University C		
1st Years	49	16
4th Years	52	19

Considering the importance given to number and number operations in the Foundation and Intermediate phases, this overall average performance of PSTs in such a fundamental domain of mathematics is reason for concern. Even though Bowie et al. (2019) do not discuss common answers to the additive reasoning task *(*623 − 298 = 622 −__), the overall performance by both first and fourth-year students suggests a prevalent difficulty with such problems.

Based on Bowie et al's. (2019) recommendations, Essien et al. (2023) reported on a study that identified knowledge gaps in first-year pre-service teachers (PSTs) on number structure and mathematical equivalence – two fundamental concepts in the context of whole numbers and additive operations. Their findings support the need to intervene in PST's mathematical content knowledge, corroborating the conclusions and recommendations made by Bowie et al. (2019).

METHODOLOGY

With the overarching aim of developing design principles to guide and improve how future PSTs understand and teach concepts of mathematical equivalence and number structure in additive relations, this article looks across three cohorts that have been involved in this intervention study over the past three years.

Research Design

This study employed a design research approach – an approach aimed at developing innovative solutions in particular problem situations and constructing guidelines for future development efforts (Reeves, McKenney and Herrington 2011). This included the four-stage framework of design research consisting of an iterative process of 1) analysing the problem, 2) developing a solution informed by existing design, 3) iterative cycles of testing and refining the solution, and 4) reflections to produce design principles (Reeves 2006). Stage 3 lasted for five weeks (for each cohort) as an intervention to develop PSTs' understanding of mathematical equivalence and number structure. The first week started with a pre-test to assess PSTs' initial understanding. The next three weeks involved teaching sessions based on 1st year primary mathematics concepts, specifically focusing on additive relations and number concepts. These lessons were tailored to address the PSTs weak points identified in the pre-test. Collaboratively, strategies were devised to enhance the quality of teaching and learning related to additive relations, using well thought-out examples to improve teaching and learning quality. Two fundamental principles guided the development of these lessons. Firstly, concepts such as place value, part-part-whole relations, balance, and the number line formed the basis of instructional design. Secondly, Skemp's (1976) seminal work on the conceptualisation of instrumental and relational understanding alongside Carpenter et al.'s (2005) and Matthews et al.'s (2012) extensions of these constructs served as a theoretical framework, informing the intervention, the selection, and the development of examples to be used in the classroom (See Essien et al. 2023 for more details). Subsequently, a post-test was conducted in the fifth week.

In 2022, the design of the course included an additive reasoning pre-test followed by four hours of lectures and two-hour tutorials over two weeks, ending with a post-test. In 2023, the same intervention model was replicated. In terms of the teaching, lessons drawn from 2022 were used to inform the 2023 teaching in line with design research. Before 2022, we focused on strategies to efficiently solve addition and subtraction problems in teaching additive reasoning. However, in 2022, while we continued exploring practical ways of solving these problems, we also strongly emphasised the role of the equal sign. We encouraged using the equal sign to produce answers in some cases and to balance both sides of the equation. In the final year, 2024, the model was replicated again, with the only difference being a new lecturer on the course inducted into the intervention model run the previous two years.

As indicated in the introductory section, this study investigates three distinct cohorts of first year PSTs to see if changes in the design of the PSTs course led to (or not) consistent improvement in PSTs' performance in understanding additive relations. Given this constraint of not being able to track the same students over time, we compared different cohorts each year, following a quasi-longitudinal approach (Campbell and Stanley 1966). Kifer (1975) asserts that where school populations remains stable and students' rank order of achievement remain constant, the results of a quasi-longitudinal study of achievement should mirror those of a proper longitudinal study. In our study, this entailed selecting three separate and successive cohorts of PSTs. This approach enabled the analysis of trends and patterns in their performance within each cohort over the three-year period.

