

Evaluating the effectiveness of an educational programming intervention on children's mathematics skills, spatial awareness and working memory

David Messer¹  · Lucy Thomas² ·
Andrew Holliman² · Natalia Kucirkova³

Received: 31 October 2017 / Accepted: 25 May 2018 / Published online: 2 July 2018
© The Author(s) 2018

Abstract This investigation concerns two questions: i). is simple educational programming with children, compared to working on mathematical tasks, more effective in increasing scores in mathematical abilities, spatial awareness and working memory? ii), is educational programming on a digital device, compared to similar paper and pencil programming activities, more effective in increasing mathematical abilities, spatial awareness and working memory? Forty-one 5 to 6 year olds from a UK infant school were randomly allocated to one of three groups: programming+iPad technology, programming using paper and pencils, and a comparison condition involving pencil and paper mathematical addition and subtraction tasks. Two 10-min intervention sessions took place each week, over a period of six weeks, with pre-intervention and post-intervention tests administered to assess children's mathematical abilities, spatial awareness and working memory. A series of mixed analyses of variance revealed that all three groups increased their mathematical abilities and spatial awareness. However, there were no significant increases in working memory, and there were no significant differences between any of the groups. These findings suggest that even a relatively short period of experience with programming can benefit other abilities and that the effects are similar to more directly targeted interventions. It was not found that programming using iPads resulted in higher scores than programming with paper and pencils.

✉ David Messer
david.messer@open.ac.uk

¹ Centre for Research in Education and Educational Technology, The Open University, Milton Keynes MK7 6AA, UK

² Psychology Department, University of Coventry, Coventry CV1 5FB, UK

³ UCL Institute of Education, University College London, 20 Bedford Way, London WC1H 0AL, UK

Keywords Programming intervention · Children education · Working memory · Spatial awareness · Mathematical ability · Spatial awareness · Bee-bot

1 Introduction

Educational research documenting the use of iPads in schools has been concerned with supporting children's story-making and writing (e.g., Rowe and Miller 2016), creativity and literacy (e.g., Dezuanni et al. 2015) or multiple outcomes for specific groups of children (e.g., children with autism, Allen et al. 2015). However, less attention has been paid to using iPad-supported programming to enhance children's other educational abilities. For previous generations of technology, it has been argued that the activity of programming could have general educational benefits beyond the acquisition of specific programming skills. In the 1980s, Seymour Papert created the Logo educational programming language to develop children's higher-level thinking (DiSessa 2000; Lye and Koh 2014; Papert 1980). Logo enabled children to solve problems in a playful environment, by giving commands to make a small mobile robot, called a 'turtle', move. Papert believed that using educational programming results in general cognitive gains as the experience advanced problem solving and other cognitive skills, and then these transfer to other domains, an argument that is becoming widely accepted (Scherer 2016).

Papert's ideas have provided the theoretical framework for our research. The instructions and context needed to make Papert's Turtle move provided the necessary environment for children to programme, verbalise and visualize mathematical concepts: further, children were encouraged to relate abstract calculations to real life and their agency was enabled by the computer tool, use of which was coordinated by the teacher. It was through the simple educational programming of the Turtle (e.g. instructions such as forward 4, right 3) and through peer conversations that children tapped into community knowledge. This model provides a useful theoretical orientation for our work because it focussed on the way simple programming involving mathematical calculations could develop general abilities in children, and the context ensured that children's learning was intrinsically motivated and yet, open enough to incorporate novel, periodically added, concepts. Inspired by Papert's theory, we have adopted an experimental approach to address the question of whether educational programming has a positive effect on children's abilities involving mathematics, spatial relationships and working memory. In some important respects, the programming task we used differed in implementation from the research of Papert which usually involved a physical turtle; our programming involved digital screen based activities on an iPad as these and similar devices are becoming available in primary schools. As with Logo the children gave simple programming instructions in this case instructions to a bee about the route needed to reach some honey.

