Using a digital tool to improve students' algebraic expertise in the Netherlands: crises, feedback and fading

Christian Bokhove
University of Southampton

Enhancing ways of developing students' algebraic expertise remains an important focus for research. This paper reports on a design research study which involved a digital intervention for 17-18 year old students, implemented in nine schools in the Netherlands (N=324). For the intervention, algebra tasks for the conceptual and procedural components of algebraic expertise were placed in a sequence based on three design principles: (i) 'crisis' items that intentionally questioned the use of standard algorithms, (ii) feedback provided by the digital system, and (iii) the 'fading' of feedback during the sequence to increase transfer. Data collected included results from student pre- and post-tests, questionnaires, and scores and log files of their digital work. Results from the study show that the intervention was effective in improving algebraic expertise, and that the aforementioned design principles have merit. This paper reports on the effects and illustrates the design principles through a case example. The intervention shows a significant effect in improving algebraic expertise. It shows that well-thought-out design principles augment learning. The paper fits in a broader discussion on how to integrate algebraic expertise and ICT use in the classroom through the use of educational design.

Keywords: algebraic expertise, digital intervention, Netherlands, design principles.

Introduction

The distinction between procedural skills and conceptual understanding is a highly researched field of interest. In *Adding it up* (Kilpatrick, Swafford and Findell, 2001) synthesised the research in this area with the concept of *mathematical proficiency*. Mathematical proficiency comprises five strands: *conceptual understanding, procedural fluency, strategic competence, adaptive reasoning* and *productive disposition.* Conceptual understanding is defined as "the comprehension of mathematical concepts, operations, and relations" (p. 116), and procedural fluency as the "skill in carrying out procedures flexibly, accurately, efficiently, and appropriately" (ibid.). Furthermore, "the five strands are interwoven and interdependent in the development of proficiency in mathematics" (ibid.).

Both conceptual understanding and procedural fluency have been discussed extensively in research. For example, Arcavi (1994) introduced the notion of symbol sense, which includes "an intuitive feel for when to call on symbols in the process of solving a problem, and conversely, when to abandon a symbolic treatment for better tools" (p. 25). Arcavi (1994) exemplifies eight behaviours that describe symbol sense, and posits that these behaviours show the intertwinement between procedural skills and conceptual understanding, both being complementary aspects. In line with this work, Drijvers, Goddijn and Kindt (2010) define algebraic expertise as a dimension ranging from basic skills to symbol sense. Basic skills involve procedural work with a local focus and emphasis on algebraic calculation, while symbol sense involves strategic work with a global focus and emphasis on algebraic reasoning.

Acknowledging the potential of ICT for mathematics education (e.g. Heid and Blume, 2008a, 2008b; Pierce and Stacey, 2010), this study uses an ICT intervention for acquiring, practising and assessing aforementioned algebraic expertise. This paper reports on the study by first describing the design research approach and the algebra content of the intervention, then elaborating the three guiding principles behind the intervention, and its implementation in Dutch classrooms. Finally it presents the results, draws conclusions and presents some challenges for discussion. The focus of this paper is on the question whether three carefully chosen design principles for feedback have an effect on the acquisition of algebraic expertise. To further exemplify these principles, apart from quantitative data, a case example of one student is provided.

Methodology

The complete study 'Algebra with Insight' followed a design research approach with four phases, one preliminary phase and three intervention cycles (Van den Akker, Gravemeijer, McKenney and Nieveen, 2006). The preliminary research phase concerned the choice of the digital tool (Bokhove and Drijvers, 2010a). The first intervention cycle focused on whether it was possible to design a tool in such a way that it would allow symbol sense activities (Bokhove and Drijvers, 2010b). The second cycle consisted of a small-scale field experiment with two teachers in one school. The third and final cycle involved a large-scale classroom experiment. This research set-up shows a progress from small-scale to large-scale in 'layers of formative evaluation' (Tessmer, 1993). The last cycle, aimed at the intervention effects, is the focus of this paper, and in particular the effects of the three design principles for the intervention.

