Piloting national diagnostic assessment for strategic calculation

Mellony Graven¹ · Hamsa Venkat²

Received: 13 November 2018 / Revised: 30 September 2019 / Accepted: 10 October 2019 /

Published online: 18 November 2019

© The Author(s) 2019

Abstract

In this paper, we share the results of the piloting of national diagnostic assessments for strategic calculation with Grade 3 learners in South Africa. The diagnostic assessment pilot intervention was focused on promoting strategic use of calculation strategies and aimed to move learners on from the concrete one-to-one counting methods that persist in this grade and into higher grades. Working with a broader team of mathematics education specialists, including international leaders in the field and department of education representatives, we designed a series of assessments that focus on key calculation strategies such as bridging through ten. Each assessment, comprised a pre- and repeat post-assessment, was accompanied by interim lesson starters for teacher use in eight 10-min mental mathematics sessions. These were designed to develop learner fluency in each of the focal strategies and related skills. The pre- and post-assessments allow teachers to track improvement in student learning. The pilot focused on the “bridging through ten” strategy in seven classes across the Eastern Cape and Gauteng provinces. Positive results indicate that the format of these assessments paired with the lesson starters is potentially useful for broader national trialling and implementation. This work feeds into the need for diagnostic assessments that inform teaching focused on strategic efficiency rather than simply accepting correct answers produced through highly inefficient methods.

Keywords Diagnostic assessment · Strategic calculating · Fluency · Mental strategies, number sense

✉ Mellony Graven
m.graven@ru.ac.za

Hamsa Venkat
hamsa.venkatakrishnan@wits.ac.za

¹ Rhodes University, Makhanda, South Africa

² Wits School of Education, University of the Witwatersrand, Johannesburg, South Africa

Introduction and rationale

A majority of learners continue to perform below grade level expectations in mathematics in post-apartheid South Africa. Spaul and Kotze (2015) show that by Grade 4 a majority of learners are already two grades behind curriculum expectations. A key factor in the shortfall is a lack of number sense and a dominance of concrete methods of calculation. For example, Schollar (2008) found that “79.5% of Grade Five and 60.3% of Grade 7 children still rely on simple unit counting to solve problems to one degree or another, while 38.1% and 11.5% respectively, of them rely exclusively upon this method.” (p.iii). This study gathered data from over 150 schools across all provinces and included addition, subtraction, division and multiplication problems with both small and large numbers. Schollar argued from his findings that the inability of the majority of learners to either mentally or in writing manipulate numbers (including large numbers) and their lack of understanding of place value and the base-10 number system was “clearly the single most important cause of poor learner performance in our schools” (p.iii). Our own work and research across schools in Gauteng and the Eastern Cape similarly reveals widespread evidence of learners using tallies for simple low range calculations such as $10 + 10 + 10$ (e.g. Weitz and Venkat 2013). This points to weak number sense and an absence of strategic calculating where the focus is on efficient working that pays attention to number relationships and/or the use of known facts at the level of fluent recall to derive further results.

Unit counting is widely accepted as key in the initial stages of developing number sense and calculating with number. Learners do however need to move on from this approach towards more flexible and efficient strategies as they work with larger numbers and move up the grades—there is strong evidence that not doing this impacts on further learning (e.g. Wright et al. 2006a; Bobis et al. 2005). The persistent use of finger or written tally unit counting for calculating is, of course, not particular to South Africa. Indeed, researchers in other contexts have noted learners in the later years of primary schooling using these methods for adding or subtracting single digit numbers to a number (e.g. see Thevenot et al. 2016; Hopkins and Bayliss 2017). This said, there is little research that suggests that many learners continue to use these methods for calculations involving large numbers. In this respect, we argue that our and Schollar’s findings in the South African context represent an extreme situation requiring urgent attention at a national systemic level.

Structural approaches to working with number (understanding numbers as being constituted as part-part-whole relations rather than individual units) are widely considered important in developing fluent and flexible ways of calculating with number as well as developing conceptual understanding in mathematics (Björklund et al. 2019). Mulligan and Mitchelmore (2009) note that “Recent educational research has turned increasing attention to the structural development of young students’ mathematical thinking” and that “There is increasing evidence that an awareness of mathematical structure is crucial to mathematical competence among young children” (p. 33). The predominance of unit counting as the preferred method of calculation across the South African primary landscape points to a seeming absence (or at least lack of emphasis) of teaching approaches that pay attention to calculation strategies that build on understanding of the structure of number.

South Africa's Curriculum and Assessment Policy document (CAPS) for Grades 1 to 3 states that mathematics should “develop mental processes that enhance logical and critical thinking, accuracy and problem solving that will contribute in decision making” (DBE 2011a, 8–9). It includes developing mental models and strategies for computation that rest on fluency with a range of basic facts that learners should know almost instantly (such as single-digit doubles, bonds to ten and ten plus a single-digit number). The listed mental processes include a range of strategic calculation strategies such as using bridging through ten or compensation for efficient working. Due to our experience of an absence of these strategies being used or taught in classrooms, we focused on these strategies in our professional development (PD) projects. We paired the focus on strategies with use of a range of mental models that support a structural understanding of number such as part-whole diagrams and the empty number line. In an effort to move students beyond one to one counting methods of calculation, we emphasise the relationship between fluency with some basic facts and using calculation strategies flexibly and efficiently. For example, if we automatically know the basic fact that double 20 is 40 then we can find $19 + 19$ by doubling 20 and subtracting 2 (using doubling and compensation as a strategy).

It is widely accepted that national assessments and assessment policies greatly impact teaching (Elmore et al. 1996). Ruthven (1994) argues that assessment plays a key role in enabling reform and that a “change in public assessment is the key to wider change in curriculum and pedagogy; more colloquially expressed as the ‘WYTIWYG’ syndrome: namely ‘what you test is what you get’” (p. 433). The Mathematics Annual National Assessments (ANAs) of the Department of Basic Education (initiated in 2011) for Grades 1–6 and 9 were criticised for perpetuating weak number sense and their failure to support teachers in addressing poor performance. A focus on correct answers fed into acceptance of inefficient counting-based strategies thus perpetuating, rather than addressing, problems of progression (Graven and Venkat 2014). The ANAs were abandoned in 2016 which created the space for new forms of assessment that would support teachers throughout the year. We saw the lack of assessment of the number sense that underlies fluent, flexible and strategic mental (and written) working in the previous Annual National Assessments (ANAs) and the way in which the assessments were influencing teaching practice as problematic (Graven and Venkat 2014). We decided that in order to shift teacher practice nationally we should attempt to influence national assessment practices to promote progression beyond counting strategies to developing a rich and connected understanding that draws on understanding of the structure of number to calculate flexibly and efficiently.