Data was collected using a 27-item pre-test and post-test assessments carefully designed on additive reasoning. The tests contained various item types, including questions on completing missing numbers in equivalent additive relations, such as "37 + 56 + 13 = __ + 50". Additionally, questions were included to assess PSTs capacity to comprehend additive word problems and grasp place value concepts. There were also questions about part-part-whole understanding, number structure, and logical mathematical reasoning in additive relations.

A total of 498 first-year PSTs participated in the study, with 174 PSTs in 2022, 171 in 2023, and 153 in 2024. These PSTs were enrolled in a Primary Mathematics Content course at a University in South Africa and all participants consented to participate in the study and had completed both the pre-test and the post-test.

Method of Analysis

Descriptive statistics, including means and percentages, were employed to summarise PST's performance on additive reasoning in both pre- and post-tests across cohorts. Trends and patterns in PSTs' performance trajectories over the three-year period were examined using a confidence interval plot and a line graph. Subsequently, a one-way analysis of variance was conducted to assess significant differences within and across the cohorts' pre-and post-test scores. Additionally, PSTs' performance on specific types of additive reasoning problems across cohorts was compared to identify items where performance was consistently high and consistently low. In doing so, we arranged PSTs performance in each assessment based on the item's facility in quartiles. Items in the upper quartiles indicated higher performance, while items in the lower quartiles indicated lower performance. We then looked at common items in the lower and upper quartiles in the pre-and post-tests for the three-year cohorts. We found 4-items each in the lower and upper quartiles. Characteristics of these items in the two groups are closely examined and discussed. We see this type of systematic cohort quartile item analysis as crucial for understanding what needs more or less emphasis, more specifically, in the subsequent design research intervention and more generally, in the teaching and learning of additive reasoning in primary teacher education.

While quasi-longitudinal studies are useful for observing trends over time, they have limitations that affect their ability to establish causality, control for confounding factors, and ensure the validity and generalisability of their findings. However, in our study, the stability of the pre-service teacher population over the three years is consistent with Kifer's (1975) observation that the results of a quasi-longitudinal study of achievement closely mirror those of a true longitudinal study, enhances the validity, reliability, and overall credibility of our findings.

Ethical Considerations

Permission was first sought from the University's ethics committee and then from the Faculty of Humanities' registrar, and subsequently from the Head of the School of education given that the study was on students. Following formal permission from the University, the students were asked for consent to be part of the study.

ANALYSIS AND FINDINGS

Trends in PSTs' Performance Trajectories

Table 1 presents the mean and standard deviation of the pre-and post-tests across the cohorts on the 27 test items. Across all three cohorts, mean scores consistently increased from the pre-test to the post-test, indicating improvement in additive reasoning performance over time.

Table 1: Mean and Standard deviation of pre-test and post-test performance across the three-year cohorts

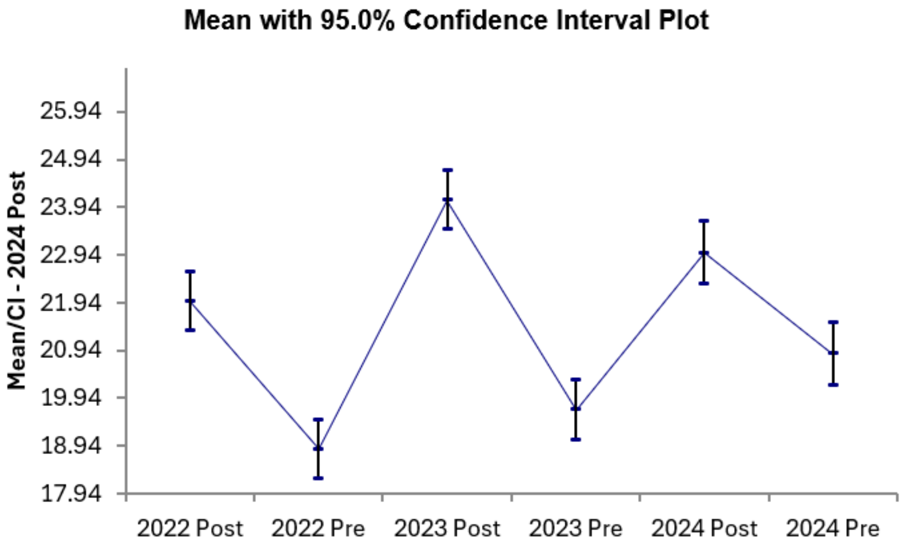
Cohorts	Pre-test		Post-test		Mean diff
	Mean	Std dev	Mean	Std dev	
2022 (n = 174)	18.87	4.61	21.97	4.39	3.10 (14%)
2023 (n = 171)	19.69	4.92	24.08	2.51	4.39 (18%)
2024 (n = 153)	20.87	4.39	22.97	3.33	2.10 (9%)

