Abstract

Background: Although previous correlational studies have shown that both symbolic and non-symbolic abilities relate to mathematical abilities, correlational studies cannot show the cause and effect of these abilities for mathematical success. Aims: The current study examined the effect of a non-symbolic training programme, called PLUS and a symbolic training programme, called DIGIT, to provide further insight into the causal nature of domain specific factors that contribute to mathematical abilities. Methods and Procedures: Forty-nine preschool children who had limited symbolic and non-symbolic knowledge were recruited and randomly allocated them to the DIGIT and PLUS training programmes. Performance on a number of mathematical tasks was compared to 20 preschoolers with no mathematical difficulties. Outcomes and Results: Performance in both training programs improved on the Test of Early Mathematical Abilities as well as on a non-symbolic Approximate Number Sense task, counting tasks, and digit recognition tasks, immediately after five weeks of training and this improvement remained six months later. Conclusions and Implications: This study provides further evidence that symbolic and non-symbolic abilities bi-directionally impact on each other and that ordinality knowledge is an important factor of mathematical development.

What this paper adds:

Previous correlational studies have debated the importance of non-symbolic and symbolic abilities for the development of general mathematical abilities in young children. The current study is the first to examine the impact of a training programme that focuses on non-symbolic abilities and directly compares it to a training programme that targets symbolic knowledge in a group of preschool children with low mathematical abilities. The results

provide further evidence about the foundations for mathematical achievement in preschoolers. They suggest that both symbolic and non-symbolic abilities bi-directionally impact on each other and that this bi-directional relationship might be driven by ordinality knowledge.

Keywords: symbolic knowledge, Approximate Number Sense, intervention, preschoolers, low-achievers.

1. Introduction

Strong mathematical abilities are required for everyday living and influence employability, wages, and wellbeing (Rivera-Batiz, 1992). In addition, children who engage actively in fun mathematical games early on in life are more likely to engage with mathematics later on in life (van de Walle & Lovin, 2006). However, between 15 to 35% of preschoolers are typically classified as low achievers (LA) for mathematical abilities and are at an increased risk of being diagnosed with mathematical learning difficulties later on in life (Geary, Hoard, Nugent, & Bailey, 2013).

It has been shown that good mathematical abilities rely on both symbolic abilities, such as counting and digit knowledge, and non-symbolic abilities, which include magnitude knowledge (see Schneider et al., 2017 for a review). However, most of these studies have mainly relied on correlational methods which cannot provide any insight about the underlying cognitive mechanisms that drive these relations. Instead, these findings should be followed-up by intervention studies. In addition, it is not clear what type of intervention works best for children who perform low on mathematical achievement tasks. The current study compares the effectiveness of two different training programs in LA preschool children who have received little formal education and have limited symbolic knowledge: one training program concentrated on symbolic knowledge whilst the other focused non-symbolic knowledge only. This allowed us to examine the impact of symbolic and non-symbolic abilities on mathematical abilities in LA preschoolers and to make suggestions for educational practice.

1.1. Symbolic versus non-symbolic mathematical abilities

Mathematical abilities rely on a number of domain specific skills including symbolic abilities, such as counting, cardinality, and digit recognition, as well as non-symbolic abilities, such as the approximate number system (ANS). The ANS is a noisy, imprecise, nonverbal system that allows discrimination of large numerosities without counting or numerical symbols (Feigenson, Dehaene, & Spelke, 2004). This system relies upon the ratios presented, known as Weber's fraction (*w*), and over development children become better at discriminating between smaller ratios. For example, at 6 months of age infants can, using their ANS, discriminate between 8 and 16 dots (ratio 1:2) presented but not 8 versus 12 (ratio 3:4) (Xu & Spelke, 2000). However, by 9 months of age, infants can discriminate between 8 and 12 sounds but not between 8 and 10 sounds (ratio 4:5), providing evidence that numerosity discrimination increases in precision over development (Lipton & Spelke, 2003). Studies in both typical (see Chen & Li, 2014; Fazio, Bailey, Thompson & Siegler, 2014 for a review) and atypical development (see Van Herwegen & Karmiloff-Smith, 2015 for a discussion) have shown that ANS is an important foundation of children's number abilities. ANS precision measured in infancy and preschoolers predicts mathematical abilities later on in life (Starr, Libertus, & Brannon, 2013). The ANS mapping account (Feigenson et al., 2004), argues that the ANS is a critical building block for mathematics upon which symbol knowledge can be mapped upon and that the ANS functions as a checking system for exact symbolic arithmetic. Thus, improved precision of the ANS representations, i.e. the ability to discriminate between smaller ratios, allow for more accurate processing and checking when completing symbolic tasks, such as when adding two numbers.