As the two South African Numeracy Chairs (SANCs) at Rhodes University (in the Eastern Cape) and Wits University (in Gauteng), we are mandated to both search for ways forward to the challenges of mathematics teaching and learning in primary schools and to feed tried and tested innovations into the national landscape. Our Chairs have worked collaboratively since inception in 2011. An advantage of working across two very different provinces (Gauteng being the wealthiest and most urban and the Eastern Cape being one of the poorest and most rural) is that we have been able to identify both commonalities and differences in the needs of teachers and learners across the provinces. Our research and PD work across these provinces pointed to (i) the challenge of weak number sense and strategic thinking when calculating in both teaching and learning and (ii) opportunities for early intervention that focus on

progression of learners towards use of more efficient strategies (see Graven and Venkat 2017). Responding to these challenges and opportunities, within the national context of an emerging space for new forms of assessment that feed back into teaching and learning (as opposed to the ANAs which were summative and at the end of the year), we jointly initiated and led the Grade 3 diagnostic assessment collaboration. The project began in 2016 with a 5-day working session attended by members of both Chair teams, two international experts in early mathematics teaching and learning (Bob Wright and Mike Askew), with representation from the Department of Basic Education (DBE at national, provincial and district level), the professional Association of Mathematics Education of South Africa (AMESA), the Southern African Association of Research in Mathematics Science and Technology Education (SAARMSTE) and the NGO community. Classroom-based diagnostic assessments were chosen for two reasons: firstly, their key purpose is to feedback formatively into teaching, a purpose that our DBE partner indicated that they were looking to support; and secondly, because this model provided a low-stakes assessments as a national alternative to the relatively high-stakes ANAs that had previously been rejected as detrimental to teaching and learning development.

Grade 3 was chosen because it is the last grade of the Foundation Phase (Grade R, 1, 2 and 3) and in curriculum terms, by the Intermediate Phase (Grades 4–6), fluent flexible calculation strategies are largely assumed and emphasis shifts towards the use of written methods of breaking down into place value parts and the vertical algorithm for addition and subtraction calculations (DBE 2011b). At this point, pen and paper instructions of algorithms often replace learners' mental strategies (Heirdsfield and Cooper 2002), even for problems such as $98 + 2$ though often with errors leading to nonsensical answers such as 910 (see Graven et al. 2013). While algorithmic thinking can involve reasoning about structure and generality (see Stephens 2018, drawing on Clark 2016), in our context, rule based following of the vertical addition and subtraction algorithms frequently sidelined any attention to quantitative relations. Establishing strong number sense by the end of the Foundation Phase of learning is thus critically important.

In this paper, we share the ways in which the focus and format of the diagnostic pre- and post-assessments and interim lesson starters were informed by our experience of drawing on the international literature based on number sense development in our intervention work with teachers and learners in the South African context. Furthermore, we share experiences from the early trialling that allowed us to think more comprehensively about the full network of skills and basic facts needed to support development of the focal strategies. This was particularly important in a context which is marked by extremely low attainment. (See Reddy et al. (2016) for the report on South Africa's performance in the most recent Trends in International Mathematics and Science Study (TIMSS).) Thereafter, we report on the findings of our initial pilot of the implementation of the "bridging through ten" assessment and lesson starter cycle across three classrooms in the Eastern Cape and four classrooms in the Gauteng province.

Framing and literature review

Askew (2012), drawing on Kilpatrick et al.'s (2001) work on the intertwined strands of mathematical proficiency, suggests a focus on fluency, reasoning and problem-solving

as the “most visible” strands in mathematical working. Askew’s contextualisation of these three aspects in an example is useful to illustrate how they are interwoven. For working out $36 + 7 = 43$ with bridging through ten you might first add 4, making the 36 up to 40, and then add the remaining 3. Askew (2009) notes that: “This is an effective and efficient strategy, but only if you are fluent in knowing what to add to a number to make it up to the next multiple of ten –speedily and confidently being able to say 36, 4; 58, 2; 67, 3. If children have to use their fingers to count-on, the strategy is pointless; they might as well carry on counting-on in ones.”(27–28). He goes on to argue that regular, short and speedy focus on these basic skills is needed because children must be fluent in these to free up working memory (Askew 2012). Reasoning about the relationships underlying strategic calculating allows attention to equivalences such as $36 + 7 = 36 + 4 + 3$.

Given our specific focus on mental calculation skills, we used the terms “rapid recall”, “strategic calculating” and “strategic thinking” to reflect Askew’s fluency, problem-solving and reasoning categories in our PD work and in the assessments. In our focus on supporting learners to progress from concrete inefficient methods of calculating (usually in the form of unit counting) towards using a range of strategies, we identified a network of basic facts or skills needed for learners to use each strategy efficiently. We refer to questions on the assessment of these basic skills or known facts that need a level of automaticity for efficient use of a calculation strategy, as “rapid recall” items in our assessments. This is because learners are required to rapidly recall such facts when learning about and using the various mental calculation strategies noted in the curriculum. As Askew’s example above shows, a lack of fluency in such basic facts and skills is likely to render a strategy no more efficient than counting in ones.

In designing our assessment items, we drew on the seminal assessment work of our two international participants Mike Askew and Bob Wright and their colleagues in England and Australia respectively. In particular, we drew on the types of assessment items included in the study published by Askew et al. (1997) as *Effective Teachers of Numeracy*. They included in their design various calculation items that pushed for forms of reasoning based on the “appreciation of relationships between numbers and operations” (p. 104) that freed students from calculation. For example: Given that $86 + 57 = 143$ then what is $87 + 57$? $86 + 56$? $860 + 570$? The findings from their study indicated that having a connected understanding of mathematics and using corresponding teaching approaches that emphasise the importance of developing mental skills, strategies and reasoning are important for effective teaching of numeracy. They concluded that highly effective teachers believed that being numerate requires:

- Having a rich network of connections between different mathematical ideas
- Being able to select and use strategies, which are both efficient and effective

These teachers used corresponding teaching approaches that:

- Connected different areas of mathematics and different ideas in the same area of mathematics using a variety of words, symbols and diagrams
- Used pupils’ descriptions of their methods and their reasoning to help establish and emphasise connections and address misconceptions

- Emphasised the importance of using mental, written, part-written or electronic methods of calculation that are the most efficient for the problem in hand
- Particularly emphasised the development of mental skills (p. 3)

The background to the approach that we took built on international evidence pointing to the need to support moves beyond counting-based working into approaches based on part-whole number relationships in early number learning (e.g. Bobis et al. 2005). Van den Heuvel-Panhuizen's (2008) attention to provision and use of structured representations of number, such as part-whole bar diagrams and empty number lines, both of which support attention to number relationships rather than to counting, were integrated into our design of assessment tasks and lesson starter activities for use in the 3-week block between the pre- and repeat post-assessment. The integration of these representations facilitated, in turn, a useful feedback loop into teaching that emphasised multiple ways of representing number and number relationships, a further feature that has been acknowledged as important for gaining a good number sense (McIntosh et al. 1992).