Algebra content of the intervention

In the design of the study intervention, we want to address algebraic expertise, both basic algebraic skills and symbol sense. To do so, the digital intervention needs to offer symbol sense opportunities (Bokhove and Drijvers, 2010b). Tasks were sourced from exit and entry examinations, textbooks, journals and remedial courses. Several suitable 'symbol sense type items' were identified and selected, with the main criterion being that items covered both basic skills and symbol sense. With this content, an intervention called 'Algebra met Inzicht' [Algebra with Insight] was designed in the Digital Mathematical Environment (http://www.fi.uu.nl/dwo/en) of the Freudenthal Institute.

The complete cycle consisted of a pen-and-paper pre-test (eight items), a digital practice module (sections d1-d4, 45 items, excluding randomisation), a digital diagnostic test (section d5, 23 items, excluding randomisation), a digital summative test (section d6, 23 items, excluding randomisation) and, finally, a pen-and-paper post-test (10 items). The time needed to complete the intervention was estimated at six hours, excluding pre- and post-tests.

The intervention was used in fifteen 12th grade classes from nine Dutch secondary schools (N=324), involving eleven mathematics teachers. The design set-up did not include control groups. The schools were spread across the country and showed a variation in school size, pedagogical and religious backgrounds. The participating students were pre-university level 'wiskunde B' students, of whom 43% were female and 57% were male. The participating schools subscribed after an open invitation in several bulletins for mathematics education. Schools received an example course plan and some suggestions on using the intervention, they were however free to adapt the intervention to their own requirements. Schools deployed the intervention in the last three months of 2010, just before preparations for the final national exams would start. Teachers received mailings on a regular basis, and could visit a project website with support materials like screencast instructions.

Data collection for the intervention included results from a pre- and post-test, and the scores and log files of the digital activities. The log files record information on students' individual item scores, feedback, answers, and number of attempts per step. Apart from marking for correct and incorrect, pre- and post-test were also marked with regard to symbol sense behaviour. Using a second marker, inter-rater reliability was good with an alpha=.91 for all items of 5% of the students' pre-tests and 5% of the post-tests.

Three main design principles

The three underlying principles for the intervention are answers to three challenges that arose from the three design cycles prior to the last cycle: (i) students learn a lot from what goes wrong, (ii) but students will not always overcome these difficulties if no feedback is provided, and (iii) that too much of a dependency on feedback needs to be avoided, as summative assessment typically does not provide feedback. These three challenges are addressed by principles for crises, feedback and fading, respectively.

The principle of using crises is based on, as the poet John Keats so eloquently described in the early 19th century, failure being 'the highway to success'. The same idea has had many forms during the years. Piaget (1964) used the concept of disequilibrium and equilibrium. Tall (1977) refers to 'cognitive conflicts'. Van Hiele (1985) distinguishes a 'crisis of thinking' with a need for challenge. More recently, Kapur (2010) coined the term 'productive failure'. Most sources, however, see crises as an inherent part of learning when solving problems; in this case the idea was to embed tasks that could intentionally cause a crisis (Bokhove and Drijvers, 2012b). In this intervention intentional crisis tasks are added to sequences of near-similar tasks, as depicted in Table I. The general structure of a sequence is: pre-crisis items, crisis item, post-crisis items. First, students are confronted with familiar equations (pre-crisis items).

1.1	Tasks: "Solve the following equation:"	**Pre-crisis items** In the initial items students are confronted with equations they have experience with. Students may choose their own strategy. Many students choose to expand brackets as that is the strategy that they have used often: work towards the form $ax^2 + bx + c = 0$ and use the Quadratic Formula. There is some limited feedback on the task.
1.2	$(4x-3)\cdot(4x-1)=(4x-3)\cdot 2$	
1.3	$\sqrt{3x+2}\cdot(3x+3)=\sqrt{3x+2}\cdot(6x-2)$	
1.4	$(x-4)\cdot(2x-5)=(x-4)\cdot(-3x+3)$	