Table 1 indicates the difference in gains varied across the cohorts, with the 2023 cohort showing the highest mean difference (4.39, an 18% increase) and the 2024 cohort showing the least mean difference (2.10, a 9% increase) from pre-test to post-test. Cohort 2024 generally showed higher mean scores in the pre-tests compared to Cohort 2022 and 2023 (see Figure 1). Students in the 2024 cohort came into the B.Ed programme with a higher Admission Point Score (APS) of a minimum of 37 compared to the APS requirement of 36 in 2022 and 2023.

The pre-test scores' standard deviation was higher than the post-test scores across all three cohorts. A higher standard deviation in the pre-test suggests a broader spread of the PSTs performance prior to the intervention, indicating more varied abilities. Conversely, the lower standard deviation in the post-test implies that PSTs performance was more consistent after the intervention, reflecting a reduction in variability and suggesting that the intervention may have helped to standardise their additive reasoning skills.

Figure 2 depicts a confidence interval plot showing the trend of mean performance from pre-test to post-test across the three cohorts. The figure shows a more visualised positive trend of performance across the three cohorts and a higher gain from the 2023 cohorts. There is also a narrower confidence interval in the post-test compared to the pre-test, again suggesting a reduced variability and more consistent performance in the post-test. Figure 2 also shows a consistent increase in the pre-test from the 2022 to 2024 cohorts, but still the 2023 cohort exhibits greater learning gains. The 2024 cohort starts at higher performance but still shows learning gains.

Figure 2: Trends in performance in pre and post-test across the three-year cohort



As indicated elsewhere (Essien et al. 2023), while it is not the intention of the larger study (from which this article is drawn) to ascribe gains made in the post-tests to the impact or the effectiveness of the intervention, what Figure 2 shows is that despite some variations in the improvement between cohorts, the PSTs have been able to make learning gains in their understanding of additive reasoning problems over time. Nevertheless, further analysis (such as inferential tests) was necessary to confirm the significance of these differences. Hence, a one-way ANOVA test was conducted for PSTs performances across the three-year cohort, as shown in Table 2.

Table 2: One-way ANOVA test results for PSTs performances across the three-year cohort

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Pre-Test Scores	Between Groups	607.098	2	303.549	12.488	.000
	Within Groups	12178.328	495	24.308		
	Total	12785.427	497			
Post-Test Scores	Between Groups	270.788	2	135.394	11.977	.000
	Within Groups	5663.448	495	11.304		
	Total	5934.236	497			

The ANOVA results suggest that there are significant differences between the pre-and post-test assessments across the three-year cohorts ($F(2,501)=12.488$, $p=0.000$ and $F(2,501)=11.977$, $p=0.000$). The significant p-values (< 0.05) for pre-test and post-test scores imply a significant learning gain on participants' scores. In other words, the variation in scores observed in pre-and post-test across the different cohorts is unlikely due to chance, suggesting consistent learning gains in PSTs' performance. However, Table 2 does not tell us what additive reasoning tasks PSTs performed better at from pre-test to post-test, which items are consistently poorly attempted, and which we now turn to.