Research involving previous generations of technology have documented positive effects of programming. In a meta-analysis of 65 studies with a range of programming languages used by kindergarten to college students, Liao and Bright (1991) reported that 89% of these studies resulted in positive effects of computer educational programming interventions compared to control groups. However, the reviewers pointed out that there was a lack of information about whether programming was as effective as other comparable interventions which involve similar non-computer based tasks.

Limited research interest has continued in the effects of programming when used by young children (see Fessakis et al. 2013). Scherer (2016; page 2) has argued, largely based on research conducted in the 1980s and 1990s, that evidence about the ‘transfer effects of computer programming skills on other skills such as problem solving and creativity is reasonable’, but goes on to state that ‘there is a strong need for empirical evidence supporting this, particularly in the context of the recent advancements of digital technologies’. This particularly applies to the UK primary school educational system, where research has been limited because programming was not until recently part of the UK National Curriculum. Thus, it is timely to revisit the issue of whether learning to programme has more general effect on children’s abilities in the early primary years (5–6 years). This topic of research also connects to other current technology based initiatives which seek to enhance education, by using the power of technology (Sarker et al. 2017).

Our investigation concerned the effect of educational programming on three different abilities, mathematics, spatial awareness and working memory. Mathematics was chosen partly because Papert assumed educational programming involving spatial relationships to have specific effects on mathematical abilities due to the related numeric content of the programming instructions. Indeed, it has been suggested that programming skills are part of what has been termed computational thinking (Lye and Koh 2014). In addition, mathematical abilities were chosen as an outcome measure as these have a central place in education and programming (e.g. use of algorithms) and programming appears to facilitate the learning of mathematics (Clements 2002; Computing at School 2013; Johnson 2000; Tynker 2016). Additionally, computer assisted learning may be an effective way to improve mathematical abilities (Cox et al. 2003; Cheung and Slavin 2013).

Because the educational programming task involved trying to work out the best route to reach a target location it was thought that there might be improvements in the children’s spatial awareness, which involves an understanding of an individual’s location in relation to other objects in the environment (Baddeley 1996). Spatial awareness has been found to be related to mathematics achievement (Raghubar et al. 2010), and related to programming abilities (Ambrosio et al. 2014).

The central executive of the working memory system was chosen as the third outcome variable. The central executive involves higher-order allocation of attentional resources in the presence of distraction or interference and is linked to academic success (Baddeley 1996). An example of a working memory task is hearing a spoken sequence of numbers and being able to repeat back the sequence in reverse order, a task that involves both keeping information in memory and being able to manipulate the information at the same time, a process that often occurs in mathematical calculations. Thus, if programming has a general effect on the capacity for thinking through developing the capacity to access and hold information, then there is the possibility of gains in working memory. A number of studies have demonstrated that interventions can improve working memory capacity in children (Buschkuhl et al. 2012). Consequently, we had tentative expectations that the experience of educational programming would increase working memory capacity.

The choice of comparison groups can be difficult in interventions where the effects of a learning process are being investigated (i.e. programming), and the intervention is presented using information technology (i.e. via iPads) which itself could have an effect

on outcomes through motivation or related processes (Ciampa 2013; Passey et al. 2004). Furthermore, it is of relevance that the technological determinist perspective supposes that technology by itself can enhance educational processes (Oliver 2010). In the light of these considerations it was decided to include a second intervention that involved the same Beebot programming operations, but which took place using paper and pencil tasks, to investigate whether programming on iPads was more effective in increasing mathematics, spatial awareness and working memory than similar paper and pencil tasks. In addition, there was a third group who did not carry out any programming, but were given addition and subtraction tasks to solve using paper and pencils (i.e. an active comparison group). This allowed us to examine whether the effects of programming were comparable or more effective than a teaching intervention which targeted mathematical abilities.