In terms of supporting teaching for attention to number structure and relation, we borrowed approaches suggested by Fosnot and Dolk's (2001) notion of "problem strings": "a structured series of problems that are related in such a way as to develop and highlight number relationships and operations" (Location 2519). This idea of using carefully structured sets of examples to draw attention to structure and relationships through the use of variation has been described extensively in the work of Watson and Mason (2005). We incorporated the approach into our development of what we came to call "reasoning chains". Reasoning chains are sets of connected problems carefully designed on the basis of "root" problems that are amenable to solution by recalled facts. The root problem is followed by related problems that can be solved by reasoning about the relationship of the related problem to the root problem, e.g.:

$$25 + 25$$

$$25 + 26$$

$$25 + 27$$

Our focus was on encouraging teachers to point out to learners that strategic calculating involves seeing connections between the set and using rapid recall of answers to root problems for working efficiently on the follow-up problems.

In addition, we drew on Bob Wright and his colleagues' Learning Framework in Number (LFiN) and particularly the Stages of Early Arithmetic Learning (SEAL) that support first assessing the stage of calculation efficiency that learners are at and then encouraging progression up the stages (see for example, Wright et al. 2006a; Wright et al. 2012; Wright et al. 2006b). For this paper, we focus on Wright and colleagues' emphasis on beginning intervention for learners who have fallen behind, with diagnostic assessments focused on strategic efficiency of the response rather than on the

correctness of the answer. However, since their assessments are based on one to one interviews with learners and individualised intervention, we needed to translate this idea into a model for whole class assessments and intervention. Recall that in South Africa it is the majority of learners rather than a few who have fallen behind (discussed above). Furthermore, classes of more than fifty learners are common and so in this respect we needed to move to assessment and intervention formats that could be administered in whole class contexts. Earlier research in both Chair projects pointed to successful outcomes in adapting Wright and colleagues' individual assessment and remediation approach to group-based assessment and remediation (e.g. Morrison 2018; Wasserman 2017).

Our intention for the assessments and linked lesson starters to be used nationally in whole class contexts led us to consider a written assessment format that could maintain attention to the strategic approach used. We also needed written assessments and reasoning chain activities that could fit within the timeframes of the lesson starter part of lessons. (Lesson times can differ across primary schools ranging from 30 to 50 min) The curriculum and assessment policy suggests that lessons begin with:

Whole class activity: where the focus will mainly be on Mental Mathematics, Consolidation of concepts and allocation of independent activities for 20 minutes per day at the start of the mathematics lesson. During this time the teacher works with the whole class to determine and record (where appropriate) the name of the day, the date, the number of learners present and absent, and the nature of the weather (including temperature in Grade 3). Mental Mathematics will include brisk mental starters such as “the number before 8 is; 2 more/less than 8 is; $4+2$; $5+2$; $6+2$ etc. (2011a, b p.9)

These contextual factors led us to the development of low-stakes diagnostic assessments to be set in a time-limited format, and over the following 3 weeks, a series of 10-min lesson starters focused on linking rapid recall facts and strategic calculating and thinking, with a particular strategy in focus (the bridging through ten (BTT) strategy, in the case of this paper). In this format, the extent of effective completion of items across the rapid recall, strategic calculating and strategic thinking sections of the assessment provided a proxy indicator of the extent of each of these skills, based on marking of the assessments simply according to a right or wrong answer. Differences in the extent of effective completion between the pre- and post-assessments provided indications of the efficacy of the interim lesson starters teaching and learning. This model thereby provided a format that was likely to be straightforward enough for larger-scale roll-out in the event of promising results, while maintaining focus on the flexible number sense that we were seeking to promote.

In the next section, we discuss firstly our research design and the six strategies we selected to focus on. Thereafter, we elaborate on the categories of assessment items developed and the way in which we collated these into 2- to 3-week assessment-lesson starter cycles for each of the six strategies. We then provide an example of the first of the six assessment-lesson starter packages which we used as a pilot to gauge the feasibility of the “packages” in providing teachers with support and guidance on a

coherent assessment-teaching cycle of the bridging through ten strategy. Thereafter, we discuss the data and findings from this pilot.

Research design

Our research methodology relates to developmental teaching experiment research in that it included the three central elements of instructional planning and design, ongoing analysis of classroom events and retrospective analysis of data generated from the course of a teaching experiment (Cobb 2000). That is, we first designed our assessments and series of lesson starters, informed by research in established literature; second, we implemented these in several informal settings with only a few learners and adapted both the content and format of the assessments; third, we implemented the assessments and lesson starter instructional sequence in three different classrooms. Reflections on the trialling in the three classes were summarised for discussion with the broader team and again several adaptations were made to the format and content of both the assessments and the lesson starter sequence of activities. Ethical approval was sought and obtained by all relevant stakeholders including our universities, the provincial departments of education, the schools, teachers and the parents of the learners.

The following six strategies for performing calculations with numbers were chosen as most appropriate for the development of Grade 3 assessment items and lesson starters:

- Bridging through ten
- Jump strategy
- Doubling and halving
- Understanding the relationship between addition and subtraction
- Re-ordering
- Compensation

These were chosen from the Curriculum and Assessment Policy (CAPS) document for the Foundation Phase (Grades R, 1, 2 and 3) (DBE 2011a) and, as noted in the literature, were seen to be important strategies for enabling learner progression from one-by-one counting and for developing a rich and interconnected number sense. Number lines and part-whole bar diagrams were considered to be essential models of and tools (whether as a written representation or as a mental image) for making sense of, and working across, these strategies. Several linked basic facts were identified as essential for the successful use of these strategies. Data from the piloting of the first assessment-lesson starter cycle for the “bridging through ten” strategy is shared in this paper.