1.5	**Opgave 1.5** Los de volgende vergelijking op:	$(5x-13)\cdot(4x-3)-(5x-13)\cdot(-2x+3)=0$
1.6	$(x^2+3x-3)\cdot(8x-6)=(x^2+3x-3)\cdot(4x+1$ $8x^3+18x^2-42x+18=4x^3+24x^2+24x-3$ $4x^3-6x^2-66x=-54$ $4x\left(x^2-1\frac{1}{2}x-16\frac{1}{2}\right)=-54$ ✓	**Crisis item** Students are then confronted with an intentional crisis: if a student uses his/her conventional strategy of expanding the expression. The yellow tick at the bottom of the screen denotes that the equation is algebraically equivalent to the initial one, but that it is not the final answer (half-correct). This is accompanied by a partial score for an item and some feedback in Dutch: 'You are rewriting correctly'. Although these students showed good rearranging skills, in the end they were not able to continue, as they did not master the skill to solve a third order equation. There is some limited feedback on the task.
1.7	**Opgave 1.7** Los de volgende vergelijking op:	tip stap losop $(2x^2-3x-2)\cdot(7x-3)=(2x^2-3x-2)\cdot(3x+12)$
1.8	$(x^2-3x-2)\cdot(6x-3)=(x^2-3x-2)\cdot(4x+12)$	**Post-crisis items** After the crisis item students are offered help by providing a 'voorbeeldfilm', an instructional screencast, and buttons to get hints ('tip'), the next step in the solution ('stap') or a worked solution ('losop'). These features have in common that they provide feed-forward information at the task level and self-regulation.
1.9	$\sqrt{3x+3}\cdot(2x+4)=\sqrt{3x+3}\cdot(6x-5)$	
1.10	$(4x+4)\cdot\sqrt{-2+3x}=\sqrt{3x-2}\cdot(7x-5)$	
1.11	$(-5+{}^{2}\log(x-2))\cdot(6x-6)=(-5+{}^{2}\log(x-2))\cdot(3x+14)$	
1.12	$(4x-13)\cdot(3x-3)=(4x-13)\cdot(-3x+2)$	
1.13	$(-4x+5)\cdot(8x-5)=(-4x+6)\cdot(3x+14)$	

Table 1 Sequence of items illustrating crises and feedback. The sequence starts with conventional pre-crisis items, then a crisis item that cannot be solved with the 'standard' procedure and ends with feedback to overcome the crisis and further practice items (post-crisis items).

Then the student encounters a carefully designed 'crisis item': this item intentionally confronts conventional strategies head on, meaning that the 'standard procedure' will not work. Finally, having experienced a 'crisis' students are offered help to overcome the crisis by providing feedback; the second design principle. As Hattie and Timperley (2007) pointed out, one effective action for learning would be to provide hints and corrective feedback. Feedback would then very much have the role of aiding assessment for learning, formative assessment. Black and Wiliam (1998) define assessment as being 'formative' only when feedback from learning activities is actually used to modify teaching to meet the learner's needs. Feedback in this intervention is provided at different levels: corrective through green, yellow and red symbols, and supportive by providing screencast movies, hints, next steps and worked solutions.

However, in my personal experience as a teacher I have seen there can be an over-reliance on feedback that is provided: when students take an exam there is no feedback present, so can students still solve tasks correctly, without any feedback? The third principle of fading addresses this. The digital intervention initially provides a lot of feedback, but the amount is decreased towards the end (Renkl, Atkinson and Große, 2004; Bokhove, 2008). Figure 1 shows how this principle was implemented in the intervention. At the beginning of the intervention, in sections d1 to d4, feedback is provided for all intermediate steps of a solution. The next part of the intervention, section d5, concerns self-assessment and diagnostics: the student performs the steps without any feedback and chooses when to check his or her solution by clicking a ‘check’ button. Feedback is then given for the whole of the exercise.