However, not all studies have found that non-symbolic ANS abilities predict mathematical abilities (see De Smedt, Noël, Gilmore, & Ansari, 2013; Schneider et al., 2017 for a review) and it has been argued that performance on symbolic magnitude tasks, in which children have to say which symbolic digit is larger, rather than non-symbolic tasks, in which children have to say which set contains the largest amount, correlates with mathematical achievement (Lyons, Price, Vaessen, Blomert, & Ansari, 2014). Therefore, the dual representation view (Carey, 2009) argues that symbolic and non-symbolic abilities develop completely independently and that small numerical symbols are initially mapped onto precise representations based on symbol-symbol associations, which through increasing knowledge of the counting list can then be applied to larger numbers (see Sasanguie, De Smedt, & Reynvoet, 2017 for a discussion).

Others have suggested that the relationships between symbolic and non-symbolic abilities change over time and that whilst non-symbolic ANS abilities early on in mathematical development, symbolic abilities become an important prerequisite for higher-order arithmetical achievement as children progress through the school years (Xenidou-Dervou, Molenaar, Ansari, van der Schoot, & van Lieshout, 2017).

In addition to the views presented above, it is possible that better symbolic abilities improve non-symbolic abilities and vice versa (for a review see Chen & Li, 2014; Mussolin, Nys, Content, Leybaert, 2014; Schneider et al., 2017). We will refer to this as the interaction view. Seeing the potential bi-directional relationship, correlational studies alone cannot provide full insight into the importance of ANS abilities or symbolic abilities for further general mathematical abilities. Intervention studies, on the other hand, that specifically target specific mathematical abilities allow for examination of the effect of improved ANS abilities or improved symbolic abilities on general mathematical performance.

1.2. Improving maths abilities in typically developing preschoolers

A number of studies in both children and adults have shown that training on approximate, non-symbolic magnitudes only improved performance on symbolic mathematical tasks (Hyde, Khanum, & Spelke, 2014; Kuhn & Holling, 2014; Park & Brannon, 2013, 2014). However, the relationship between ANS and symbolic knowledge is likely to be affected by the amount of symbolic knowledge a child already has and thus, the amount of schooling can influence ANS acuity (Mussolin et al., 2014). Therefore, it is important to examine intervention outcomes in very young children who have received very little or no formal maths instruction yet.

Recently, a number of studies have focused specifically on improving non-symbolic ANS abilities in preschoolers and these studies have found that improving preschoolers' non-symbolic ANS acuity had a positive impact on children's mathematical abilities (Odic, Hock, & Halberda, 2014; Wang, Odic, Halberda, & Feigenson, 2016). A recent study by Honoré and Noël (2016) compared a symbolic ANS computerised training program that used Arabic numerals to an identical non-symbolic ANS training program that included collections of dots with 5-year-old children. Although non-symbolic ANS comparison abilities improved in both training groups, only the symbolic ANS training group showed improved arithmetic abilities. These results suggest that symbolic abilities can regulate non-symbolic abilities but improved non-symbolic abilities do not necessarily impact on symbolic abilities, which would support a dual-representation view. However, in many studies the task used to train participants was the same task as the assessment task. Although children do not expect symbolic abilities to improve after ANS training tasks (Dillon, Pires, Hyde, & Spelke, 2015), presenting pre-and post-assessment tasks with the same structure as the training tasks can result in a placebo effect (Boot, Simons, Stothart, & Stutts, 2013). Therefore, it is unclear whether any training effects found are due to the method of assessment or to the true effect of training.

