



The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Pournara, Craig and Hodgen, Jeremy and Sanders, Yvonne and Adler, Jill (2016) Learners' errors in secondary algebra: insights from tracking a cohort from Grade 9 to Grade 11 on a diagnostic algebra test. *Pythagoras*, 37 (1). pp. 1-10. ISSN 1012-2346

Access from the University of Nottingham repository:

<http://eprints.nottingham.ac.uk/38675/1/Pournara%20et%20al%20Tracking%20algebra%20errors%20Pythagoras%202016.pdf>

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see: <http://creativecommons.org/licenses/by/2.5/>


A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

Learners' errors in secondary algebra: Insights from tracking a cohort from Grade 9 to Grade 11 on a diagnostic algebra test

**Authors:**

Craig Pournara¹ 
 Jeremy Hodgen²
 Yvonne Sanders¹
 Jill Adler¹

Affiliations:

¹School of Education,
 University of Witwatersrand,
 Johannesburg, South Africa

²School of Education,
 University of Nottingham,
 United Kingdom

Corresponding author:

Craig Pournara,
 craig.pournara@wits.ac.za

Dates:

Received: 04 Apr. 2016

Accepted: 02 Sept. 2016

Published: 31 Oct. 2016

How to cite this article:

Pournara, C., Hodgen, J., Sanders, Y., & Adler, J. (2016). Learners' errors in secondary algebra: Insights from tracking a cohort from Grade 9 to Grade 11 on a diagnostic algebra test. *Pythagoras*, 37(1), a334. <http://dx.doi.org/10.4102/pythagoras.v37i1.334>

Copyright:

© 2016. The Authors.
 Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

It is well known that learner performance in mathematics in South Africa is poor. However, less is known about what learners actually do and the extent to which this changes as they move through secondary school mathematics. In this study a cohort of 250 learners was tracked from Grade 9 to Grade 11 to investigate changes in their performance on a diagnostic algebra test drawn from the well-known Concepts in Secondary Maths and Science (CSMS) tests. Although the CSMS tests were initially developed for Year 8 and Year 9 learners in the UK, a Rasch analysis on the Grade 11 results showed that the test performed adequately for older learners in SA. Error analysis revealed that learners make a wide variety of errors even on simple algebra items. Typical errors include conjoining, difficulties with negatives and brackets and a tendency to evaluate expressions rather than leaving them in the required open form. There is substantial evidence of curriculum impact in learners' responses such as the inappropriate application of the addition law of exponents and the distributive law. Although such errors dissipate in the higher grades, this happens later than expected. While many learner responses do not appear to be sensible initially, interview data reveals that there is frequently an underlying logic related to mathematics that has been previously learned.

Introduction

This article has its origins in the discourse of poor performance and failure in school mathematics in South Africa and other developing countries. That said, not all learners in South Africa are failing mathematics but the vast majority are not coping (Spaull, 2013). Attempts to seek and gather evidence of improved learner performance are often thwarted by a lack of longitudinal data or instruments that are not sufficiently sensitive to detect change at lower performance levels. Furthermore, in developing countries where these problems are most acute, there is seldom adequate funding to develop more appropriate and sensitive measures.

In our project we seek to shift the discourse from 'learners can't do' to 'what learners can do is ...'. This requires sensitive diagnostic instruments that enable us to measure change over time, even when learners' performance is poor. However, there are no adequate local diagnostic instruments available for this purpose. Consequently, we drew on a selection of algebra items from the Concepts in Secondary Maths and Science (CSMS) tests (Hart & Johnson, 1983) and tracked the performance of a group of 250 learners on these items across their Grade 9, 10 and 11 years of schooling. This has enabled us to investigate the kinds of algebraic errors learners make and how these change over time.

Our findings show that many learners in Grade 11 still have difficulty with aspects of basic algebra. This reinforces the comments made on the performance of Grade 12 learners on the National Senior Certificate Paper 1 each year. For example, in 2012 the examiners commented:

Many of the errors made in answering this paper have their origins in poor understanding of the basics and foundational competencies taught in the earlier grades. For example, algebraic manipulation, factorisation, solution of equations and inequalities. (Department of Basic Education, 2012, p. 12)

In this article we draw on quantitative and qualitative data to illuminate the kinds of errors that learners make in relation to basic algebra. Thus we go beyond merely noting *that* learners make errors to provide insights into what errors are being made and the extent to which these decrease or persist over time. In addition, we provide evidence of the kinds of rational descriptions that learners provide of their thinking as they work on algebraic tasks. We pay particular attention to errors of conjoining, difficulties with negatives and brackets and learners' tendency to evaluate expressions rather than leaving them in the required open form.

Read online:

Scan this QR code with your smart phone or mobile device to read online.

Our analysis also shows evidence of curriculum impact, that is, how learners apply new procedures and laws they have been taught, but in inappropriate ways. Typically, this changes over time as learners become more familiar with the new procedures and when they should be applied but the cycle continues as they learn further procedures and then apply these inappropriately. This finding gave rise to more substantial concerns regarding the use of the CSMS items outside the UK and with older learners. Based on a simple Rasch analysis of the Grade 11 data we show that the items performed sufficiently well for the purposes of this study although the nature of the errors is clearly impacted by learners' exposure to more mathematics as they move from Grade 9 to Grade 11.

Literature review and theoretical perspectives

Internationally, there has been much interest in learners' errors in mathematics for the past 30 years. The early work drew strongly on constructivist perspectives (e.g. Ben-Zeev, 1998; Borasi, 1994; Olivier, 1989; Radatz, 1979), others adopted a socio-cultural perspective (e.g. Ryan & Williams, 2007) and more recent work comes from a discursive approach (e.g. Brodie & Berger, 2010).

In South Africa, there has been a resurgence in the focus on learner errors and error analysis in mathematics in recent years (Brodie, 2014; Brodie & Berger, 2010; Herholdt & Sapire, 2014; Makonye & Luneta, 2014; Shalem, Sapire, & Sorto, 2014). This is not surprising given the poor performance in mathematics in SA and the associated focus on mathematics professional development. The work continues to reflect a range of theoretical perspectives to describe and explain errors.