In sum, working memory, spatial awareness, and mathematics are important, inter-related abilities. Few studies have explored the impact of programming, and the way programming is taught, on these three abilities. Our research questions were:

1. Are increases in children's mathematical abilities, spatial awareness and working memory higher after educational programming than after working on mathematical tasks? We hypothesised that the two groups using programming would have higher scores after the intervention than the group who had to solve mathematical problems.
2. Is educational programming on iPads more effective than similar programming with paper and pencil tasks? We tentatively predicted that the motivating properties of iPads might result in higher post-test scores in the group who programmed with an iPad.

2 Methodology

2.1 Participants

Forty-one pupils aged between 5 and 6 years (Mean age = 73.0 months, $SD = 3.07$, at the start of testing) were recruited from an infant school in Warwickshire, UK. This was a convenience sample obtained from two classes. Participants were randomly assigned to one of the three intervention groups after pre-testing. Approval was obtained from the appropriate University Ethics Committee. Written consent was obtained from the parents and verbal consent from the children.

2.2 Measures

2.2.1 Mathematics

A school based measure of mathematics was developed. The test was carefully designed in close collaboration with the Year 1 teachers at the school using ideas from the British Ability Scales-3 (BAS-3) number skills subtest (Elliott and Smith 2011), and the Wide Range Achievement Test math computation subtest (Wilkinson and Robertson 2006). Twelve addition and subtraction questions were chosen which were thought to differentiate between the abilities present in the year group, and which were relevant to the

national curriculum. The most simple item was $1 + 2 = ?$, a medium difficulty items was $8 - ? = 3$, and the most difficult question was $? = 46 - 6$ (children had to fill in an empty box that was positioned at the location of the question marks). Internal reliability (Cronbach's α) at pre-intervention was .84, and at post-intervention was .88.

2.2.2 Spatial awareness

This ability was assessed using the standardised BAS-3 recognition of designs subtest (Elliott and Smith 2011). Children were shown an abstract line drawing for five seconds, it was removed and they had to identify what they had seen from four options. This process involved the children's analysis of shape, size, and orientation, together with visual-spatial memory. The maximum score was 24. Elliott and Smith (2011) report internal reliability (Cronbach's α) of .77.

2.2.3 Working memory

This ability was assessed using the BAS-3 recall of digits backwards subtest (Elliott and Smith 2011). Children listened to a sequence of spoken digits and attempt to repeat them back in reverse order. The maximum score was 25. Elliott and Smith (2011) report internal reliability (Cronbach's α) of .89.

2.3 Procedure

The spatial awareness and working memory standardised tests were administered in the school library during a single one-to-one session with the researcher. Teachers administered the maths questions during normal class time when participants worked independently and without teacher support. Assessments were administered in the following order during pre- and post-testing, recall of digits backwards and then spatial awareness. Scoring for all assessments was based on receiving one point for each correct response.

2.4 Intervention tasks

During the intervention, all three groups took part in two 10-min sessions per week, for six weeks. The children either worked on educational programming or mathematical tasks. The children worked individually, with three or four children seated together. All participants received instructions and support was available after there had been at least one attempt at a problem. The participant needed to complete the questions at each level of difficulty before attempting a more difficult question. The three interventions are described below.

2.4.1 Programming+technology group

The free Bee-bot iPad app (see Fig. 1 of the app's screenshot) is available from TTS Group Limited, 2012. The app involves giving instructions to a bee so that it can reach a flower. Only the following commands could be used: forwards, backwards, turn right by 90° , turn left by 90° . Commands were given by pressing



Fig. 1 Bee-bot user interface

touch-sensitive buttons. One command, or a sequence of several commands, could be given at any time, with reset buttons clearing previous commands, and a ‘Go’ button to execute the commands. Eighteen progressively difficult levels were included, with a timer and stars awarded for a quick, successful completion of a level. These activities, involving paths, maze solving and program debugging, were developmentally appropriate for the age of the children.