The Preschool Number Learning Scheme (PLUS) is the only intervention program to date that targets non-symbolic magnitudes using stimuli from a range of modalities, including visual, auditory, tactile and vestibular stimuli, in preschool children. The PLUS games are based on games familiar to preschoolers, such as playing musical instruments and playing with domino's, and include two types of games that match the ANS functions, namely estimating and matching large numerosities. A recent study showed that playing PLUS games for five weeks improved preschooler's ANS abilities, in contrast to an active control group that read books (Van Herwegen, Costa & Passolunghi, 2017). However, the study did not include a formal mathematical ability task and thus, it is not clear whether improving preschoolers' ANS abilities would result in improved general mathematical abilities. In addition, it is not clear how the PLUS games compare to games that target symbolic abilities in order to improve general mathematical abilities in preschoolers.

1.3. The current study

The current study randomly assigned LA preschoolers who had low symbolic knowledge and performed below chance on an ANS task to two different types of training programs; the PLUS program (Van Herwegen et al., 2017) focussed on non-symbolic abilities and the DIGIT program (developed specifically for this study) targets two aspects of symbolic knowledge, namely counting and digit recognition.

It was predicted that if, in line with the ANS mapping account, that ANS abilities are important for the development of mathematical abilities then children who played PLUS-games would show improved ANS abilities as well as improved symbolic knowledge. If, however, symbolic knowledge is more important for mathematical abilities, then children who played DIGIT games should show higher improvements after the training program than the PLUS group. In addition, in line with the dual representation view, improving symbolic abilities would not necessarily improve non-symbolic abilities, as according to this view the numerical meaning of symbolic numbers is not acquired through a mapping process onto the ANS (Reynvoet & Sasanguie, 2016). Therefore, it would be predicted that improved symbolic abilities would not improve ANS abilities. However, in line with the interaction view, it would also be possible that improving the counting sequence would improve children's understanding of the numerical system which in turn may improve ANS acuity (Mussolin et al., 2014). In this case, it was predicted that improved symbolic abilities would also improve ANS performance. In addition to comparing performance scores before (pre-assessment) and after the training sessions (post-assessment), participants completed all assessments again six months later to examine any long-term effects of the training programs (follow-up assessment). Finally, performance of LA children was compared to a group of typically performing (TP) controls in order to assess whether either of the training programmes would allow the LA children to catch up with their peers.

2. Methods

2.1. Participants

Forty-nine children between the ages of three and five years old out of 539 children from fourteen preschool settings (seven private and seven free public settings) from Greater London were identified as LA. These children had general intellectual abilities but performed at or below the 45th percentile on the Test of Early Mathematical Abilities (TEMA-3; Ginsburg & Baroody, 2003)[1] and scored lower than 30 out of 48 trials correct on the ANS task[2]. These children did not have any developmental issues reported by parents in a parental questionnaire. However, only 38 children (19 in PLUS and 19 in DIGIT group) were included in the final analyses due to high attrition at follow-up (see Figure 1 for full details). These children were compared to a TP group of twenty children who scored within the typical range for mathematical abilities. A group of 20 TP children who did attend the preschool settings as usual and did not participate in any training programs was selected from

[1] The TEMA-3 scores from all the children screened showed that TEMA-3 is insensitive at the lower range and that even a high percentile score is based on getting just a few items correct. Based upon the fact that 45% of the total sample of children in the current study scored a raw score of 6 or less on the TEMA-3, where the score range is 0-32, we opted for a higher percentile cut-off for TEMA-3 to define LA children.