We work from a sociocultural perspective but draw on some aspects of discursive approaches (Sfard, 2008). In agreement with all the work referred to above, we consider learners' errors to be rational attempts to make sense of mathematics. We acknowledge that, for the most part, errors are systematic, persistent and resistant to 'fixing' through instruction.

Research on the transition from arithmetic to algebra has been ongoing since the 1970s (e.g. Collis, 1978) and much has been learned about learners' difficulties in making the transition. Following the CSMS study, the Increasing Competence and Confidence in Algebra and Multiplicative Structures (ICCAMS) study used the CSMS algebra items to investigate changes in learners' understanding of algebra 30 years later. It was found that more learners were achieving low marks, fewer were scoring very high marks and there was little change in the kinds of errors learners made (Hodgen, Brown, Coe & Küchemann, 2012).

Linked to the CSMS study, Küchemann (1981) identified six ways in which learners interpret letters; these ideas have influenced much of the work that has followed. One interpretation is that letters are abbreviations for the names of objects. Asquith, Stephens, Knuth, and Alibali (2007) have

shown that this error tends to disappear by the 10th grade in the United States. Many have reported the persistent tendency for learners to think that a letter stands for a single, specific number (Asquith et al., 2007; Booth, 1984; Collis, 1975). Christou and Vosniadou (2012) have shown that 10th grade learners in Greece tend to substitute natural numbers in algebraic expressions even when they recognise that letters can stand for any number. There is also a great deal of evidence which shows that prior knowledge of arithmetic has a negative impact on the learning of algebra. This is typically manifest in errors relating to 'lack of closure' where learners do not accept expressions such as $a + 2$ as final answers. (e.g. Booth, 1984; Christou & Vosniadou, 2012; Collis, 1978; MacGregor & Stacey, 1997).

Despite all that has been learned, algebra remains a substantial obstacle for many learners of secondary mathematics internationally as borne out in the performance of SA learners on the CSMS items 45 years after they were first developed.

The design and content of the Concepts in Secondary Maths and Science algebra test

The algebra items used in this study were first developed and administered to a nationally representative sample of English learners in the 1970s as part of the CSMS study. The design of the items was strongly informed by Piagetian ideas that were dominant at the time and were intentionally designed to be recognisably connected to the UK mathematics curriculum of the time, but required learners to: use methods 'which were not obviously 'rules' (Hart & Johnson, 1983, p. 2), avoid 'excessive computation' (p. 22) and make use of simple numbers in non-routine problems.

The CSMS algebra test focuses on generalised arithmetic, the use of symbols to denote numbers and letters as variables (Collis, 1975; Küchemann, 1981). Consider the following item: 'If $e + f = 8$, $e + f + g = \dots$ '. This was designed to test whether learners would accept the 'lack of closure' (Collis, 1975) of the expression $8 + g$, that is, to see it as an entity in its own right (Sfard, 1991). Without this, learners tend to see the expression as an instruction to do something and give numerical responses (such as 9 or 12) or the 'compressed' response $8g$.

Research design and methodology

Over the period 2010 to 2014 our project worked in 10 schools in the Johannesburg area in South Africa. Six of the schools are in so-called townships where there are only basic resources for teaching and learning. The other four schools are low-fee schools located in suburban areas. Here too resources are limited although these schools are typically better equipped than those in the townships. In all schools, English is the language of instruction but almost no learners speak English as their main language.

Initial selection of sample

We set out to track a cohort of learners from Grade 9 to Grade 11 (approximately 15 to 17 years old) across all project

schools, regardless of whether they progressed to the next grade at the end of the year and regardless of whether they chose Mathematics or Mathematical Literacy¹ in Grade 10 or Grade 11. In 2011 we selected a sample of approximately 1500 Grade 9 learners with the sole criterion being that their mathematics teachers in that year had been participating in the professional development activities offered by the project. While we anticipated some attrition over the three-year period, the attrition rate was far greater than expected and in 2013 we ended up with a tracked cohort of only 250 learners. There are many reasons for this attrition, including learners moving schools, teachers not administering the test to learners who had failed or who were taking Mathematical Literacy and the withdrawal of one school from the project.

The limitations of this sample are that it became a small, opportunistic, non-random sample with a smaller proportion of weaker learners than the intended sample. Nevertheless, it provides a unique snapshot of learners' performance in algebra across the transition from Grade 9 to Grade 11.

The test instruments and coding

The test instruments did not consist only of CSMS items. We designed three tests, one for each grade, containing curriculum-related algebra and functions items and a selection of CSMS algebra items which were common to all three tests. We focus here only on the CSMS items, which have been validated in the UK in both studies. In selecting the items, we chose only those clusters of items that had been 'levelled' in the CSMS hierarchy (see Hart, 1981, for more details). However, we excluded some of the more difficult items, partly because of the length of the test and partly because we anticipated that very few learners would cope with them. The order of the items differed from the UK test.

In 2010 the CSMS items were piloted with Grade 8 and Grade 10 learners in schools similar to the project schools. The results of this pilot indicated that the items and the associated coding scheme were appropriate in the SA context. The three full tests were then piloted in project schools with different cohorts prior to administering to the tracked cohort.

The tests were administered in October each year (i.e. near the end of the school year). In 2011 and 2012 they were administered by project team members. In 2013 they were administered by team members in some schools and by teachers in other schools. We deliberately wanted to include teachers in the research process but, with hindsight, this was not wise because some teachers did not follow the directions given regarding who was to write the test, which ultimately impacted on the size of the tracked cohort.

The test codes were adapted from the CSMS codes. In the main, we used the most prevalent CSMS codes and then

1. In South Africa, all students are required to take mathematics for all 12 years of schooling. Typically they will choose between Mathematics or Mathematical Literacy at the end of Grade 9. Many learners change from Mathematics to Mathematical Literacy in Grade 10 or Grade 11 if they are not coping with Mathematics.

added additional codes to capture the wide range of responses that would otherwise have merely been coded as 'other'. Coding was carried out by trained research assistants and moderated by senior members of the project team.