2.4.2 Paper and pencil programming group

A new paper-based task was designed to replicate the programming+technology task. Pictures of the eighteen Bee-bot levels were presented to children on paper. On the paper were lines where commands could be written and the tasks were very similar to those on the iPad. The commands could be drawn as icons and this minimised the influence of literacy abilities.

2.4.3 Mathematics group

The active comparison task involved simple mathematical calculations which did not involve technology or programming. Age-appropriate images (underwater, outer space etc.) were presented, with a range of mathematical questions about these images. For example, the children had to count the number of umbrellas, trees and clouds in a picture, they then had to add or subtract some of the numbers they had counted. The questions involved addition and subtraction.

3 Results

Table 1 shows the number of correct responses for mathematics ability, spatial awareness and working memory for all three groups at pre- and post-test. There were increases in all the scores from pre-test to post-test except for working memory in the mathematics intervention. Inspection of graphs of these data suggested that the scores were normally distributed apart from the programming group's working memory pre-tests, which showed some positive skew. Box's M values concerning homogeneity of covariance across the groups were found to be non-significant.

To investigate whether there were group differences (Programming+iPad, Programming, and Mathematics) and differences according to when the tests were administered (pre-test and post-test), mixed analyses of variance were conducted. The 3 (group) \times 2 (test times) analyses of variance were conducted separately on the scores for mathematical ability, spatial awareness and for working memory to give three sets of findings.

For mathematical ability and spatial awareness there was a similar pattern of findings with the only significant effect being the increase in scores from pre- to post-test, there were no significant group differences or interactions (for mathematics ability, there was a main effect of Time, $F(1, 38) = 8.57, p = .006$, partial $\eta^2 = .18$; no significant effect of Group, $F(2, 38) = .79, p = .463$, partial $\eta^2 = .04$, and no interaction, $F(2, 38) = .22, p = .804$, partial $\eta^2 = .01$. For spatial awareness, there was a significant effect of Time, $F(1, 38) = 26.54, p < .001$, partial $\eta^2 = .41$; no effect of Group, $F(2, 38) = .50, p = .610$, partial $\eta^2 = .03$, and no interaction, $F(2, 38) = .50, p = .610$, partial $\eta^2 = .03$). Lastly, for working memory, there was no significant effects indicating that none of the three types of intervention resulted in increases in working memory scores, nor were there group differences (Time, $F(1, 38) = 2.00, p = .166$, partial $\eta^2 = .05$; Group, $F(2, 38) = 1.85, p = .171$, partial $\eta^2 = .09$, and interaction effect $F(2, 38) = 1.79, p = .180$, partial $\eta^2 = .09$).

4 Discussion

To our knowledge, this study provides the first evaluation of the effectiveness of an educational programming intervention on improving mathematics, spatial awareness

Table 1 Correct responses by intervention group (programming+technology, programming, and mathematics) and time (pre-tests and post-tests) on the measures of working memory, spatial awareness, and mathematics ability

Group	Mathematics (/24)		Spatial awareness (/25)		Working memory (/12)	
	Pre-test Mean (SD)	Post-test Mean (SD)	Pre-test Mean (SD)	Post-test Mean (SD)	Pre-test Mean (SD)	Post-test Mean (SD)
Programming +Technology	5.36 (3.03)	6.14 (3.88)	9.36 (2.50)	11.93 (4.10)	7.79 (2.33)	8.57 (2.47)
Programming	6.07 (3.56)	7.43 (4.03)	9.36 (2.31)	12.50 (4.03)	8.57 (2.93)	9.43 (3.01)
Mathematics	4.54 (3.18)	5.85 (3.24)	8.92 (1.50)	10.85 (3.78)	7.46 (2.22)	7.08 (1.94)
Mean	5.34 (3.25)	6.49 (3.72)	9.22 (2.12)	11.78 (3.94)	7.95 (2.50)	8.39 (2.65)

and working memory in young children. All three groups had significant increases from pre- to post-test in mathematics and spatial awareness, but none of the groups had a significant pre- to post-tests increase in working memory. There was no evidence that the two programming intervention groups were significantly more effective in increasing scores on the three tests than the mathematics group, and there was no evidence that programming on an iPad was significantly more effective than a similar pencil and paper programming intervention.