We chose a 2- to 3-week cycle with a pre- and post-assessment format focused on each of the six strategies (noted above). The length was chosen because we had thought this would enable two strategies to be done in each of the first three school terms and the fourth term could then allow for consolidation and working across strategies. Furthermore, we considered that eight lesson starter sessions would be sufficient for the teaching of a strategy and that these, with the pre- and post-assessment sessions, could be achieved over a 2- to 3-week period. The extra week would allow teachers to

consolidate a session if they felt it was needed and for unforeseen disruptions to the timetable. The 2- to 3-week cycle was comprised as follows:

- Cycle begins and ends with an assessment in time-limited format
- Marking assessments guides lesson starter “reasoning chain”
- Lesson starter teaching activities aimed at developing fluencies and strategies in 10-min sessions across the 2–3 weeks of lessons following the pre-assessment
- Re-assessment provides feedback on learning and, hence, success of teaching

In respect of the above, for each of the six strategies, we were broadly guided by Askew’s (2012) focus on fluency, problem-solving and reasoning to design three categories of assessment items, namely:

Rapid recall items: designed as items we would expect learners should be able to almost instantly “recall” when asked. These items are noted in the CAPS curriculum particularly under the heading of mental mathematics and include “mental” calculation skills such as adding and subtracting 1, 2, 3, 4, 5 and 10 to any number; place value decompositions of number; and key fact triples between 1 and 20.

Strategic calculating items: designed as items that are too time consuming to do in a unit counting or “procedural” calculation orientation but can be efficiently done by employing one of our six strategies. For example, to do a problem such as $99 + 99$ would be extremely slow through counting on in ones and relatively inefficient using the vertical algorithm (and these methods are highly error prone). Efficient calculation involves the strategy of doubling or bridging through 100. That is, if it is recognised as a problem that is either (i) amenable to rounding to 100, doubling and then compensating (i.e. $99 + 99$ can be thought of as 99×2 which is the same as $100 \times 2 = 200$ then $200 - 1 - 1 = 198$) or (ii) amenable to using a number line—written or mental—to “bridge through 100” (i.e. $99 + 99$ can be pictured as $99 + 1 + 98 = 198$) or “compensate” by subtracting 1 after adding 100 to 99.

Strategic thinking items: designed as items which focused on understanding and using the structure of number, i.e. number properties and number relationships, including knowledge of how operations behave. These items focus on using such knowledge to reduce the amount of calculation needed or to eliminate the need for calculation entirely. For example, given $43 + 138 = 181$ then what is $181 - 43 = \underline{\quad}$? or What is $43 + 139 = ?$

The process of refining assessments and lesson starters prior to piloting

Prior to piloting the first author trialled the use of the bridging through ten draft assessments and lesson starters by administering the assessments and the linked lesson starters to three classes in one Eastern Cape school. This involved spending approximately 10 to 15 min in three different Grade 3 classes at the start of their daily mathematics lesson everyday for 2 weeks (10 school days). On the first day, the assessments were administered. From the second to ninth day, the lesson starter activities for the bridging through ten strategy were taught for both adding and subtracting a single digit number to/from a two digit number. Each day the lesson starter was taught to one class by the first author

with the three teachers of the three classes observing. The lesson starters for the latter two classes were taught by the teachers with the first author observing and providing support. On the tenth day, the assessments were re-administered. The teachers observed the administering of the assessments and provided feedback on their experiences after each lesson starter session. While most of the assessment items and lesson starter activities ran smoothly, some challenges arose that we built into the reformatting and refinement of our assessments and lesson starter activities. The initial format of the assessment had one page of 10 rapid recall items followed by a single second page of ten combined strategic calculating (first 5 items) and strategic thinking items (second five items). One minute was given for the first page and another 2 min given for the second page. As mentioned above, the limited time was intended to discourage slower unit counting or procedural methods. While many students finished the first page before the minute was up, most learners did not get past the first two or three items on the second page, commonly becoming bogged down in unit counting. This meant learners did not get to attempt the strategic thinking items. For the pilot, we thus changed the format of the assessments to three separate pages for each of the three categories of items (with each page timed separately).

Some predictable learner nervousness around being “tested” was noted and since timed assessments or “speed tests” can be particularly stressful (Boaler 2012) we emphasised to learners that the assessments were low stakes (they would not affect their report marks) and that the aim of the assessments was to gauge individual improvements over time. Thus, the aim for each learner was to improve on their performance from the pre- to the post-assessment and it was this improvement that was the focus rather than the mark for the assessments. The explanation to learners appeared to work well as learners seemed comfortable and quite excited to complete as much of the assessment as they could in the given time. This concurred with Stott and Graven’s (2013) experience that it was possible to reduce learners’ negative perceptions and experiences of timed tests when the emphasis is placed on their own improvement rather than comparing performance with others.

When teaching the BTT lesson starters, a range of rapid recall skills were noted as absent among learners. These were not in our rapid recall assessments or in our lesson starter plans as we had assumed them to be in place. So, for example, in order to use the BTT strategy to calculate $26 + 7$, learners must be able to rapidly identify the closest next ten to a number (i.e. the next ten after 26 is 30); know the bond to ten that will get you to the next ten ($26 + 4$ will get me to 30); know the various bonds that make up the number added (i.e. $4 + 3 = 7$); and be able to add the “leftover” part of the number to the multiple of 10 (i.e. $30 + 3 = 33$). While we had predicted that rapidly recalling bonds to ten would be weak and in need of practice, we had assumed stating the next ten after a given number and adding a single digit number to a multiple of ten would be well-established known facts easily rapidly recalled by Grade 3 learners. This was not the case and many learners used fingers and or unit counting to add a single digit number to a multiple of ten. Thus, having realised that many learners did not know these basic facts, we re-designed both our assessments and our lesson starters to explicitly focus on practicing, establishing and assessing these. Thus, the rapid recall assessment was expanded to twenty items (see Fig. 1 Appendix) and the lesson starters built into them rapid recall questions focused on establishing each of these rapid recall elements in support of the strategy. In addition, we realised that learners needed an opportunity to practice working with the strategies outside of the whole class lesson starter setting. In

this respect, we devised two take home worksheets given at the end of session 4 and session 7. These worksheets had items similar to the ones in the assessments and included empty number line and part-whole bar diagram items.