Interview data

In 2012 and 2013, following the analysis of pilot data and the test scripts, several small studies were conducted by postgraduate students (Honours level) who were supervised by senior members of the project team. These studies involved an analysis of scripts and one-on-one task-based interviews where each postgraduate student interviewed learners from project schools on questions similar to selected test items in order to investigate learners' thinking in relation to introductory algebra. We draw on these studies to illustrate our claims about learners' thinking and strategies in relation to their errors. We explicitly acknowledge the postgraduate students as authors of the work reported here. However we have re-transcribed and re-analysed all interview extracts that are included in this article.

Ethical issues

Ethical clearance was obtained from the university (2010ECE60C) and from the Gauteng Department of Education for this research. Informed consent was obtained from schools, parents and learners for both the tests and the interviews. Names of learners and schools have been kept confidential and we have not reported separately on particular schools. Pseudonyms have been used in all learner transcripts. Permission was obtained from the postgraduate students to use their data as part of the wider project research.

Reliability and validity

As noted above, the validity of the items for the SA context was confirmed through piloting. The reliability of the coding was increased through a moderation process: all learner interviews were re-transcribed and they were then re-analysed by two of the authors to confirm the initial interpretation.

Since the CSMS items had been developed for Year 8 and Year 9 learners in the UK, it was necessary to validate it for older learners in SA. We therefore conducted a simple Rasch analysis, using Winsteps (Version 3.73)², to assess the validity of the scale for older learners. We chose the Grade 11 data since these learners had performed better on the test, thus enabling better discrimination of items at the higher and lower ranges of difficulty.

We used three fit statistics to judge the quality of the selected CSMS items as an instrument for measuring learners' understanding of algebra in the SA context: infit mean square (MNSQ), outfit MNSQ and item-scale correlation. Since the test is a low stakes diagnostic test, we took a relatively liberal approach and so values of infit and outfit MNSQs higher than

2. Following Linacre (2015), this analysis was re-performed after the data was cleaned to recode the most unexpected responses as *missing*.

2.0 or lower than 0.5 were regarded as a cause for possible concern (Linacre, 2002). For the item-scale correlation, the discrimination index should ideally be above 0.4. A discrimination index below 0.2 is regarded as low and items with such a score may be candidates for deletion (Wu & Adams, 2007). There were four items with problematic fit statistics: Q2, Q5.1, Q9.1 and Q10.1. However, none of these questions is considered to be a threat to the scale.

Learners' performance on the Concepts in Secondary Maths and Science items

A summary of learners' performance on the items across the three years is provided in Table 1. The figures indicate the item facility (i.e. the percentage of learners who answered each question correctly). This is followed by a brief discussion of the results.

The general performance of the learners over the three years is low but this was not unexpected. There is a general increase from Grade 9 to Grade 11, although there are six items where performance dipped from Grade 9 to Grade 10 and two items

where performance dropped from Grade 10 to Grade 11. Several of these are discussed in later sections. The four items that Grade 11 learners found most difficult were Q2, Q8.4, Q10.4 and Q11. More than 16% did not attempt the latter three items in Grade 11. By contrast fewer than 1% of learners omitted Q2 in Grade 11 although only 1% answered it correctly.

It is surprising that Q1.5, Q1.6 and Q1.9 were among the most difficult items for learners at Grade 11 level since these look similar to questions found in Grade 8 and Grade 9 textbooks in SA. Reflecting on Question 1 as a whole, it is noticeable that the performance on questions with brackets or negatives is well below the other items for all three years. For example Q1.3 and Q1.5 are very similar except for the operation in the bracket. Yet the performance on these two items is considerably different, particularly in Grade 9 and Grade 10. Other errors include inappropriate application of the addition law of exponents and the distributive law. These errors are discussed in more detail in the next section. The poor performance on Q1.2 appears to be related to learners' expectation that they should do something with the unlike terms to produce an answer.

TABLE 1: Percentage of correct responses to each item per grade.

Item	Percentage of correct responses		
	Grade 9	Grade 10	Grade 11
1.1 Simplify $2a + 5a$	86	73	86
1.2 Simplify $2a + 5b$	34	37	46
1.3 Simplify $(a + b) + a$	25	21	35
1.4 Simplify $(2a + 5b) + a$	44	43	52
1.5 Simplify $(a - b) + b$	6	11	30
1.6 Simplify $3a - (b + a)$	8	14	30
1.7 Simplify $a + 4 + a - 4$	31	42	59
1.8 Simplify $3a - b - a$	34	43	56
1.9 Simplify $(a + b) + (a - b)$	7	12	30
2 Which is larger, $2n$ or $n + 2$? Explain.	0	1	1
3.1 Add 4 to 8	82	74	78
3.2 Add 4 to $n + 5$	36	46	61
3.3 Add 4 to $3n$	26	37	56
4.1 Multiply 8 by 4	70	70	77
4.2 Multiply $n + 5$ by 4	9	22	34
4.3 Multiply $3n$ by 4	64	64	66
5.1 If $a + b = 43$, $a + b + 2 = ?$	82	85	87
5.2 If $n - 246 = 761$, $n - 247 = ?$	30	40	46
5.3 If $e + f = 43$, $e + f + g = ?$	15	21	41
6.1 Determine a , if $a + 5 = 8$	78	85	86
6.2 Determine b , if $b + 2 = 2b$	8	14	30
7.1 If $u = v + 3$, $v = 1$, $u = ?$	50	60	71
7.2 If $m = 3n + 1$, $n = 4$, $m = ?$	32	52	66
8.1 Area of rectangle 3×4	52	59	61
8.2 Area of rectangle 6×10	52	58	61
8.3 Area of rectangle $n \times m$	62	70	67
8.4 Area of rectangle $5 \times (e + 2)$	2	5	11
9.1 Perimeter, if sides $9 + 1 + 10 + 2$	95	92	89
10.1 Perimeter, if sides $e + e + e$	81	86	88
10.2 Perimeter, if sides $h + h + h + h + t$	56	64	75
10.3 Perimeter, if sides $u + u + 5 + 5 + 6$	56	50	64
10.4 Perimeter, if n sides of length 2	12	18	23
11 Cakes cost c rand and buns cost b rand each. If I buy 4 cakes and 3 buns, what does $4c + 3b$ stand for?	20	22	28

The slight drop in performance on Q3.1 from Grade 9 to Grade 11 is surprising, as is the small size of the gain on Q4.1, Q8.1 and Q8.2. One possible explanation is that asking simple arithmetic questions in the context of an algebra test confuses learners. Alternatively, given that the first part of our test consisted of curriculum-level questions, some learners might have expected that these questions were more complex than intended.