[2] A score of 30 equals a score significantly higher than chance level. However, children below this score are not necessarily giving random answers as they all passed the training task, suggesting they understood the task, and they all showed better performance on the larger ratios compared to the smaller ratios pre-intervention; $F(3,165)= 5.301$, $p = .002$, $\eta^2_p = .088$.

those children who performed at or above the 50th centile on the TEMA-3 and had scores above 30 out of 48 trails correct for the ANS task.

Figure 1 about here

All children came from a variety of Socio-Economic Status (SES) backgrounds. SES was measured using mother's highest level of education as parental education is considered to be one of the most stable aspects of SES (Sirin, 2005).

Tables 1 and 2 provide an overview of the participants' screening and demographic information. As can be seen from the tables, the groups included a similar ratio of boys versus girls and there was no significant difference for maternal education between the three groups; $\chi^2(18)= 19.720$, $p = .349$.

Table 1 about here

Table 2 about here

2.2. Materials

2.2.1. Screening for mathematical difficulties

British Ability Scales (BAS3) was used to measure children's general intelligence**.** Six core scales (Verbal Comprehension, Picture Similarities, Naming Vocabulary, Pattern Construction, Matrices, and Copying) from the BAS*3* Early Years cognitive battery were used to derive a General Cognitive Ability (GCA) score (*BAS3*; Eliot & Smith, 2011). This summary score has a mean of 100 and a standard deviation of 15. The technical manual reports an average test-retest reliability coefficient for the composite GCA scale of .93 (range= .91-.94) for the early year age range.

Test of Early Mathematical Abilities-3 (TEMA-3) assessed children's mathematical difficulties. TEMA-3 is a norm-referenced diagnostic instrument to determine mathematics strengths and difficulties for children aged 3 to 8 years old. Administration takes about 40 minutes and includes symbolic and non-symbolic items. The test includes A and B forms that can be used interchangeably to evaluate training programs. Internal consistency reliabilities are all above .92; immediate and delayed alternative form reliabilities are in the .80s and .90s. Percentile scores were a measure of interest to include children in the training programs.

ANS Task. In this computer task children were presented with a set of dot presentations on the left and right of the screen. The presentations included between 5 and 28 dots with ratios 0.5, 0.6, 0.7 or 0.8. The task included 48 trials, including both congruent trials (visual properties correlated with the number of dots) and incongruent trails (visual properties did not correlate with the number of dots). The presentation with 'more' dots was counterbalanced and appeared on either the left or right side of the screen (See Gebuis and Reynvoet, 2011 for similar task). Children were asked to select the dot presentation that had 'more' using a touch screen. Participants received a score of 1 for each correct trial and the maximum score was 48.

Prior to the actual ANS task, participants were administered a practice task in which it was assessed whether children understood the concept of 'more' in a numerical sense. Each training trial showed two dot presentations that had a ratio difference of 1/3 between them and included congruent and incongruent trials. Children received feedback when they picked the incorrect answer (see Negen & Sarnecka, 2015 for a similar approach). The practice task required children to complete 8 consecutive trails correct and none of the children required more than 12 trials to reach this criterion. This shows that when the trials were easy all children understood the task.

2.2.2. Pre-, Post- and Follow-up Assessments

Mathematical Achievement. In addition to the screening at pre-assessment, the TEMA-3 also was administered at post-assessment and at follow-up assessment in order to evaluate whether the training programs improved the participants' overall mathematical knowledge. As not all participants completed the exact same set of items, ability scores were also compared between the three groups across the different time points.

ANS Abilities. The ANS –task was also administered at post-assessment and follow-up assessment, in addition to the screening assessment, to examine any training effect on ANS abilities. Raw scores were used to examine any change in abilities.

Counting Abilities. To assess counting abilities a *Counting Task* and an *Enumeration Task* were carried out. In the Counting Task children counted out loud from 1 and the highest number correctly counted was recorded. In the *Enumeration Task* children counted a line with 20 equally spaced dots whilst pointing at each of the dots counted. The highest number counted correctly was recorded.