The marking of the assessments according to right or wrong answers, then entered into a spreadsheet, provided an adequate, straightforward and well-understood descriptive overview by which we could judge overall learner improvement. However, as indicated above, the combined assessment of strategic calculating items with strategic thinking items in this initial trial meant that we could not distinguish between improvements in these two categories. The assessment results of the three Grade 3 Eastern Cape classes ($n = 94$) showed an average improvement in correct answers for the ten rapid recall items from 5.5 to 8.4 while the average improvement on the ten combined strategic calculation and strategic thinking items improved from 2.6 to 3.1. While the average improvement on the strategic thinking and strategic calculating items was relatively small in comparison with the pleasing improvement on the rapid recall items, we felt that we had sufficient evidence as to the value of our approach to proceed to piloting with learners and schools across our two provinces. Furthermore, the three teachers involved in the trialling indicated that they found the process valuable and particularly appreciated the focus on developing more efficient calculation strategies.

The final assessments and lesson starters for the pilot

The redesigned assessment items and format involved three single pages of items administered separately, with the times available for each category of items stated, as follows:

- 20 rapid recall items to be completed in 2 min (e.g. $10 = 7 + _$; $50 + 6 = _$ and $40 - 7 = _$)
- 5 strategic calculating items to be completed in 1 min (e.g. $56 + 8 = _$ and $93 - 7 = _$, the first two items were accompanied by a number line for example:

- 5 strategic thinking items to be completed in 1 min (e.g., $67 + 5 = 67 + 3 + _$)

Appendix provides the full assessments for each of the three categories of items. Each of the required basic facts was explicitly built into the rapid recall part of the assessment and into the reasoning chain so that even the most basic of these were not taken for granted (see Appendix Fig. 1). This was important given the possible national context of implementation further down the line. Furthermore, in the assessments and in the reasoning chains, there was emphasis on working across representations and building variation into the items and examples. For example, the empty number line is included in all three categories of assessment items (see Appendix Fig. 1 question 6, Fig. 2 questions 1 and 2, Fig. 3 questions 3 and 5). Part-whole bar diagrams are included in the rapid recall assessment (see Appendix Fig. 1 questions 5 and 9). Variation in terms

of the missing number position is built into all three categories of assessment items (see for example $56 + 8 = _$ $67 + _ = 73$ and $_ + 7 = 82$ in Appendix Fig. 2).

As indicated above, the pre-assessments were followed by eight lesson starter sessions which each began with some fast paced mental warm-ups establishing and practising connected known facts. Variation and the use of a variety of representations were built into the “reasoning chain” sequence over the eight sessions. After the pre-assessment and the eight lesson starter sessions, learners were re-assessed on the assessments in the same way. Day one’s lesson starter outline for the bridging through ten strategy is given below:

Day 1: Bridging through ten – lesson starter reasoning chain

First minute mental warm up: play ‘quiz games’ bonds to ten and to multiples of ten; find the next ten and adding a single digit number to a multiple of 10

Examples:

For bonds to 10 ‘I say 3 you say 7 to make up 10; I say 6 you say ? [4]’ etc. Bonds to 50 ‘I say 44 you say 6 to get to 50. I say 46 you say? [4] I say 41 you say? [9]’ etc.

What is the first multiple of ten (‘the next ten’) that comes after 47? [50] 58? [60] 32? [40] – emphasise that this is not rounding to the nearest ten but finding the next ten on the number line.

What is $50 + 7$? $50 + 9$? $50 + 3$? $50 + 2$? Etc.

Start of the reasoning chain:

Consider: $46 + 7$. We can show this on a number line:

We have to jump forwards 7. Let us jump to the next ten rather than jumping in 1s. What is the next ten after 46? *Then show the 50 on the number line above. What do we add to get to the 50? Show this +4 with an arrow in the number line. We have added 4 but we need to add 7. How much more must we add? Show the +3 with an arrow on the number line from 50 to ? So $50 + 3$ is ? Show the 53 on the number line. So $46 + 7 = 46 + 4 + 3 = 53$. Do another example and have learners solve a few more using this method independently.*

Table 1 indicates how each of our three categories of items (rapid recall, strategic calculating and strategic thinking) relates to one another and the focal strategy (BTT). We also provide in the table examples of items for each of these categories for the example given above in the first lesson starter session of adding a one digit to a two-digit number ($46 + 7$) using the bridging through ten strategy:

The pilot: sampling and the process

The authors sought and obtained access to two government primary schools in Gauteng and one government primary school in the Eastern Cape to pilot the above-described pre-assessment–lesson starter–post-assessment cycle. Below we share the results of our pilot based on administering the assessment at the start and at the end of a 2- to 3-week period in which the teachers in these schools taught the eight lesson starter sessions. South African state schools are classified into five quintiles based on a range of catchment area factors (e.g. income, unemployment rate). Quintile 1 schools are the poorest, while quintile 5 the

Table 1 The relationship between three categories of assessment items/lesson starter examples

Focus	Assessment item/ lesson starter example type	Possible examples
Strategy: using BTT to calculate	strategic calculating	$46 + 7 =$
Related known facts: set of known facts that should be automated for using BTT efficiently	rapid recall	The next ten after 46 is ? How far to next 10 $46 + _ = 30$ Suitable breaking down of addend: $7 = 4 + _$ 10s + a single digit facts: $50 + 3 = ?$
Structural reasoning: using reasoning about the relationships between numbers and operations	strategic thinking	$46 + 7 = 46 + 4 + _$

“least poor” (DoE 1998). The two Gauteng schools were a township quintile 1 school and a suburban quintile 5 school. One township quintile 3 school participated in the Eastern Cape trial. In this Eastern Cape school, a Grade 2 teachers also requested to participate. While the assessment and lesson starter cycle was designed at Grade 3 level, all fluencies and strategies assessed apply also to Grade 2 learners in curriculum terms. Furthermore, since the pilot occurred towards the end of the academic teaching year, it was considered appropriate to include these Grade 2 learners. In terms of the language of instruction, the pilot was conducted in the classroom language of instruction, which across the three schools was a language that the learners were fluent in either as a first or second language (isi-Xhosa in the Eastern Cape school; Sepedi in the Gauteng township school and English in the Gauteng suburban school).