Analysis of Items with High or Low Levels of Performance

High level of performance

As indicated earlier, the PSTs' performance in the 27 items in both pre- and post-test in each cohort was arranged into quartiles. Four items consistently appeared in the fourth quartile across all three cohorts. These items have an average score ranging from 94 per cent to 98 per cent, as shown in Table 3. The scores clearly indicate that the items were relatively easier for the PSTs.

Table 3: Items with consistently high performance

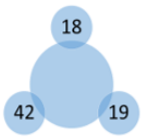
	Item	Pre-Test (%)	Post-Test (%)
a)	56 + __ = 102	94	95
b)	__ + 23 = 80	95	97
c)	17 + __ = 42	95	98
d)	7 + __ = 12	96	98

The four items flagged as consistently high-performing need missing addend calculations with number ranges within the Grade 3 CAPS curriculum (Department of Education 2011). All four calculations are addition calculations, making it easier to find the missing addend. Nevertheless, this finding supports the notion of a growing understanding of the equal sign as signifying equivalence rather than as a symbol inviting students to find the answer (Essien and Setati 2006). Again, there is a positive shift in performance across all four items from pre- to post-test. The results in Table 3 indicate that both our intervention and primary teacher education, in general, require greater emphasis on areas of additive reasoning beyond additive missing addend problems, such as those presented in items a-d above.

Low level of performance

In relation to the items that consistently demonstrated low performance, two items appeared in the first quartile across all three cohorts from pre-test to post-test. As shown in Table 4, these items have an average score as low as 28 per cent.

Table 4: Items with consistently low performance

	Item	Pre-Test (%)	Post-Test (%)
a)	The sum of the 3 numbers in the 3 small circles needs to go in the bigger centre circle.  Select the two numbers that are quickest to add first. Why are these two numbers the quickest to add first?	28	35
b)	I calculated a) 658 + 236 = 894 and b) 658 – 236 = 422. But I was meant to calculate 658 + 237 and 658 – 237. To correct my answers, I must: i. add 1 to both my answers. ii. subtract 1 from both my answers. iii. add 1 to my answer in (a) and subtract 1 from my answer in (b) iv. subtract 1 from my answer in (a) and add 1 to my answer in (b)	55	65

There is a positive shift in performance across the two items from pre-test to post-test. In item (a), the idea was to assess PSTs fluency in understanding number structure. 18 and 42 would easily give a multiple of ten, which could be added to 19. However, many PSTs often select 18 and 19 as smaller numbers, not paying attention to the number structure and fluency in forming a multiple of ten, which is easier to add (Threlfall 2009; Rathgeb-Schnierer and Green 2019). Item (b) (in Table 4) involves simultaneously thinking of both addition and subtraction. In item (b), many PSTs selected either "add one to both my answers" or "subtract one from both my answers", given that 236 was changed to 237 in both calculations. Students were meant to draw on additive reasoning of one more/one less but failed to draw on relational reasoning to find the answers to the calculations.

While these calculations fall within the Grade 3 CAPS curriculum range, equality, number structure and fluency, as well as relational additive reasoning are areas of concern for many PSTs. This suggests that more emphasis is needed in these areas to develop a deeper and conceptual understanding of additive relations and equivalence (Wilmot 2018). The above common answers provided by the PSTs suggest that a need for further development of what Carpenter et al. (2005) refer to as relational thinking which involves a deeper understanding of number structure. This is necessary not only within our intervention programme but also more broadly in primary teacher education, particularly in the context of additive reasoning.

IMPLICATIONS AND CONCLUDING REMARKS

The analysis of trends in PST's performance trajectories in additive reasoning across the three years (2022-2024) reveals several noteworthy findings. Firstly, there is a consistent improvement in mean scores from pre-tests to post-tests across all three cohorts, indicating a positive trend in additive reasoning skills development over time. This finding is particularly significant within the South African educational context given the previous findings of Bowie et al. (2019). The key significance lies in their improved mathematical knowledge, specifically regarding additive reasoning.