With regard to Question 5, there is a noticeable contrast in performance between Q5.2 and Q5.3 in Grade 9 and Grade 10, although this was expected. It is also worth noting that responses to Q5.2 were unstable with many learners moving from a correct response in one year to an incorrect response the following year. For example, in Grade 9, 80 learners gave the correct answer but of these 44% got it incorrect in Grade 10. Similarly in Grade 10, 100 learners gave the correct answer but of these, 30% got it incorrect in Grade 11. It may be that this item is more error-prone than initially appreciated. If one merely glances at the question it is quite easy to focus on the increase on the left side and assume the same on the right side whereas it requires paying more careful attention to the relationships to recognise that such reasoning is faulty.

Question 6 and Question 7 involved substitution. Learners had particular difficulty with Q6.2. Given the prevalence of premature closure in their responses across the test, it is possible that many learners could not make sense of the equation $b + 2 = 2b$ since for them $b + 2$ can be written as $2b$ and so they did not see the statement as one of mathematical equivalence in the way that was intended.

Question 9 and Question 10 required algebraic expressions for the perimeter of polygons. While performance improved systematically on Q10.2, performance on Q10.3 dropped in Grade 10 and was still poorer than Q10.2 in Grade 11. The higher number of errors is likely due to the combination of letters and numbers in Q10.3 with 23% of Grade 11 learners making conjoining errors and hence expressing the perimeter as $18u$.

Question 11 was the only CSMS item selected that focused explicitly on letter as object. Not surprisingly many learners treated the letters as objects with at least 35% of learners each year stating that $4c + 3b$ stands for 4 cakes and 3 buns.

Common learner errors

While learners' poor performance is reason for concern, an analysis of their errors provides insight into their reasoning

and shows evidence of rational albeit incorrect strategies. These insights enable us to focus on what learners can do rather than on what they can't do. In this section we focus on key errors and illustrate these with examples from Questions 1, 3, 4 and 5. We also include interview data from the postgraduate students' research projects.

The most common errors involve conjoining and premature closure, negatives and subtraction, multiplication and indices, the equality relationship and evaluating letters rather than accepting an open expression as a final answer. We also provide evidence that learners change their strategies depending on particular details of a question.

Errors involving conjoining and premature closure

Errors involving conjoining or premature closure occur in the responses to many items as shown in Table 2. While the actual number of conjoining errors decreases by Grade 11, it still constitutes a large proportion of the errors on several questions.

In Q1.2 it could be argued that the nature of the question prompts learners to do something and so a reasonable response is to conjoin $2a$ and $5b$. The interview data provides insight into how learners produce conjoined answers. Mashazi (2012) asked Lizwe in Grade 9 to 'add 5 to $3x$ ' and to 'multiply $3x + 1$ by 5':

- Lizwe: They said add 5 to $3x$ so I said 5 plus 3 equal to 8 then after that I took the x and put it next to 8 to get $8x$. (p. 20)
- Lizwe: They say multiply $3x + 1$ by 5 so I first took the three and one and I added it because 3 plus 1 is 4 then I put the x and times by 5 and got $20x$ (p. 23)

When asked to 'multiply $t + 2$ by 3' Shenaaz, also Grade 9, adopted the same strategy (Govender, 2012):

- Shenaaz: Then [question] 2.2 was $t + 2$. So I said, 2 times 3. And isn't 2 times 3 which is 6. So then after I put the 6 and then I put the t .
- Govender: Okay, how did you get the 6?
- Shenaaz: Okay, I said 2 times 3 which is 6. Then I brought down the t . Cos you can't multiply the t by something else. It's gotta come after you know.

The learners 'simplify' the binomial by attending first to the numbers, then appending the letter and then multiplying by the constant. Their use of language shows explicitly that when they conjoin unlike terms, they ignore the letter and then append it as one writes units in a measurement problem. Both learners use 'put' to explain how the letter comes to be

TABLE 2: Percentage of responses showing conjoining error.

Item	Typical errors	% of all responses			
		Grade 9	Grade 10	Grade 11	
1.2	Simplify $2a + 5b$	$7ab$	53	50	42
1.4	Simplify $2a + 5b + a$	$7ab, 8ab$	25	18	14
3.2	Add 4 to $n + 5$	$9n$	11	12	7
3.3	Add 4 to $3n$	$7n$	25	32	17
5.3	If $e + f = 8$ $e + f + g = ?$	$8g$	17	11	10

placed next to the number and Shenaaz also says 'I brought down the l ' illustrating that the letter is attended to last.

The error of attending first to numbers is persistent in terms of the proportion of errors, as illustrated in Table 3. The numbers in the questions are either added or multiplied. In some cases the letter is conjoined (e.g. $7n$) while in others it is not (e.g. $n + 9$).

In the sections that follow we continue to point out errors involving conjoining although they are no longer in the foreground. It is worth noting at this point that when learners conjoin, their talk typically reflects an operational view of the expression they are dealing with. This is evidenced in treating subtraction as 'take away' and in counting the numbers of things they are working with, for example the numbers of b 's in an expression.

Errors involving subtraction and negatives

Learners' errors in simplifying algebraic expressions suggest they may not be paying attention to signs and operations. This is more visible when the question involves subtraction and negatives and it appears that learners are focusing mainly on letters and numbers. For example in Q1.5 more than 10% of Grade 9 and Grade 10 learners gave answers of $a \pm 2b$ and one of the most common errors in Q1.6 was $4a - b$. Such responses come from collecting terms and ignoring the syntax of the expression.