We provided all teachers involved in the pilot with an introduction to the assessments and the lesson starters. We emphasised that the assessments were low stakes and aimed at assessing where learners were at so as to inform teaching and to assess progress over time. The logical flow of the lesson starter outlines was explained (and provided) for each of the eight sessions for use by teachers in class. We emphasised that teachers could adapt these depending on the results of the pre-assessments and learner needs (e.g. they could do more or less of the warm-ups of known facts depending on the knowledge of the class and do additional examples of using the strategy). The pre- and post-assessments in the above format were administered, marked and entered into spreadsheets by our Chair teams. This was done in order to ensure that the timing of the assessments was strictly kept to and that the data was accurately captured. Feedback on the pre-assessments was provided to teachers in terms of identifying the strengths and weaknesses of the learners in the class. A simple total of the number of students in their class who had succeeded in each question was shared with teachers as this quickly pointed to the general strengths and weaknesses across the various items in each of the three assessments. In some cases, exemplar scripts were used to exemplify some of the inefficient methods learners were seen to be using on the scripts.

Results

As the assessments were aimed at mental calculation and learners were encouraged to only write the answer, the pre- and post-assessments were marked simply according to right or wrong answers. As noted in the trialling, this provided sufficient descriptive statistics to indicate whether the intervention led to improvements in each category of items.

Tables 2 and 3 provide the results for the mean number of correct answers in the pre- and post- assessments across the two provinces. These are given as both average marks and percentages—the latter enables comparison in terms of the percentage point improvement for the three categories of assessment items.

The differences in performance on assessments across the two provinces cohere with national data which indicates Gauteng as one of the top performing provinces in the country and the Eastern Cape as one of the lowest. This also reflects the difference in economic wealth of these provinces with Gauteng being the wealthiest most urban province and the Eastern Cape being among the poorest and most rural.

There are several interesting points to make in relation to the pre-assessment results in the above tables. Predictably, and in both provinces, pre-assessment performance was highest in the rapid recall cluster of items and lower in the other two item cluster types. What the pre-assessment results also point to though is problematic weaknesses with the rapid recall items. Since these items are based on basic skills that curriculum expectations indicate Grade 3 children should be fluent in, not having these skills stifles the opportunity for developing strategic calculating and strategic thinking. The Eastern Cape pre-assessment data for rapid recall items indicates that the average for the two classes of Grade 3 learners is just under 27%. This indicates that in 2 min most learners were unable to answer more than 6 of the 20 basic questions (see Fig. 1 Appendix) correctly. While the Grade 2 results in the Eastern Cape are similar, though surprisingly slightly better than the Grade 3 results (31%), the Gauteng data is quite a bit stronger for rapid recall items (56.3% average in the pre-assessment across Grade 3 learners). However, since Gauteng learners mostly only managed to answer just over half of the 20 basic items correctly, this suggests that learners will face challenges when answering questions that require them to use strategies and strategic thinking that are built on and depend on rapid recall of these basic facts. In this respect, it is unsurprising that across all schools, classes and provinces the average pre- and post-assessment results of the strategic calculating and strategic thinking items are

Table 2 Eastern Cape Township school Grade 3 and Grade 2 outcomes of correct responses

	Rapid recall (20 items; 2 min)		Strategic calculating (5 items; 1 min)		Strategic thinking (5 items; 1 min)	
	Mean pre mark (%)	Mean post mark (%)	Mean pre mark (%)	Mean post mark (%)	Mean pre mark (%)	Mean post mark (%)
Grade 3 (<i>n</i> = 65)	5.34 (26.7)	8.14 (40.7)	0.585 (11.7)	1.125 (22.5)	0.095 (1.9)	0.325 (6.5)
Grade 2 (<i>n</i> = 30)	6.24 (31.2)	11.28 (56.4)	0.515 (10.3)	1.065 (21.3)	0.055 (1.1)	0.375 (7.5)

Table 3 Gauteng township and suburban school Grade 3 outcomes of correct responses

	Rapid recall (20 items; 2 min)		Strategic calculating (5 items; 1 min)		Strategic thinking (5 items; 1 min)	
	Mean pre mark (%)	Mean post mark (%)	Mean pre mark (%)	Mean post mark (%)	Mean pre mark (%)	Mean post mark (%)
Township school (<i>n</i> = 50)	9.5 (47.5)	12.26 (61.3)	1.36 (27.2)	0.76 (15.2)	0.04 (0.8)	1.16 (23.2)
Suburban school (<i>n</i> = 84)	12.38 (61.9)	16.64 (83.2)	2.19 (43.8)	2.405 (48.1)	0.465 (9.3)	1.26 (25.2)
All (<i>n</i> = 134)	11.26 (56.3)	15.06 (75.3)	1.88 (37.6)	1.79 (35.8)	0.305 (6.1)	1.225 (24.5)

weak. That is, for the pre- and post-assessments, in the Eastern Cape, most learners were unable to answer more than one strategic calculating item correctly in 1 min and most did not answer any strategic thinking items correctly in 1 min. In Gauteng, most learners were unable to answer more than 2 of the 5 strategic calculating questions correctly in the minute and were unable to answer more than 1 of the strategic thinking items correctly.

When comparing scores of the pre- and post-assessments, we see some improvements across all categories of assessment items, except for the Gauteng township school where for strategic calculating items, there was a drop from getting 2.7 of the 5 correct to only getting 1.5 correct in the post-assessment. Such a drop is not entirely unexpected if one considers that, in the pre-assessment, a learner might have successfully answered the first two or three strategic calculation items ($56 + 8$; $83 - 4$; $93 - 7$) correctly using counting up or down within the minute provided. Learning a newer, more conceptual method could, in the beginning, slow learners down or be more error prone initially than this counting up/down method. Thus, we would expect that more time would be required for fluency in the bridging through ten strategy to emerge and thus improvement in this category of items might take somewhat longer. We expect that this would not be the case when moving onto the “jump strategy” assessment-lesson starter cycle as it would be unlikely for learners to perform better in the limited time on strategic calculation items with addition or subtraction of larger numbers (e.g. 27 ± 45) using counting strategies. In this respect, the benefit of the bridging through ten strategy will really come into effect when it is used in combination with the jump strategy as fluent use of the jump and bridging through ten strategy will allow calculation in a few seconds if each of the rapid recall facts are automatically known (i.e. $27 + 40 = 67$; the next ten is 70; $67 + 3 = 70$; $5 = 3 + 2$; $70 + 2 = 72$). Thus, being able to choose from a range of strategies and then apply the strategies flexibly to a range of calculations is the goal rather than simply accuracy and speed.