Furthermore, cohort 2024 exhibited higher mean pre-test scores compared to cohorts 2022 and 2023. This difference could be linked to the higher APS of the PSTs in this cohort, indicating potentially stronger academic backgrounds upon entry into the program. Despite the higher baseline performance in cohort 2024, there is still an improvement in their post-test scores.

Secondly, analysing items with consistently high or low-performance levels provides valuable insights into specific areas of strength and weakness in PSTs' additive reasoning skills and what needs to be prioritised in the subsequent teaching intervention and primary teacher education more generally. Items such as missing addend calculations consistently demonstrated high performance across all cohorts, suggesting relative ease for PSTs in solving these types of problems. This shift is highly significant within the South African context, as it represents a notable increase in the ability to solve equivalence tasks with one missing addend. According to Vergnaud (2020), this improvement also indicates a greater familiarity with additive structures, facilitating the development of additive relations. We included such missing addend problems because learners in South Africa often need help solving this problem type (Venkat and Roberts 2022; Weitz and Venkat 2017). However, PSTs demonstrate a growing understanding of the equal sign as signifying equivalence rather than as a symbol inviting students to find the answer (Essien and Setati 2006). This insight is valuable as it suggests that future interventions should prioritise other areas of additive relations rather than focusing on missing addend problems. Since PSTs have a good understanding (based on the findings) of missing addend problems, while learners continue to struggle with them, it is essential to enhance PSTs' pedagogical skills to better support learners in this area.

Conversely, items involving deeper ideas of equality, number structure, fluency and relational additive reasoning are areas of concern as PSTs consistently demonstrate relatively poor performance. These difficulties were also seen in Bowie et al. (2019), wherein PSTs in their first and fourth year of study experienced difficulty working structurally with numbers. Venkat and Askew (2018) contend that with an intentional focus on structuring, especially in relation to numbers, students can understand the relationships within the number system they are learning about. This finding is crucial for future intervention to address these areas in PSTs additive reasoning skills, particularly where students perform consistently poorly.

Specifically, the study highlights the necessity for focused interventions aimed at enhancing two key aspects: conceptual understanding of equivalence and procedural fluency within number structure. The identification of areas to pay close attention to in subsequent interventions and in primary teacher education more generally, is the core of the design research, aiming to achieve significant improvements in PSTs' additive reasoning skills.

ACKNOWLEDGEMENTS

This work is based on research supported by the National Research Foundation South African Numeracy Chair at Wits (grant number 74703). Any opinion, finding and conclusion or recommendation expressed in this material is that of the authors and the National Research Foundation does not accept any liability in this regard.

REFERENCES

Bowie, L., H. Venkat, and M. Askew. 2019. "Pre-service Primary Teachers' Mathematical Content Knowledge: An Exploratory Study". *African Journal of Research in Mathematics, Science and Technology Education*, 23(3), 286–297. DOI: 10.1080/18117295.2019.1682777

Campbell, D. T., and J. C. Stanley. 1966. *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally

Capraro, R. M., M. M. Capraro, Z. E. Yetkiner, S. Özel, H. G. Kim, and A. R. Küçük. 2010. "An International Comparison of Grade 6 Students' Understanding of the Equal Sign". *Psychological Reports*, 106(1), 49–53. https://doi.org/10.2466/PR0.106.1.49-53

Carpenter, T., L. Levi, M. Franke, and J. Zeringue. 2005. "Algebra in elementary school: Developing relational thinking". *ZDM*, 37, 53–59. https://doi.org/10.1007/BF02655897

Davydov, V. V. 1998. "The Renewal of Education and the Mental Development of Schoolchildren". *Thinking: The Journal of Philosophy for Children, 13* (4):5–7.