The significant increase in mathematical abilities and spatial awareness in all three groups was unlikely to have been caused by general maturation over the comparatively short period of six weeks (Piaget 1971), and though it is possible that other school teaching could have increased these scores, we were not aware of any initiatives run in parallel to our study. Thus, the experiences of all three groups appears to have resulted in significant increases in mathematics and spatial awareness. This provides support for the prediction that carrying out and learning simple programming will increase mathematical abilities and spatial awareness. The lack of a significant increase in working memory is probably because such increases are more difficult to achieve because they involve changes to basic information processing and may require a longer intervention (Melby-Lervåg et al. 2016).

It was expected that the two programming groups would have better post-test performance than the mathematics group. There was no support for this prediction and this suggests that: i) the experience of addition and subtraction had a direct effect on the children's mathematical abilities which were comparable to the experience of programming, and ii) that because the children needed to look at the pictures reasonably carefully and note the presence of different types of objects this appears to have given similar experiences with the paper and pencil mathematical tasks to the two programming conditions involving a spatial task. Thus, both programming directly targeting mathematical abilities, resulted in gains in mathematics and spatial awareness. These are encouraging findings as they suggest that young children's experience of programming can have effects on other important abilities, which is consistent with the idea that including programming experiences in the primary curriculum may lead to gains not only in programming, but also in other related domains (Papert 1980; Scherer 2016). However, this interpretation needs to be supported by future research which would include a comparison group that carries out another type of activity (e.g. story construction) to rule out the possibility that some common experience was causing the increase in post-test scores of mathematics and spatial awareness.

In relation to our second prediction, contrary to expectations, we did not find that the programming with iPads was more effective than programming with paper and pencils. In other words, programming on an iPad did not confer any special advantage in terms of outcomes, this is consistent with arguments that the content of educational experiences is more important than the medium through which it is delivered (Clark 1994). However, it should be acknowledged that the absence of significant difference between paper and pen programming and iPad based programming could have been due to a lack of statistical power. It also is worth noting that in practice, it is likely that programming will be taught on digital devices, so that although educational programming might be equally effective when taught using paper and pencil tasks, it is unlikely to be employed in this way.

A limitation of the present study was the sample size. Although the number of participants was similar to previous studies which have detected effects (Henry et al.

2013; Holmes et al. 2009; Melby-Lervåg and Hulme 2012), the sample size may have made a type II error more likely. Another limitation of the present study was that maturation of general experience could have resulted in the improvements seen in all three groups.

To summarise, this investigation revealed that three types of learning experiences namely, simple programming on an iPad, simple programming with paper and pencil tasks, and practice with mathematical problems all resulted in significant pre- to post-intervention gains on assessments of mathematics and of spatial abilities. These are encouraging findings concerning the relevance of programming to other abilities. Further research is needed to support this interpretation, and in particular, an evaluation of whether similar gains occur when there is a comparison group that has a learning activity unrelated to the outcome assessments. In relation to our second research question, there was no support for the prediction that programming using iPads would be more effective than programming with paper and pencils, which is consistent with the argument that the content of instruction is more important than the medium through which it is delivered.

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.