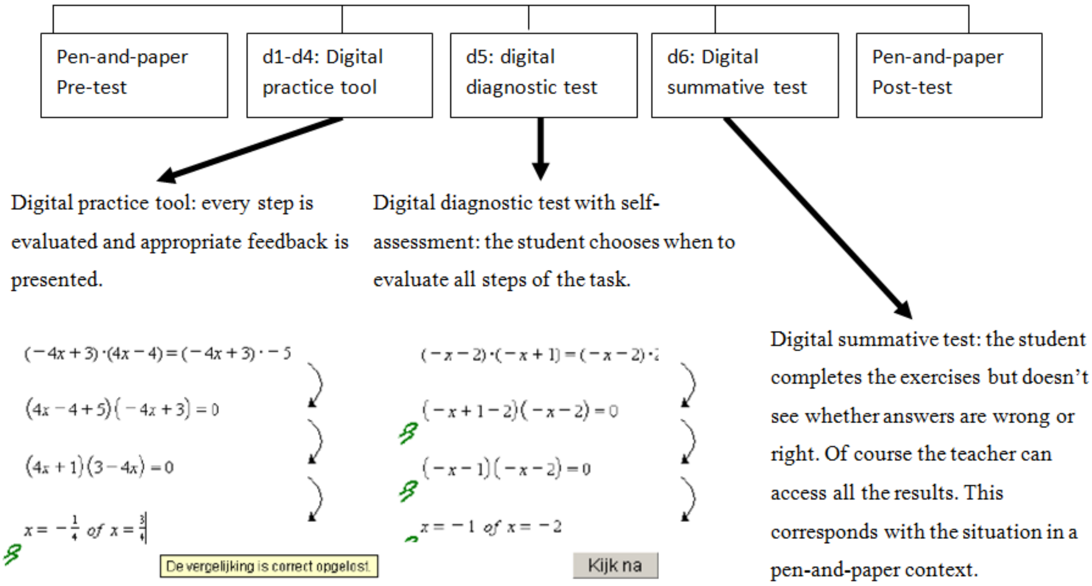

Figure 1 Outline of fading feedback in formative scenarios. The boxes at the bottom say ‘The equation has been solved correctly’ and ‘Check’, respectively.

In section d6, students get a final summative test with no means to see how they performed. Just as is the case with a pen-and-paper test, the teacher will check and grade the exam (in this case automatically) and give students feedback on their performance. In sum, the complete narrative behind the three design principles is that (i) *intentional crises* are provoked in students, (ii) enable students to overcome these crises by *providing feedback*, and (iii) to avoid a dependency on feedback *fade the feedback* in the course of the digital intervention.

Results: a case example

The following sequence of events during the intervention, concerns a student named Paula. She starts with a pre-test. Apart from the calculation error on the right hand side of the equation, Figure 2 shows that Paula’s strategy here is to expand the expressions, similar to students in earlier phases of the study (Bokhove and Drijvers, 2010b).

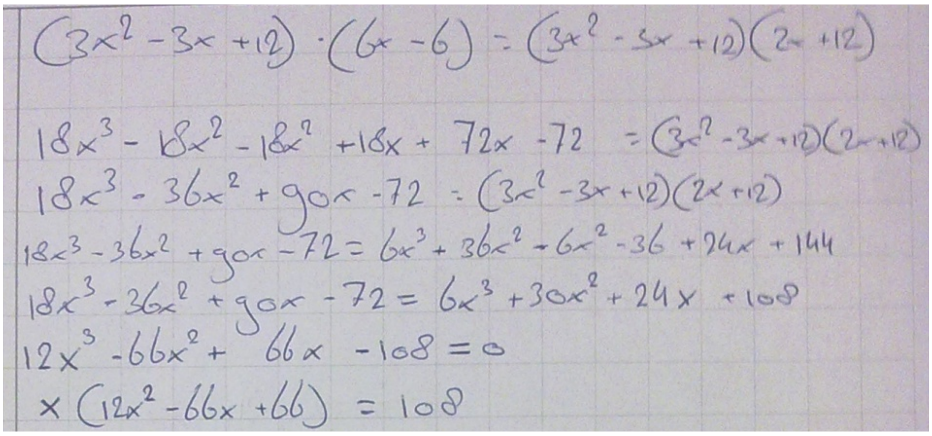

Figure 2 Example of Paula's pre-test pen-and-paper work.

Not surprisingly this strategy fails in the case of this equation. Paula scores low during the whole pre-test, only 14 out of 100. With regard to symbol sense, Paula scores poorly as well. Paula then starts with the sequence of digital tasks. In the first task she has to get acquainted with the digital environment. The pre-crisis items pose no problem for most students, including Paula. On arriving at the crisis item students exhibit three behaviours, roughly corresponding with the ones already observed in the pre-test: (i) students solve the equation correctly, (ii) students recognise the pattern of the equation but subsequently make mistakes (for example by losing solutions in the process), and (iii) students expand the expressions and get stuck with an equation of the third power. Figure 3 shows that Paula exhibits the third type of behaviour, quite similar to what she did in the pre-test.