Donovan, A.M., A. Stephens, B. Alapala, et al. 2022. "Is a substitute the same? Learning from lessons centering different relational conceptions of the equal sign". *ZDM Mathematics Education* 54, 1199–1213. https://doi.org/10.1007/s11858-022-01405-y

Essien, A. A., & Adler, J. 2016. Operationalising Wenger's communities of practice theory for use in multilingual mathematics teacher education contexts. In A. Halai & P. Clarkson (eds), *Teaching & learning mathematics in multilingual classrooms: Issues for policy, practice and teacher education* (pp. 173–194). Rotterdam: Sense.

Essien, A & Setati, M. 2006. Revisiting the equal sign: Some Grade 8 and 9 learners' interpretations. *African Journal of Research in Mathematics, Science and Technology Education*, 10:1, 47-58, DOI: 10.1080/10288457.2006.10740593.

Essien, A. (2009). An analysis of the introduction of the equal sign in three Grade 1 textbooks. *Pythagoras,* 69, 28-35. doi:https://doi.org/10.4102/pythagoras.v0i69.43

Essien, A., Mathews, C., Tshesane, H., Weitz, M., Abdulhamid, L., Hoosen, T., & Lavans, L. 2023. Developing design principles to enhance pre-service teachers' understanding of number structure and mathematical equivalence in early grade mathematics. *African Journal of Research in Mathematics, Science and Technology Education*, 27:3, 239-254, DOI: 10.1080/18117295.2023.2216608

Fleisch, B. 2008. *Primary Education in a crisis: Why South African school children underachieve in Reading and Mathematics.* Juta and Company Ltd.

Graven, M., 2016. "When systemic interventions get in the way of localized mathematics reform". *For the Learning of Mathematics*, *36*(1), 8–13

Hunter, J. 2007. "Relational or calculational thinking: Students solving open number equivalence problems". *Proceedings of the 30th annual conference of the mathematics education research group of Australasia*, pp. 421–429.

Jones, I., M. Inglis, C. Gilmore, and M. Dowens. 2012. "Substitution and sameness: Two components of a relational conception of the equals sign". *Journal of Experimental Child Psychology*, 11, pp. 166–176, 10.1016/j.jecp.2012.05.003

Kieran, C. 1981. "Concepts associated with the equality symbol". *Educational Studies in Mathematics*, *12*(3), 317– 326.

Kifer, E. 1975. "Relationships between academic achievement and personality characteristics: A quasi-longitudinal study". *American Educational Research Journal*, *12*(2), 191–210.

Knuth, E., A. Stephens, N. McNeil, and M. Alibali. 2006. "Does understanding the equal sign matter? Evidence from solving equations". *Journal for Research in Mathematics Education,* 36, pp.297–312

Machaba, F. 2017. "Grade 9 learners' structural and operational conceptions of the equal sign: A Case study of a secondary school in Soshanguve". *Eurasia Journal of Mathematics, Science and Technology Education, 13* 2017, pp. 7243–7255, 10.12973/ejmste/78017

Matthews, P., B. Rittle-Johnson, K. McEldoon, and R. Taylor. 2012. "Measure for measure: What combining diverse measures reveals about children's understanding of the equal sign as an indicator of mathematical equality". *Journal for Research in Mathematics Education*, *43*(3), 316–350. https://doi.org/10.5951/jresematheduc.43.3. 0316

McNeil, N., C. Hornburg, H. Brletic-Shipley, and J. Matthews. 2019. "Improving children's understanding of mathematical equivalence via an intervention that goes beyond nontraditional arithmetic practice". *Journal of Educational Psychology, 111*, pp. 1023–1044, 10.1037/edu0000337

Rathgeb-Schnierer, E. and M. Green. 2019. "Developing Flexibility in Mental Calculation". *Educação & Realidade, Porto Alegre*, *44*(2), e87078. http://dx.doi.org/10.1590/2175-623687078

Reddy V, M. Visser, L. Winnaar, F. Arends, A. Juan, C. Prinsloo, and K. Isdale. 2016. *TIMSS 2015 highlights of mathematics and science achievement of grade 9 South African learners*. Pretoria: HSRC

Reddy, V., L. Winnaar, F. Arends, A. Juan, J. Harvey, S. Hannan, and K. Isdale. 2022. *The South African TIMSS 2019 Grade 9 Results: Building Achievement and Bridging Achievement Gaps*. HSRC Press, Cape Town.