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

References

- Allen, M. L., Hartley, C., & Cain, K. (2015). Do iPads promote symbolic understanding and word learning in children with autism? *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00138>.
- Ambrosio, A., Almeida, L., Macedo, J., & Franco A. (2014). *Exploring core cognitive skills of Computational Thinking*. University of Sussex http://users.sussex.ac.uk/~bend/ppig2014/3ppig2014_submission_13.pdf. Accessed 6 June 2016.
- Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, 49, 5–28.
- Buschkuhl, M., Jaeggi, S., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience*, 25, 167–179.
- Cheung, A., & Slavin, R. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: a meta-analysis. *Educational Research Review*, 9, 88–113.
- Ciampa, K. (2013). Learning in a mobile age: an investigation of student motivation. *Journal of Computer Assisted Learning*, 30, 82–96.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29.
- Clements, D. (2002). Computers and early childhood mathematics. *Contemporary Issues in Early Childhood*, 2, 160–181.
- Computing at School. (2013). 160–181. *Computing in the National Curriculum: A guide for primary teachers*. Swindon: Computing at School, part of BCS.
- Cox, M., Abbott, C., Webb, M., Blakeley, B., Beauchamp, T., & Rhodes, V. (2003). ICT and attainment: A review of the research literature. In *ICT in Schools Research and Evaluation Series, No 17*. Coventry: Becta/DfES.
- Dezuanni, M., Dooley, K., Gattenhof, S., & Knight, L. (2015). *iPads in the early years: Developing literacy and creativity*. New York: Routledge.
- DiSessa, A. (2000). *Changing minds: Computers, learning, and literacy*. Boston: MIT Press.
- Elliott, C., & Smith, P. (2011). *British ability scales III*. London: GL Assessment.

- Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5-6 years old kindergarten children in a computer programming environment: a case study. *Computers & Education*, *63*, 87–97.
- Henry, L. A., Messer, D. J., & Nash, G. (2013). Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. *Infant and Child Development*, *23*(1), 84–103.
- Holmes, J., Gathercole, S., & Dunning, D. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*, 9–15.
- Johnson, D. (2000). Algorithmics and programming in the school mathematics curriculum: support is waning - is there still a case to be made? *Education and Information Technologies*, *5*, 201–204. <https://doi.org/10.1023/A:1009658802970>.
- Liao, Y., & Bright, G. (1991). Effects of computer programming on cognitive outcomes: a meta-analysis. *Journal of Educational Computing Research*, *7*, 251–266.
- Lye, S., & Koh, J. (2014). Review on teaching and learning of computational thinking through programming: what is next for K-12? *Computers in Human Behaviour*, *41*, 51–61.
- Melby-Lervåg, M., & Hulme, C. (2012). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291.
- Melby-Lervåg, M., Reddick, T., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of “far transfer”: evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*, 512–534.
- Oliver, M. (2010). Technological determinism in educational research: some alternative ways of thinking about the relationship between learning and technology. *Journal of Computer Assisted Learning*, *27*, 373–384.
- Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. Cambridge: Basic Books.
- Passey, D., Rogers, C., Machell, J., & McHugh, G. (2004). *The motivational effect of ICT on pupils*. England: DfES/University of Lancaster.
- Piaget, J. (1971). *Genetic epistemology*. New York: W.W. Norton.
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: a review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122.
- Rowe, D. W., & Miller, M. E. (2016). Designing for diverse classrooms: using iPads and digital cameras to compose eBooks with emergent bilingual/biliterate four-year-olds. *Journal of Early Childhood Literacy*, *16*(4), 425–472.
- Sarker, S., Mohapatra, S., & Sundarakrishan, J. (2017). Assessing impact of technology based digital equalizer programme on improving student learning outcomes. *Education and Information Technologies*, *22*, 195–213.
- Scherer, R. (2016). Learning from the past—the need for empirical evidence on the transfer effects of computer programming skills. *Frontiers in Psychology*, *7*, 1390. <https://doi.org/10.3389/fpsyg.2016.01390>.
- Tynker (2016). *Programming = Better Math Skills + Fun*. <http://www.tynker.com/blog/articles/ideas-and-tips/programming-better-math-skills-fun>. Accessed June 22 2016.
- Wilkinson, G., & Robertson, G. (2006). *Wide range achievement test* (Fourth ed.). London: Pearson Education Limited.