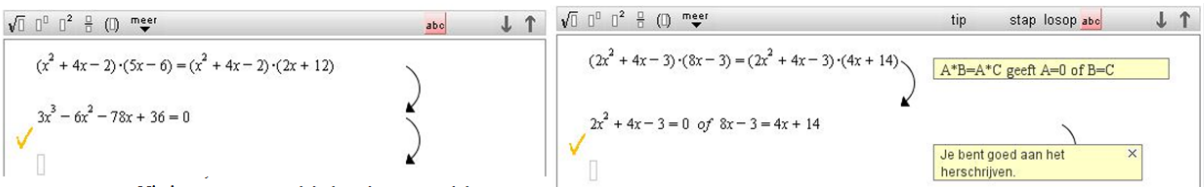

Figure 3 Paula's digital work. Left: crisis item. Right: post-crisis item. The boxes to the right say 'A*B=A*C yields A=0 or B=C' and 'You are rewriting correctly', respectively.

At this item feedback is still restricted to correct/incorrect. In addition, students are allowed to choose their own strategies, even when they aren't efficient or would lead to problems. In the post-crisis items, as well as feedback correct/incorrect, Paula is provided with buttons for hints and worked solutions, and the option to watch a screencast demonstrating the solution. The log-files from the online environment show that Paula fails at the crisis item (0 out of 10 points), but is successful at the post-crisis items with feedback (10 out of 10 points). Looking at the attempts made, intermediate steps for the equation that were sent to the system, Paula attempts the crisis-item 73 times, and the post-crisis items, aided by feedback, only three times. Finally, in the post-test Paula shows a significant increase in the total score (70 out of 100) and symbol sense behaviour. Even though mistakes are made they were not caused by a lack of symbol sense any more but errors in calculations. Focusing only on similar types of equations it becomes clear that Paula manages to solve these equations correctly. As the general results have shown, Paula is not a unique case in this school.

Overall results

Overall, dependent t-tests with pairwise exclusion if data was missing, show that students in participating schools improved on their scores and symbol sense behaviour. The score on the post-test (M=78.71, SE=15.175) is significantly higher than the pre-test score (M=51.55, SE=21.094), $t(286)=-22.589$, $p<.001$, $r=.801$, $d=-1.34$. For symbol sense behaviour scores on the post-test (M=1.462, SE=1.504) also is significantly higher than the pre-test score (M=-1.493, SE=2.339), $t(285)=-20.602$, $p<.001$, $r=.773$, $d=-1.22$. According to Cohen's benchmark (1992) this is a large effect. Both specific case examples and more quantitative analyses show that crises together with feedback decrease the number of step attempts needed, while fading feedback does not prevent a large effect (Bokhove and Drijvers, 2012b).

Conclusion and discussion

In this article I focused on the question whether three carefully chosen design principles for feedback have an effect on the acquisition of algebraic expertise. Overall, the use of the intervention for an average of six hours has a large effect on improving algebraic expertise. This effect did not only entail an improvement in score, but also an improvement in recognising patterns and having a sense for symbols. The question whether the three main design principles were the cause of this is much harder to answer with a 'yes' or 'no'. The principles seem to have merit: the *crises* together with *feedback* decrease the number of step attempts needed, and even with less and less feedback through *fading,* the effect on higher scores remains strong (see also Bokhove and Drijvers, 2012a; 2012b).