Overall, we consider that the post-assessment results point to adequate gains in both provinces in all groups of learners to indicate that there is value in this approach. Furthermore, the improvement in Grade 2 results indicate possible broader applicability of the diagnostic assessment/lesson starter format beyond Grade 3.

Concluding discussion: where to from here?

In this paper, we have focused on the piloting of only one of our six identified strategies whereas, as argued above, the likely benefit and power of these strategies are when the learner can select from the range of strategies and flexibly draw on them and/or combine them. In this respect, following the assessment and teaching of each of these six strategies, we will need to develop a series of combined assessments and activities that engage learners in making decisions about which and what combination of strategies to use. Once all six strategies have been developed, it would be possible to design assessments that enable learners to reflect on their strategy use. For example, they might decide that bridging through ten is preferable to using a doubling and compensation strategy for a problem such as $99 + 99$, and their choice could depend on their personal preference or fluency in a particular strategy. Such flexibility in and meta reflection on the choice of strategy is indeed the longer term aim and coheres with the NCTM's (2014, 1) more recent definition of procedural fluency as being able to "apply procedures accurately, efficiently, and flexibly... and to recognize when one strategy or procedure is more appropriate to apply than another" cited in Russo and Hopkins (2018, 661). Creative assessment forms exist that enable students to engage in meta reflection on a range of strategies (e.g. Russo and Hopkins 2018). In the South African context of evidence showing that the majority of learners do not use any strategy other than inefficient unit counting even for calculations with larger numbers, the first step is for learners to become familiar and fluent in the use of each of the six calculation strategies.

While we would like, in time, to see much stronger post-assessment gains in the strategic calculating and strategic thinking items, as these remain low even after the eight lesson starter sessions, we do expect that these will strengthen as learners (i) become more automated in the necessary rapid recall items, (ii) have more exposure to these types of items across the six strategies and (iii) have more time to practice and develop fluency in each of the strategies used. Particularly pleasing was that all the teachers involved in the small-scale pilot were positive about the feasibility of the assessment lesson-starter activity format and noted that the attention to flexible number skills generated via the assessment/starter combination was both important and useful.

Our preliminary trial and pilot findings shared above suggest that the diagnostic assessment-lesson starter model can support improvement in learner numeracy performance and the development of number sense through a focus on key calculation strategies with linked mental models emphasising a structural understanding of number. The next stage is a broader and more nationally representative Department of Education-led trial of the model with more than one strategy. In these trials, more provinces and a much larger and more representative sample of schools and learners will be selected which will enable statistical analysis of the results. Our suggestion for this trialling is to work with Foundation Phase provincial and district mathematics specialists to support the running of trials in a wide range of districts. These trials would use the same model as shared in this paper and as was used in the initial pilot. This work and our collaboration with the national Department of Basic Education continues in this vein.

Acknowledgments Our thanks to DBE and the NRF for their support of this work. Thanks to our entire team who grappled together in developing these formats, assessments and reasoning chains namely Lawan Abdulhamid, Mike Askew, Nolunthu Baart, Lynn Bowie, Mark Chetty, Corin Mathews, Zanele Mofu,

Samantha Morrison, Debbie Stott, Lyn Webb, Marie Weitz, Bob Wright. Also thanks to Lawan Abdulhamid, Sameera Hansa, Lynn Bowie and Pam Vale who worked with us on the data collection, collation and analysis.

Compliance with ethical standards Ethical approval was sought and obtained by all relevant stakeholders including our universities, the provincial departments of education, the schools, teachers and the parents of the learners.

Appendix

Below, in Fig. 1, Fig. 2 and Fig. 3, we provide the assessment items relating to the three assessment item categories.

NAME							
PART ONE							
2 MINUTES FOR THIS PAGE							
1 $6 + 4 = \square$	11 $50 + 7 = \square$						
2 $2 + 8 = \square$	12 $60 + 3 = \square$						
3 $10 = 7 + \square$	13 $40 + 8 = \square$						
4 8 less than 10 is $\square$	14 $40 - 7 = \square$						
5 <table><tr><td>2</td><td>$\square$</td></tr><tr><td colspan="2">10</td></tr></table>	2	$\square$	10		15 The next 10? <div><div></div><div>48</div><div>$\square$</div></div>		
2	$\square$						
10							
6 <div><div></div><div>5</div><div>$\xrightarrow{+5}$</div><div>$\square$</div></div>	16 $100 + 27 = \square$						
7 $10 - 5 = \square$	17 The 10 before? <div><div>$\square$</div><div></div><div>34</div></div>						
8 $10 - 3 = \square$	18 $\square + 8 = 50$						
9 <table><tr><td>1</td><td>$\square$</td><td>9</td></tr><tr><td colspan="3">$\square$</td></tr></table>	1	$\square$	9	$\square$			19 $30 - \square = 27$
1	$\square$	9					
$\square$							
10 $\square + 2 = 10$	20 $87 = 80 + \square$						
Total out of 20:							

Fig. 1 The bridging through ten rapid recall assessment items

PART TWO 1 MINUTE FOR THIS PAGE	
1 $56 + 8 = \square$	
2 $83 - 4 = \square$	
3 $93 - 7 = \square$	
4 $67 + \square = 73$	
5 $\square + 7 = 82$	
<div style="display: flex; justify-content: space-between; border-top: 1px solid black;"> Total out of 5: </div>	

Fig. 2 The bridging through ten strategic calculation assessment items

PART THREE 1 MINUTE FOR THIS PAGE	
1 $67 + 5 = 67 + 3 + \square$	
2 $98 + 56 = 98 + 2 + \square$	
3	<div style="display: flex; align-items: center; justify-content: center;"> <div style="text-align: center; margin-right: 20px;"> $+ \square$ </div> <div style="text-align: center; margin-right: 20px;"> $+ \square$ </div> </div> <p style="text-align: center;">$98 + \square + \square = 136$</p>
4 $96 + \square = 96 + 4 + 32$	
5	<p style="text-align: center;">$103 - \square = 95$</p>
<div style="display: flex; justify-content: space-between; border-top: 1px solid black;"> Total out of 5: </div>	