Cardinal Principle. *Give A Number* task (Wynn, 1992) was used to assess children's understanding of counting. Over fifteen randomized trials, children were asked to give the experimenter exactly 1, 2, 3, 4, and 5 beads from a clear bag with different sized beads. The children were then classified depending on the highest number they could correctly give on at least two out of the three trials, i.e. 1-knower, 2-knower, 3-knower, 4-knower, or 5-knower. Cronbach's alpha for reliability was .865.

Digit knowledge. Children were shown flashcards with digits 1 to 20 in random order and asked to name the digit. The total number of digits correctly named was recorded.

Letter Knowledge. In this control task children were shown all 26 lower-case letters of the alphabet in randomised order and asked to either name the letter or sound. As neither training programs included any letters, children should not improve on letter knowledge abilities. The total number of correctly named letters was reported.

2.3. Procedure

LA children were randomly allocated, using an online random number generator, to either a non-symbolic training program, called PLUS, or a symbolic one, called DIGIT. Each training program took five weeks during which time we aimed to administer 20 sessions of approximately 10 minutes with each child. However, as preschool children in the UK are not obliged to attend preschool, the number of sessions for each child varied (see Table 1).

2.4. Intervention programs.

The PLUS and DIGIT games are based on games familiar to preschoolers and include stimuli that relate to a variety of senses, including touch, sounds, and visual stimuli. The two training programmes both focused on speed and winning and were made as similar as possible, with exception that PLUS games target non-symbolic abilities and DIGIT games target symbolic ones. The difficulty of each game was adapted to the child's learning, i.e. if a child could not yet discriminate between an easy ratio, no harder ratios were used. Similarly, if a child could only count until certain number or recognise a certain symbol, no counting names or symbols beyond this point in the number sequence were introduced. This ensured that all children were continuously practising the games within the zone of their proximal development. Therefore, the variability in performance between the children present at the start of the intervention was also present at the end of the intervention, as not all children were expected to improve to a similar extent. For all games, corrective feedback was provided by the trainer, i.e. "that is right, there are a lot of dots here and a lot of dots there so those two go together". The fact that children played a different PLUS or DIGIT game each day and that the difficulty of the games was adapted over time, prevented children from becoming bored with the games.

PLUS games. These include two types of games, each matching a function of the ANS system, namely estimation and matching ratios (see Table 3). There are four estimation games that aim to improve children's ability to guess, and four games in which children match or differentiate between different numerosities of various ratios. Variables such as contour length, density, and colour are controlled for through the use of different shapes and the sizes of the objects used. Both congruent trials and incongruent trials were included in the games. The number of dots and objects in the games varied from 5 to 20.

For five weeks, the child played one estimation and one matching game during each session. Counting was prevented by promoting short response times and showing the stimuli repeatedly but only for a short time. All children started the games using easy large ratios and the trainers made these ratios smaller once the child had mastered the easier ones.

Digit games. Children in this training program played a number of games that target counting and digit knowledge (See Table 3) for the numbers 1 to 20. Each session, the child played one counting and one digit matching game. Again, once a child had mastered a particular digit or counting sequence new digits or counting names were introduced.

Table 3 about here

The trainers were two experienced researchers who had received extensive training by the first author before the start of the intervention programs. Both trainers followed a detailed manual for each of the games, ensuring treatment fidelity.

All training sessions as well as the pre-, post-, and follow-up assessments took place in a quiet corner in the preschool setting. Pre-, post-and follow-up assessments were carried out by different researchers who were blind to the training condition or group the child was assigned to.

Parents were provided with detailed information about the project and provided written consent, whilst verbal assent was obtained from all children. This project had received favourable opinion from the xxxx Ethics Committee.