Reeves, T. 2006. "Design research from a technology perspective". In J.V.D. Akker, K. Gravemeijer, S. McKenney and N. Nieveen (Eds.), *Educational design research* (pp. 52–66). New York: Routledge.

Reeves, T.C., S. McKenney, and J. Herrington. 2011. "Publishing and perishing: The critical importance of educational design research". *Australasian Journal of Educational Technology, 27*(1). https://doi.org/10.14742/ ajet.982

Rittle-Johnson, B., P. Matthews, R. Taylor, and K. McEldoon. 2011. "Assessing knowledge of mathematical equivalence: A construct-modeling approach". *Journal of Educational Psychology, 103*, pp. 85–104, 10.1037/a0021334

Simsek, E., I. Jones, J. Hunter, and I. Xenidou-Dervou. 2021. "Mathematical equivalence assessment: measurement invariance across six countries". *Stud. Educ. Eval. 70*, 101046. doi: 10.1016/j.stueduc.2021.101046

Skemp, R. R. 1976. "Relational understanding and instrumental understanding". *Mathematics teaching*, *77*(1), 20–26.

Spaull, N., and J. Kotze. 2015. "Starting behind and Staying behind in South Africa: The Case of Insurmountable Learning Deficits in Mathematics". *International Journal of Educational Development*, *41*, 13–24.

Taylor, N., and S. Taylor. 2013. "Teacher knowledge and professional habitus, in Taylor, N., van der Berg,S., and Mabogoane, T. Creating effective schools": *Report of South Africa's National Schools Effectiveness Study*. Cape Town: Pearson.

Threlfall, J. "Strategies and flexibility in mental calculation". *ZDM Mathematics Education* 41, 541–555 (2009). https://doi.org/10.1007/s11858-009-0195-3

Venkat, H. and N. Spaull. 2015. "What do we know about primary teachers' mathematical content knowledge in South Africa? An analysis of SACMEQ 2007". *International Journal of Educational Development*, *41*, 121– 130. https://doi.org/10.1016/j.ijedudev.2015.02.002.

Venkat, H., and J. Adler. 2012. "Coherence and connections in teachers' mathematical discourses in instruction". *Pythagoras*, 33(3) https://doi.org/10.4102/pythagoras.v33i3.188

Venkat, H., and M. Askew. 2018." Mediating primary mathematics: Theory, concepts, and a framework for studying practice". *Educational Studies in Mathematics*, *97*(1), 71–92.

Venkat, H., and N. Roberts. 2022. "Children doing mathematics with confidence in the early grades by 2030: What will it take?" In H. Venkat and N. Roberts (Eds.), *Early grade Mathematics in South Africa*, (pp. 175–192). Oxford University Press.

Vergnaud, G. 2020. "A classification of cognitive tasks and operations of thought involved in addition and subtraction problems." In *Addition and subtraction* (pp. 39–59). Routledge.

Vermeulen, C., and B. Meyer. 2017. "The Equal Sign: Teachers' Knowledge and Students' Misconceptions". *African Journal of Research in Mathematics, Science and Technology Education*, *21*(2), 136–147. https://doi.org/10.1080/18117295.2017.1321343

Weitz, M., and H. Venkat. 2017. "Using Assessments to Explore Early and Later Performance in Mathematics." *Improving Primary Mathematics Education, Teaching and Learning: Research for Development in Resource-Constrained Contexts*, 27–43.

Wilmot, E. M. 2018. "Effect of instruction on children's strategies for solving addition problems". *African Journal of Educational Studies in Mathematics and Sciences*, 14, 51–69.