Fig. 3 The bridging through ten strategic thinking assessment items

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Askew, M. (2009). Do it again. *Teach Primary Magazine*, 3(3), 27–28 <http://mikeaskew.net/page3/page2/files/Mathematicalfuency.pdf>. Accessed 4th April 2019.
- Askew, M. (2012). *Transforming primary mathematics*. Abingdon: Routledge.
- Askew, M., Brown, M., Rhodes, V., Wiliam, D., & Johnson, D. (1997). *Effective teachers of numeracy: report of a study carried out for the teacher training agency*. London: King's College/TTA.
- Björklund, C., Kullberg, A., & Kempe, U. R. (2019). Structuring versus counting: critical ways of using fingers in subtraction. *ZDM Mathematics Education*, 51, 13–24. <https://doi.org/10.1007/s11858-018-0962-0>.
- Boaler, J. (2012). Timed tests and the development of math anxiety. *Education Week*. <http://www.edweek.org/ew/articles/2012/07/03/36boaler.h31.html>. Accessed 12th Feb 2013
- Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, R., Young-Loveridge, J., & Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: three large scale cases. *Mathematics Education Research Journal*, 16(3), 27–57.
- Clark, D. (2016). *Computational and algorithmic thinking, 2011–2015*. Canberra: Australian Mathematics Trust.
- Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. Kelly & A. Lesh (Eds.), *Research design in mathematics and science education* (pp. 307–334). Nahwah: Lawrence Erlbaum.
- Department of Basic Education (DBE). (2011a). *Curriculum and Assessment Policy Statement. Foundation Phase Grades 1–3: mathematics*. Pretoria: DBE.
- Department of Basic Education (DBE). (2011b). *Curriculum and Assessment Policy Statement. Intermediate Phase Grades 4–6: mathematics*. Pretoria: DBE.
- Department of Education (DoE). (1998). *National norms and standards for school funding in terms of the South African Schools Act, 1996 (Act No. 84, 1996)*. Pretoria: Government Printer.
- Elmore, R. F., Ablemann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: from process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: performance-based reform in education* (pp. 65–98). Washington, DC: Brookings Institution.
- Fosnot, C. T., & Dolk, M. (2001). *Young mathematicians at work: constructing number sense, addition and subtraction*. Kindle Edition: Heinemann.
- Graven, M., & Venkat, H. (2014). Primary teachers' experiences relating to the administration processes of high-stakes testing: the case of mathematics Annual National Assessments. *African Journal of Research in Mathematics, Science and Technology Education*, 18(3), 299–310.
- Graven, M., & Venkat, H. (Eds.). (2017). *Improving primary mathematics education, teaching and learning: research for development in resource-constrained contexts*. Hampshire: Palgrave Macmillan UK.
- Graven, M., Venkat, H., Westaway, L., & Tshesane, H. (2013). Place value without number sense: exploring the need for mental mathematical skills assessment within the Annual National Assessments. *South African Journal of Childhood Education*, 3(2), 131–143.
- Heirdsfield, A., & Cooper, T. (2002). Flexibility and inflexibility in accurate mental addition and subtraction: two case studies. *Journal of Mathematical Behavior*, 21, 57–74.
- Hopkins, S., & Bayliss, D. (2017). The prevalence and disadvantage of min-counting in seventh grade: problems with confidence as well as accuracy? *Mathematical Thinking and Learning*, 19(1), 19–32.
- Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: helping children learn mathematics*. Washington, DC: National Academy Press.
- McIntosh, A., Reys, B. J., & Reys, R. E. (1992). A proposed framework for examining basic number sense. *For the Learning of Mathematics*, 12(3), 2–8.
- Morrison, S. (2018). *Developing early number learning using mathematics recovery principles. Doctoral dissertation*. Johannesburg: Wits University.
- Mulligan, J., & Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, 21(2), 33–49.
- National Council of Teachers of Mathematics (NCTM) (2014). Procedural fluency in mathematics: A position of the National Council of Teachers of Mathematics. Reston: NCTM Ed.It can be found at https://www.nctm.org/uploadedFiles/Standards_and_Positions/Position_Statements/Procedural%20Fluency.pdf.
- Reddy, V., Isdale, K., Juan, A., Visser, M., Winnaar, L., & Arends, F. (2016). *TIMSS 2015: highlights of mathematics achievement of Grade 5 South African students*. Pretoria: DBE/HSRC.
- Russo, J., & Hopkins, S. (2018). Measuring mental computational fluency with addition: a novel approach. 2018. In J. Hunter, P. Perger, & L. Darragh (Eds.), *Making waves, opening spaces. Proceedings of the*

- 41st annual conference of the Mathematics Education Research Group of Australasia (pp. 661–668). Auckland: MERGA.
- Ruthven, K. (1994). Better judgement: rethinking assessment in mathematics education. *Educational Studies in Mathematics*, 27, 433–450. <https://doi.org/10.1007/BF01273382>.
- Schollar, E. (2008). *Final report of the Primary Mathematics Project: towards evidence- based educational development in South Africa*. Johannesburg: Eric Schollar & Associates.
- Spaull, N., & Kotze, J. (2015). Starting behind and staying behind in South Africa. *International Journal of Educational Development*, 41, 13–24.
- Stephens, M. (2018). Developing algorithmic thinking in the primary and junior secondary years. In F.-J. Hsieh (Ed.), *Proceedings of the 8th ICMI-East Asia Regional Conference on Mathematics Education* (Vol. 2, pp. 350–363). Taipei: EARCOME.
- Stott, D., & Graven, M. (2013). Research tensions with the use of timed numeracy fluency assessments as a research tool. *Perspectives in Education*, 31, 79–93.
- Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: a counting model account. *Cognition*, 146, 48–57.
- Van den Heuvel-Panhuizen, M. (2008). *Children learn mathematics: a learning-teaching trajectory with intermediate attainment targets for calculation with whole numbers in primary school*. Rotterdam: Sense Publishers.
- Wasserman, A. (2017). Adapting mathematics recovery individual assessments for group administration in resource-constrained contexts. In M. Graven & H. Venkat (Eds.), *Improving primary mathematics education, teaching and learning: Research for development in resource-constrained contexts*. Hampshire: Palgrave Macmillan.
- Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: learners generating examples*. New York: Lawrence Erlbaum Publishers.
- 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., & Stanger, G. (2006a). *Teaching number: advancing children's skills and strategies*. London: Paul Chapman Publishing Ltd..
- Wright, R. J., Martland, J., & Stafford, A. K. (2006b). *Early numeracy: assessment for teaching and intervention* (2nd ed.). London: Sage publications.
- Wright, R. J., Ellemor-Collins, D., & Tabor, P. (2012). *Developing number knowledge: assessment, teaching and intervention with 7–11 year olds*. London: Sage publications.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.