2.5. Data analysis plan

As the TP children were expected to have higher mathematical abilities than the LA children prior to the training programs, repeated measures ANOVA scores[3] were used (Huck & McLean, 1975) to examine the main effects of the intervention programs on children's assessment scores. Boxplots were used to identify any outliers and these data points were removed before the analyses. Greenhouse-Geisser correction was used when the assumption of sphericity had been violated. Pairwise comparisons were used to analyse any main effects further. One-way ANOVAs were used to assess any differences in chronological age and GCA scores between the three groups. Independent t-tests were used to examine differences for number of training sessions between the groups. For all post-hoc comparisons, Bonferroni corrections were used to adjust for multiple comparisons.

3. Results

There was no significant effect for chronological age (in months); $F(2,57) = .903$, $p = .411$, $\eta^2_p = .032$ between the three groups. A one-way ANOVA at pre-test showed an effect for group for GCA scores[4]; $F(2,56) = 13.926$, $p < .001$, $\eta^2_p = .034$. There was no significant difference between the two training groups (*p*s = n.s.) but both training groups had lower GCAs compared to the TP group (*p*s < .001)[5]. There was no significant effect for

[3] Although performance on the cardinality task was measured by what category children belonged to, due to the size of the rating scale and the fact that the intervals in the rating scale are presumed equidistant, we carried out parametric tests. Analysing the results using non-parametric tests resulted in similar conclusions as those reported.
[4] For one child in the Digit group no GCA score could be calculated due to an experimenter error on one of the sub-tests. However, for 5 out of 6 the sub-tests that make up the GCA this child performed within the typical range (T-score > 40).
[5]As there was no significant effect for GCA scores between the two training groups and the focus of the study was to compare the two training programs, we decided not to co-vary for GCA scores in the final analysis.

the number of training sessions the training groups had received; $t(36)= -1.373$, $p= .178$, $d= .44$.

A 3 Group (PLUS, DIGIT and TP) by 3 Time points (pre-, post-, and follow-up) repeated measures ANOVA for the TEMA ability scores showed that there was an effect for Time; $F(1.728, 96) = 20.944$, $p < .001$, $\eta^2_p = .30$, for Group; $F(2, 48) = 98.304$, $p < .001$, $\eta^2_p = .80$, and an interaction between Time*Group; $F(3.456, 96) = 4.232$, $p < .001$, $\eta^2_p = .15$. For PLUS, TEMA ability scores increased over time; $F(2, 28) = 17.179$, $p < .001$, $\eta^2_p = .55$, both between pre-and post-testing ($p < .001$) and between post-testing and follow-up ($p < .001$). In the DIGIT group TEMA ability scores also increased over time; $F(2,32) = 14.029$, $p < .001$, $\eta^2_p = .47$, between pre-and post-assessments ($p < .001$) as well as between post-assessment and follow-up ($p < .001$). In contrast, there was no increase in ability scores for the TP group; $F(2, 36) = .773$, $p = .469$, $\eta^2_p = .04$. Performance in the two training groups remained below the TP group at all times (see Table 4).

Table 4 about here

For the ANS task there was a significant effect for Time; $F(2,102) = 20.392$, $p < .001$, $\eta^2_p = .29$, for Group; $F(2,51) = 20.063$, $p < .001$, $\eta^2_p = .44$, as well as a significant interaction for Group*Time; $F(2,102) = 9.092$, $p < .001$, $\eta^2_p = .26$. There was a significant effect for Time within the PLUS group; $F(2,36) = 25.626$, $p < .001$, $\eta^2_p = .59$. As shown in Figure 2, ANS scores improved both between pre-and post-assessment ($p = .002$) as well as between post-assessment and follow-up ($p = .009$). For the DIGIT group there was an effect for Time; $F(3.456, 96) = 6.549$, $p < .001$, $\eta^2_p = .21$, but ANS scores only improved between pre-and post-training ($p = .008$) but not between post-training and follow-up ($p = 1.00$). There was no effect for Time in the TP group; $F(2,34) = .425$, $p = .657$, $\eta^2_p = .02$. Although there were no significant differences between the PLUS and DIGIT groups, both groups scored below the

TP group (ps <.001) at pre- and post-training but there were no significant differences for the three groups at follow-up (p = .215)[6].