Kalidheen (2012) provides evidence of Grade 11 learners who continue to struggle with algebraic syntax involving negatives. For example, she asked Simon to respond to the following question: subtract $2b$ from 8, $2b - b$, $2b - a$ and $2b - 2a$.

- Simon: It's going to be 8 minus $2b$ and the answer is going to be 8 minus $2b$. The answer will still be 8 mam ... cos... er here we have 8. We have numbers only and not letters and this side on the right we have two b 's so there is no way we can take it out from 8 while there's no b 's.
- Kalidheen: So you've got subtract $2b$ from 8 and your answer is 8. Right?
- Simon: Still 8.
- Kalidheen: So you left out the $2b$. Why?
- Simon: Cos, the $2b$ mam ... cos ... it is what, it is what, it had to be subtracted from 8 and we couldn't.
- Kalidheen: Ok, so what happens if I make this $2b$ minus b (i.e. $2b - b$)? What will happen then?
- Simon: $2b$ minus b ? No mam it's, two b 's minus one b so it's gonna be b .
- ...
- Kalidheen: And then, if I had ... ok let's say I change this to $2b$ minus a .

Simon: $2b$ minus a ? Mam you couldn't ... it can't be ... it can't be done cos there's no, there's no a 's here, it's only b 's.

Kalidheen: Ok, so, what would happen here? What will the answer be?

Simon: No mam, it will be left like this (i.e. $2b - a$).

Simon provides the correct answers when both terms contain letters. However, when only one term contains a letter, he eliminates the term because it cannot be subtracted from 8 'while there's no b 's'. Yet he does not eliminate the term in a when dealing with $2b - a$. He is one of many learners who uses a different strategy depending on whether both terms contain letters. His responses also show that he is working with a partitioning structure of subtraction, that is, as 'take away' (Haylock, 2006). Such a view of subtraction is inadequate for algebra but it goes some way to explaining why learners might claim incorrectly that $8 - 2b = 8$ but also claim correctly that $2b - a$ cannot be done 'cos there's no a 's here [to take away]'.

In contrast, other interviews show evidence of learners paying explicit attention to signs and operations yet over-generalising methods from equation solving when working with expressions (e.g. Gumpo, 2011; Mashazi, 2012). Mashazi (2012, p. 21) asked Themba (Grade 9) to explain his solution to: 'Simplify $3x + 2 + x$ ':

Themba: They said simplify. The first thing that I did, I grouped the like terms then I got the answer.

Mashazi: How did you group your like terms?

Themba: They said $3x + 2 + x$ then I said eh, when x comes between $3x$ and 2 it changes the sign to negative x , then I said $3x - x + 2$, then I got the answer for $3x - x$ which is $2x$ and I left the 2 there, then I said $2x + 2$ which gave me $4x$.

Mashazi: Let us go back to where you added the like terms. You said when x moves closer to $3x$, it changes the sign. Why?

Themba: Because that is how mam taught me that when x eh, when the equation moves to the other side it changes the sign.

Mashazi: Oh! When it is an equation. Is this an equation?

Themba: Yes, ma'am.

Mashazi: Why do you say it is an equation?

Themba: Because, ma'am, it has the variables.

Themba is confident in his selection and use of a strategy, albeit an inappropriate one. He changes the sign if a term moves to another position in the expression and his justification reveals some confusion between an equation and an expression. It also reflects inappropriate criteria for defining an equation and incorrect description of the objects he is working with, such a referring to a term as an equation.

TABLE 3: Percentage of errors showing focus on numbers.

Item	Typical errors	% of all errors			
		Grade 9	Grade 10	Grade 11	
3.2	Add 4 to $n + 5$	9n, 9	32	38	34
3.3	Add 4 to $3n$	7n, 7, 7 + n	47	59	49
4.2	Multiply $n + 5$ by 4	$n + 20, 20, 20n, n + 9, 9$	46	59	54

Errors related to multiplication and indices

There is much evidence of learners applying the distributive law and the addition law of exponents inappropriately across various items in Question 1, particularly in Grade 10. Such errors account for the dip in the percentage of correct responses in Q1.1 and Q1.3 in Grade 10. By Grade 11 these errors were less frequent and this partially explains the substantial increase in the percentage of correct responses in items involving brackets.

Errors involving the addition law for exponents, that is, $a^m \cdot a^n = a^{m+n}$, are best illustrated by Q1.1, Q1.4 and Q1.5 as shown in Table 4. In Q1.1, $7a^2$ is by far the most common error each year. The increase in the number of learners making this error in Grade 10 may suggest that learners' knowledge of algebraic simplification is unstable. However, the increase in Grade 10 may also suggest that learners expected a more complex question given that the first part of the test contained more difficult curriculum items.

Many learners treated the presence of brackets as a signal to multiply, not paying attention to the operations adjacent to the brackets. Consequently there is much evidence of the incorrect application of the distributive law in questions involving brackets. This is clearly illustrated in Q1.5 and Q1.6 as shown in Table 5. The percentage of learners giving these responses grows substantially from Grade 9 to Grade 10 and then drops off in Grade 11.

Of all the CSMS test items, Question 1 was most similar in form to SA curriculum items in algebra. Consequently, learners saw opportunity to use the algebraic laws and procedures they had been taught and this may account for the inappropriate use of such curriculum knowledge. Exponential laws are introduced informally in Grade 9, then formalised for integer exponents in Grade 10 and extended for rational exponents in Grade 11. The increased attention given to the laws in Grade 10 may explain the spike in this error. The drop in Grade 11 may reflect that this knowledge is now more stable. The trend in Q1.4 and Q1.5 is similar.

The distributive law is first introduced in Grade 8 and reinforced in Grade 9 for more complex algebraic expressions. It is therefore surprising that the misapplication of the law appears more frequently in Grade 10 and not in Grade 9 as was

the case with the exponential law. The reduction in errors involving the distributive law suggests that learners are more familiar with the correct application of the law in Grade 11.