To conclude I want to address two points that led to discussions within and about this study. Firstly, I noticed that some educators were concerned that students were 'set up to fail'. A lot of this seemed to correspond with negative perceptions towards words like 'crisis' and 'failing'. This is understandable, as the words have a negative connotation in society. I would like to emphasise that students naturally are not told about an imminent crisis item. The whole intent is that students *could* fail if they exhibit mathematical behaviours we don't want. It would be quite unethical if this was not followed up by a solution as well: detailed feedback to overcome the crisis. In fact, I would contend that the whole combination of crises and feedback strengthens the learning, as set out in the section on design principles. A second point concerns the limitation of the study that the design set-up does not include a control group. As a more philosophical and final comment, I often wonder what the control group should be when one is introducing a new approach in the classroom. If I did not have the opportunity to provide feedback automatically it would be completely unfeasible to do the same thing as a teacher, without ICT. With a collaborative approach students could help each other, but I set out to look at the potential of ICT for acquiring algebra. Is it really so useful to have a control group in studies where the discerning factor, for example use of ICT, has inherent and obvious advantages?

References

Arcavi, A. (1994) Symbol Sense: Informal Sense-Making in Formal Mathematics. *For the Learning of Mathematics*, 14(3), 24–35.

Black, P. & Wiliam, D. (1998) Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kappa*, 80(2), 139–149.

Bokhove, C. (2008) Use of ICT in formative scenarios for algebraic skills. Paper presented at the *4th conference of the International Society for Design and Development in Education*, Egmond aan Zee, the Netherlands.
Bokhove, C. & Drijvers, P. (2010a) Digital Tools for Algebra Education: Criteria and Evaluation. *International Journal of Computers for Mathematical Learning*, 15(1), 45–62.
Bokhove, C. & Drijvers, P. (2010b) Symbol Sense Behaviour in Digital Activities. *For the Learning of Mathematics*, 30(3), 43–49.
Bokhove, C. &Drijvers, P. (2012a) Effects of a digital intervention on the development of algebraic expertise. *Computers & Education*, 58(1), 197–208. doi:10.1016/j.compedu.2011.08.010.
Bokhove, C. & Drijvers, P. (2012b) Effects of feedback in an online algebra intervention. *Technology, Knowledge and Learning*, 17(1-2), 43-59.
Cohen, J. (1992) A power primer. *Psychological Bulletin*, 112(1), 155–159.
Drijvers, P., Goddijn, A. & Kindt, M. (2010) Algebra education: exploring topics and themes. In Drijvers, P. (Ed.) *Secondary algebra education. Revisiting topics and themes and exploring the unknown* (pp. 5–26). Rotterdam: Sense.
Hattie, J. & Timperley, H. (2007) The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
Heid, M.K. & Blume, G.W. (2008a) *Research on Technology and the Teaching and Learning of Mathematics: Vol. 1, Research Syntheses*. Charlotte, NC: Information Age Publishing.
Heid, M.K. & Blume, G.W. (2008b) *Research on Technology and the Teaching and Learning of Mathematics: Vol. 2, Cases and Perspectives*. Charlotte, NC: Information Age Publishing.
Kapur, M. (2010) Productive failure in mathematical problem solving. *Instructional Science*, 38(6), 523-550.
Kilpatrick, J., Swafford, J.E., & Findell, B.E. (2001) *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academy Press.
Piaget, J. (1964) Development and learning. In Ripple, R.E. & Rockcastle, V.N. (Eds.), *Piaget Rediscovered* (pp. 7-20). New York: Cornell University Press.
Pierce, R. & Stacey, K. (2010) Mapping Pedagogical Opportunities Provided by Mathematics Analysis Software. *International Journal of Computers for Mathematical Learning*, 15(1), 1-20.
Renkl, A., Atkinson, R.K. & Große, C.S. (2004) How Fading Worked Solution Steps Works – A Cognitive Load Perspective. *Instructional Science*, *32*(1/2), 59–82.
Tall, D. (1977) Cognitive Conflict and the Learning of Mathematics. In Proceedings of the *First Conference of The International Group for the Psychology of Mathematics Education. Utrecht, Netherlands: PME*. Retrieved from http://www.warwick.ac.uk/staff/David.Tall/pdfs/dot1977a-cog-confl-pme.pdf
Tessmer, M. (1993) *Planning and Conducting Formative Evaluations*. Abingdon: Routledge.
Van den Akker, J., Gravemeijer, K, McKenney, S. & Nieveen, N. (Eds). (2006) *Educational design research*. London: Routledge.
Van Hiele, P.M.V. (1985) *Structure and Insight: A Theory of Mathematics Education*. Orlando, FL: Academic Press.