Figure 2 about here

For counting there was an effect for Time; $F(1.782, 108) = 44.411$, $p < .001$, $\eta^2_p = .45$ and for Group; $F(2,54) = 20.060$, $p < .001$, $\eta^2_p = .43$ but no interaction for Time*Group; $F(3.565, 108) = 1.373$, $p = .248$, $\eta^2_p = .05$. Overall, there was a significant difference between pre-and post-assessment ($p = .03$) and between post-assessment and follow-up ($p < .001$) for all groups. There was no significant difference between The DIGIT and PLUS group (p = .374) but both groups performed below the TP group (ps < .001). However, the large variability in performance, as shown by the standard deviations, could potentially have masked any group differences. Therefore, we also examined counting abilities using an enumeration task.

For the enumeration task there was a significant effect for Time; $F(2, 76) = 16.035$, $p < .001$, $\eta^2_p = .30$ as well as for Group; $F(2,38) = 26.043$, $p < .001$, $\eta^2_p = .58$. There was no significant interaction for Group*Time; $F(4, 76) = .809$, $p = .524$, $\eta^2_p = .04$. As can be seen in Table 3, all groups showed improved scores over time between pre- and post-assessment (p = .007) and between post-assessment and follow-up ($p = .036$). There was no significant difference between the training groups ($p = .159$) but both performed lower than the TP group; Digit Group ($p < .001$) and PLUS group ($p < .001$). However, a number of children were excluded from the analysis as they were identified as outliers, showing that performance on this measure was very variable.

For cardinality, most of the children in the TP group (N = 17/20) were already classified as 5-knowers at pre-assessment on the Give-a-number task, and thus only the PLUS

[6] We also examined the differences for incongruent and congruent trials only and these results were the same as including all of the trials in the analysis.

and DIGIT training groups were included in the analyses (see Table 5). There was an effect for Time; $F(1.601, 54.418) = 30.498$, $p < .001$, $\eta^2_p = .47$, but there was no effect for Group; $F(1, 34) = .034$, $p = .855$, $\eta^2_p = .001$, and no interaction for Time*Group; $F(1.601,54.418) = 1.102$, $p = .338$, $\eta^2_p = .03$. Children in both the DIGIT and PLUS groups showed better cardinality abilities immediately after the training programs ($p = .001$) as well as at follow-up ($p < .001$).

Table 5 about here

A repeated measures ANOVA generated similar findings for the digit recognition task: there was a significant effect for Time; $F(2, 94) = 51.777$, $p < .001$, $\eta^2_p = .52$ and for Group; $F(2,47) = 13.785$, $p < .001$, $\eta^2_p = .37$ but no interaction for Time*Group; $F(4, 94) = .160$, $p = .958$, $\eta^2_p = .01$. For all groups, there was a significant improvement at post-assessment ($p = .004$) and follow-up ($p < .001$). Although there was no significant difference in performance between the training groups ($p = .686$), they both performed below the TP group (ps $< .001$).

For letter knowledge there was an overall effect for Time; $F(2, 96) = 12.937$, $p < .001$, $\eta^2_p = .21$, and for Group; $F(2, 48) = 32.216$, $p < .001$, $\eta^2_p = .57$, but there was no interaction for Group*Time; $F(4, 96) = 1.833$, $p = .129$, $\eta^2_p = .07$. Pre- assessment scores did not differ significantly from post-assessment scores ($p = .815$) but scores at follow-up differed from post-assessment ($p = .001$). Although both training groups performed lower than the TP group (ps $<.001$), there were no significant differences between their scores ($p = 1.00$).