In Q1.9 there is evidence of both the addition law of exponents and some attempt at the distributive law. It appears that many learners treated the question as the product of binomials: $(a + b)(a - b)$. The substantial increase in these errors from Grade 9 to Grade 10 is further evidence of curriculum effects since the factorising and expansion related to the difference of two squares is given more attention in Grade 10.

Errors related to evaluation of letters

When learners are not yet able to deal with a lack of closure in algebraic expressions, a common strategy is to evaluate the letters (Küchemann, 1981). Item Q5.3 was deliberately designed to test whether learners would accept the lack of closure inherent in the expression $8 + g$. As expected, many learners evaluated the letters with errors involving evaluating, accounting for 51% of all responses in Grade 9. This dropped to 31% in Grade 11. However, of all the errors committed in Grade 11, evaluating of letters constituted 52%. There were two common strategies for evaluating:

- *Equal splits*: splitting the known quantity equally depending on the number of letters on the left and then giving the new letter the same value as the others, for example $e = f = 4$ so g also has a value of 4 and hence $e + f + g = 12$.
- *Assigning a value of 1*: the new letter is given a value of 1, presumably because its value is unknown and hence can be any value, so the simplest value is chosen, giving $e + f + g = 8 + 1 = 9$.

Both strategies are reported in other studies that have made use of the CSMS items (MacGregor & Stacey, 1997; Oldenburg, Hodgen & Küchemann, 2013). We consider the equal splits strategy to be a more sophisticated strategy because it indicates that learners are looking for patterns in the relationship rather than merely adding one. Equal splits is the most common strategy constituting 29%, 31% and 33% of the errors each year respectively. It is also a persistent strategy across the three years. For example, of the 62 Grade 10 learners who used equal splits, 47% had also used it in

TABLE 4: Percentage of incorrect responses involving exponential laws.

Item	Typical responses	% of all responses			
		Grade 9	Grade 10	Grade 11	
1.1	Simplify $2a + 5a$	$7a^2$	10	17	7
1.4	Simplify $2a + 5b + a$	$7a^2b, 8a^2b, 2a^2 + 5b$	10	18	15
1.5	Simplify $(a - b) + b$	$a - b^2, ab^2$	8	17	13

TABLE 5: Percentage of incorrect responses involving distributive law.

Item	Typical responses	% responses			
		Grade 9	Grade 10	Grade 11	
1.5	Simplify $(a - b) + b$	$ab - b^2$	3	15	12
1.6	Simplify $3a - (b + a)$	$\pm 3a^2 \pm 3ab$	3	15	9
1.9	Simplify $(a + b) + (a - b)$	$a^2 \pm b^2$	10	26	21

Grade 9 and 37% continued to use it in Grade 11. Another noticeable trend was the move from the add 1 strategy to the equal splits strategy. For example, in Grade 9, 46 learners used add 1. Of these, 35% used the same strategy in Grade 10 but 27% changed to equal splits. In Grade 10, 46 used add 1. Of these, 24% moved to equal splits and 26% gave the correct answer in Grade 11. Other less common strategies included assigning consecutive whole numbers, which typically occurred when the constant was odd (Du Plessis, 2012), and solving for e .

In the interview extracts we gain insight into learners' reasoning about Q5.3 or similar items from the Honours projects. Thabo, a Grade 9 learner interviewed by Brown (2012, p. 95), argues confidently for an answer of 12 to 'if $e + f = 8$, then $e + f + g = \dots$ ':

Thabo: If you can take a look at e plus f equals to 8, it actually tells us that this e is 4 and this f is 4. That's why they got 8. And right here we'll say e plus f plus g equals to, ah, then one must say 4 plus 4 plus 4 is, 4 plus 4 plus 4, I think I'll, I will actually say 4 times 3 which is 12. That's why I got this 12.

Thabo says 'it actually tell us' indicating that he perceives the value of e and f to be given, and is not considering any other values for the letters. By contrast Lindo (Grade 11), in Madosi (2012, p. 25), is very clear that she *guessed* her answer to 'if $a + b = 6$, then $a + b + c = \dots$ ':

Madosi: So how did you get 3?
Lindo: I just guessed the 3. I just thought this 6, they are two variables, meaning I add two numbers to get 6.
Madosi: Ok.
Lindo: 3 plus 3, and here also I said 3 because there are 3 numbers, I also used the same number that I thought they used here.

In both cases learners 'know' that they need to obtain values for the letters and both make use of an equal splits strategy. Thabo assumes the values are implied in the question and recognises that he can multiply to obtain the sum. Lindo guesses a value of 3 but in another section of the interview acknowledges that she could have chosen different values.

Another strategy, which was more common among Grade 11 learners, involved treating the expression as an equation and solving for the unknown. This is illustrated in Du Plessis (2012), where Bokang solves for e in 'if $c + d = 6$ then $c + d + e = \dots$ ' (see Figure 1):

Bokang: I thought, I thought eh here neh, since c well, eh, c plus d is 6, I thought we were looking for e only. Then if we are looking for e only, then which means here was 3 and here was 3. Then you add this and this and you get 6. Then when you go to the other side then e is minus 6.

Bokang interprets the question to mean that he needs to find the value of e . He assigns equal splits to c and d and then treats the expression as an equation, assuming zero on the right side and then transposing the 6. His comment 'I thought

FIGURE 1: Bokang's response to 'if $c + d = 6$, then $c + d + e = \dots$ '

we were looking for e only' suggests he views this item as an equation to be solved and there is only one unknown. While he still evaluates the letter, his approach is more sophisticated than merely assigning equal splits or adding one.

Choosing strategies based on particular features of the item

There is evidence across several interview items that learners change their strategies depending on the specifics of the question. For example, Du Plessis (2012) shows that learners used different strategies in responding to a variation of Q5.3 depending on the choice of letters (consecutive versus non-consecutive) and constants (odd or even). Earlier we provided evidence of how Simon reasoned differently when subtracting a constant compared with subtracting a term involving a letter. In the example below Musa (Grade 11) reasons differently depending on the operation and the coefficients (Kalidheen, 2012, p. 48):

Musa: The first question says simplify where possible. The equation is $2a$ plus $3b$ plus c ($2a + 3b + c$). In this case you would have to add..., you would have to add all of them together but we have $2a$ plus $3b$ plus c so add the like terms together which is $2a$ plus $3b$ which will give us $5ab$ plus this c , ja, plus... ja, ja, so the final answer will be $5ab$ plus c .

...
Kalidheen: Which ones are the like terms?
Musa: $2a$ plus $3b$. They are like terms in, because a and b have terms in front of them and c does not.
Kalidheen: When you say term, what do you mean?
Musa: I mean the number.

When dealing with subtraction, Musa reasons differently. He is given $2b - a$ and says the answer is b (Kalidheen, 2012, p. 48):

Musa: $2b$ minus a , $2b$ minus, ok, $2b$ minus a , which it stands for 1 right? So if a stands for 1, so just, you would be

left with b . Yes, you would be left with b because, a because a in maths, if we use it as an x , x stands for actually one, so if we actually took out the a -factor which is in this case, er the one, so it means you must take out the one from the $2b$.

Musa's strategy for adding terms is to identify like terms although he does so based on an incorrect definition. When subtracting, he makes no mention of like terms. Instead he assigns 1 to a and then focuses on the numbers to 'take out one from the $2b$ '. Here again we see a view of subtraction as take away. The different strategies for addition and subtraction suggest that he does not see similarities between items that add algebraic terms and those that subtract them.

Discussion

The quantitative and qualitative analyses provide complementary insights into learners' errors in basic algebra. From the quantitative evidence we see that conjoining errors become less frequent by Grade 11 but they are still persistent. Another persistent error is the focus on numbers and the lack of attention to operations. Learners' talk confirms that they are separating numbers from letters when dealing with expressions.

Errors related to the learning of new procedures appear to depend on *when* the new procedures are taught. For example the inappropriate application of the distributive law increased in Grade 10 but dropped off by Grade 11. Although we cannot be certain, it appears that by Grade 11 learners are familiar with when the distributive law can be applied which suggests they may be paying more attention to the operations and not simply to the presence of brackets.

Learners' performance on items involving subtraction and negatives improved substantially from Grade 10 to Grade 11 but was still below performance on items without brackets and negatives. Further research needs to focus on learners' reasoning when dealing with subtraction and negatives both in number work and algebra. The interviews show evidence of learners talking of subtraction as 'take away', even when dealing with algebraic terms. This restricted view of subtraction may be a consequence of limited exposure to other subtraction structures such as reduction, comparison and inverse-of-addition (Haylock, 2006) in number work. Thus, despite evidence that arithmetic is an obstacle to learning algebra, there is need for placing greater attention on relationships between numeric quantities particularly in numeric subtraction scenarios.

There is widespread evidence in the tests and interviews of learners evaluating letters rather than accepting expressions such as $8 + g$ as a final answer. The most common evaluation strategy was equal splits and it was persistent up to Grade 11. This suggests that learners assume the letters must have values which may be a curriculum effect of the emphasis on equation solving. We have shown instances where learners are certain that the letter has a particular fixed value and other

instances where learners recognise that the letter could take on several different values. These are both steps towards the notion of letter as variable but learners need to develop this more sophisticated interpretation of letters earlier in order to cope with algebra and function.

The interview data provides evidence of learners incorrectly naming mathematical objects, for example referring to like terms as those with coefficients other than one, regardless of the letter. In other work emanating from the project we have argued for paying greater attention to the words learners use to name or refer to the mathematical symbols, expressions, graphs and for more opportunities for learners to talk mathematics (Adler & Ronda, 2015; Adler & Venkat, 2014).

One of the most alarming observations each year was the high percentage of errors that were coded 'other', meaning that there was insufficient trend to add a particular code for the error. Although the diversity of such errors reduced by Grade 11, the apparent idiosyncrasy of the errors remains a cause for concern. However, the interviews across all postgraduate projects confirmed that learners are frequently able to justify their strategies and answers with some rational connection to the mathematics they have learned. This highlights the need to gain deeper insight into learners' reasoning as they work with basic algebra and hence for teachers to provide more opportunity for learners to explain their strategies verbally.

Conclusions and implications

In this article we have taken a first step in identifying and describing the algebraic errors that learners make as they move from Grade 9 to Grade 11. We have also provided some evidence of their reasoning in relation to these errors. This enables us to see what learners *can* and *do* do as they progress to higher grades, thus going beyond the general calls from Grade 12 examiners for greater attention on the 'basics' of algebra. Since this research provides insights into learners' algebraic errors and their related thinking, it also provides a starting point to address the errors through attention to the underlying erroneous thinking rather than attempting to 'fix' the errors through reteaching.

Explanations for learners' errors depend on theoretical orientations but there is agreement across perspectives that learners need to move from operational to structural ways of thinking about the symbols and their relationships. This will require deliberate instruction with tasks that push learners towards a structural view. Furthermore, increased attention to learners' ways of speaking about algebra will provide greater insight into their thinking. There is no 'quick fix' to deal with learners' errors in algebra, particularly when one sees how some errors persist deep into senior secondary mathematics. However, the insights provided by the research reported here give some indication of useful starting points such as specific attention to the meaning of brackets and a stronger focus on negatives and different views of subtraction.

Acknowledgements

This work is based on the research of the Wits Maths Connect Secondary Project at the University of the Witwatersrand, supported by the FirstRand Foundation Mathematics Education Chairs Initiative and the Department of Science and Technology and administered by the National Research Foundation. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the institutions named above. The authors wish to thank Vasen Pillay for the vital role he played as data manager in the project.

Competing interests

The authors declare that they have no financial or personal relationship(s) that may have inappropriately influenced them in writing this article.

Authors' contributions

C.P. conducted data analysis and wrote the majority of the article; J.H. did Rasch analysis, wrote sections of the article relating to CSMS design and the validation test for SA; Y.S. re-transcribed interview data, checked accuracy of transcripts and conducted data analysis; J.A. was project leader and provided critical comments on drafts of the article.

References

- Adler, J., & Ronda, E. (2015). A framework for describing mathematics discourse in instruction and interpreting differences in teaching. *African Journal of Research in Mathematics, Science and Technology Education*, 19(3), 237–254. <http://dx.doi.org/10.1080/10288457.2015.1089677>
- Adler, J., & Venkat, H. (2014). Teachers' mathematical discourse in instruction: Focus on examples and explanations. In M. Rollnick, H. Venkat, J. Loughran, & M. Askew (Eds.), *Exploring content knowledge for teaching science and mathematics* (pp. 132–146). London: Routledge.
- Asquith, P., Stephens, A., Knuth, E., & Alibali, M. (2007). Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: Equal sign and variable. *Mathematical Thinking and Learning*, 9(3), 249–272. <http://dx.doi.org/10.1080/10986060701360910>
- Ben-Zeev, T. (1998). Rational errors and the mathematical mind. *Review of General Psychology*, 2(4), 366–383.
- Booth, L. (1984). *Algebra: Children's strategies and errors*. Windsor: NFER Nelson.
- Borasi, R. (1994). Capitalizing on errors as 'springboards for inquiry': A teaching experiment. *Journal for Research in Mathematics Education*, 25(2), 166–208.
- Brodie, K. (2014). Learning about learner errors in professional learning communities. *Educational Studies in Mathematics*, 85, 221–239.
- Brodie, K., & Berger, M. (2010, January). *Toward a discursive framework for learner errors in mathematics*. Paper presented at the 18th annual meeting of the Southern African Association for Research in Mathematics, Science and Technology Education, Durban.
- Brown, M. (2012). *Exploring algebra misconceptions in Grade 9 learners in a South African school*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Christou, K., & Vosniadou, S. (2012). What kinds of numbers do students assign to literal symbols? Aspects of the transition from arithmetic to algebra. *Mathematical Thinking and Learning*, 14(1), 1–27.
- Collis, K. (1975). *The development of formal reasoning*. Report of a SSRC sponsored project carried out at the University of Nottingham School of Education during 1974. Newcastle, New South Wales: University of Newcastle.
- Collis, K. (1978). Operational thinking in elementary mathematics. In J. Keats, K. Collis, & G. Halford (Eds.), *Cognitive development: Research based on a Neo-Piagetian approach* (pp. 221–248). Chichester: John Wiley & Sons.
- Department of Basic Education. (2012). *National senior certificate examination: National diagnostic report on learner performance 2012*. Pretoria: Department of Basic Education.
- Du Plessis, I. (2012). *Grade 11 learners' errors in introductory algebra*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Govender, R. (2012). *Exploring Grade 9 learner errors in introductory algebra*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Gumpo, L. (2011). *Learner thinking about introductory algebra and integers*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Hart, K. (Ed.). (1981). *Children's understanding of mathematics: 11–16*. London: John Murray.
- Hart, K., & Johnson, D. (Eds.). (1983). *Secondary school children's understanding of mathematics*. A report of the mathematics component of the Concepts in Secondary Mathematics and Science Programme. London: Centre for Science Education, Chelsea College.
- Haylock, D. (2006). *Mathematics explained for primary teachers*. London: Sage.
- Herholdt, R., & Sapire, I. (2014). An error analysis in the early grades mathematics – A learning opportunity? *South African Journal of Childhood Education*, 4(1), 42–60.
- Hodgen, J., Brown, M., Coe, R., & Küchemann, D. (2012, July). *Surveying lower secondary students' understandings of algebra and multiplicative reasoning: To what extent do particular errors and incorrect strategies indicate more sophisticated understandings?* Paper presented at the 12th International Congress on Mathematical Education, Seoul, Korea.
- Kalidheen, S. (2012). *Learner interpretations of letters and symbols in introductory algebra*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Küchemann, D. (1981). Algebra. In K. Hart (Ed.), *Children's understanding of mathematics: 11–16* (pp. 102–119). London: John Murray.
- Linacre, J. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 878.
- Linacre, J. (2015). *Winsteps® Rasch measurement computer program user's guide*. Beaverton, OR: Winsteps.com
- MacGregor, M., & Stacey, K. (1997). Students' understanding of algebraic notation: 11–15. *Educational Studies in Mathematics*, 33(1), 1–19.
- Madosi, T. (2012). *Typical errors that Grade 11 learners make when dealing with introductory algebra*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Makonye, J., & Luneta, K. (2014). Mathematical errors in differential calculus tasks in the Senior School Certificate Examinations in South Africa. *Education as Change*, 18(1), 119–136. <http://dx.doi.org/10.1080/16823206.2013.847014>
- Mashazi, S. (2012). *Learner thinking underlying the errors in introductory algebra at Grade 9 level*. Unpublished BSc (Hons) research report. University of the Witwatersrand, Johannesburg.
- Oldenburg, R., Hodgen, J., & Küchemann, D. (2013, February). *Syntactic and semantic items in algebra tests – A conceptual and empirical view*. Paper presented at the Eighth Congress of the European Society for Research in Mathematics Education, Antalya, Turkey.
- Olivier, A. (1989). Handling pupils' misconceptions. *Pythagoras*, 21, 10–19.
- Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education*, 10(3), 163–172.
- Ryan, J., & Williams, J. (2007). *Children's mathematics 4–15: Learning from errors and misconceptions*. Buckingham: Open University Press.
- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22, 1–36.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge, UK: Cambridge University Press.
- Shalem, Y., Sapire, I., & Sorto. (2014). Teachers' explanations of learners' errors in standardised mathematics assessments. *Pythagoras*, 35(1), 11 pages. <http://dx.doi.org/10.4102/pythagoras.v35i1.254>
- Spaull, N. (2013). South Africa's education crisis: The quality of education in South Africa 1994–2011. Johannesburg: Centre for Development and Enterprise.
- Wu, M., & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne, Australia: Educational Measurement Solutions.