4. Discussion

The importance of symbolic versus non-symbolic knowledge as a predictor for mathematical development has been the focus of many debates. However, much of the evidence in this debate relies on correlational studies. In contrast, intervention studies allow examination of the causal nature of cognitive factors that contribute to mathematical abilities. The current study assessed the effects of two different training programs: one focusing on symbolic knowledge (DIGIT) and one focusing on non-symbolic knowledge (PLUS). Participants included LA preschoolers who had not yet received formal mathematical instruction and performed low on the ANS task. We selected this small group specifically so that there would be less variability within the group as all children had both symbolic and non-symbolic difficulties. This would allow us to examine which abilities would be improved through our different training programs. It was predicted that if symbolic knowledge drives mathematical abilities than LA children in the DIGIT group should show improved mathematical abilities, whilst if non-symbolic abilities drive mathematical achievement, LA children in the PLUS group would improve.

The results showed that both training groups had improved mathematical ability scores immediately after the training as well as at follow-up, in contrast to the TP group whose scores remained constant. Letter recognition abilities did not increase during the training suggesting that the training programs made a difference to children's mathematical abilities and not their general learning abilities.

In line with the ANS mapping account, we found that non-symbolic training improved ANS abilities as well as symbolic abilities. This finding is in line with previous studies that have shown that young children use their ANS representations to solve tasks involving number words and symbols and that improved ANS abilities relate to improved mathematical performance (Wang et al., 2016). However, those children who received symbolic training, not only improved on symbolic abilities, they also showed improved ANS abilities. In contrast to the dual-representation model (Carey, 2009), this would suggest that children do access their ANS representations when learning new integers early on in development. Alternatively, it could be that they did not access their ANS representations whilst learning but that the symbolic training improved their general understanding of how numbers relate to each other and that this improved knowledge facilitated their ANS performance (Mussolin et al., 2014). The current study does not allow further insight into whether or not children accessed their ANS representations during the symbolic training program or the assessment.

There were no differences between the groups for any of the symbolic knowledge tasks and all groups improved on the counting, enumeration, and the digit recognition tasks. For cardinality, there was no difference between the two training groups. We had predicted that if symbolic knowledge drives mathematical abilities that children in the DIGIT group would show more improved digit recognition and counting abilities, beyond any improvements that would be observed in the PLUS group. First, it is important to note that although all groups significantly improve on symbolic counting and digit recognition tasks, these differences are minimal in real terms. For example, at post-assessment the groups recognized on average one or two more digits compared to the pre-assessment. It is unclear why the improvements are so small in the DIGIT group but it may be that these children were asked to count on a daily basis and that they were not motivated to count again during post- and follow-up assessments.

Recent studies have shown that ordinal processing is crucial for mathematical performance (see Lyons, Vogel, & Ansari, 2016 for a review) and that symbolic number-ordering abilities mediate the correlation between ANS and mathematical abilities (Lyons & Beilock, 2011). Further examination of the games showed that both training programs include a game in which children have to say whether or not the stimuli, either symbolic or non-symbolic, are in the correct order or not. Therefore, seeing the importance of ordinality for improving mathematical abilities, these ordinality training games may explain why both training groups improved on both symbolic and non-symbolic assessments, in contrast to previous studies, such as Honoré and Noël (2016), which only found improvements for symbolic training programs. At the time of the development of the training programs, ordinality was an under-researched facet of numerical cognition and its importance was not yet known. Although this could be seen as a confound in the current study, the current study is the first intervention study that may provide further evidence for the importance of ordinality to improve number abilities in preschoolers. This finding should be further explored in future studies by comparing symbolic and non-symbolic training studies that only include games that target ordinality processing.

The sample size in the current study was rather limited, due to the fact that preschoolers in the UK are not obliged to attend preschool and that we included a specific group of preschoolers in order to limit the heterogeneity within our sample. Previous studies have shown that children are at risk for mathematical difficulties for a number of reasons (see Costa, Nicholson, Donlan, & Van Herwegen, 2018), by applying strict inclusion criteria it is possible that depending on the cause of their mathematical difficulty not all children may have benefitted equally from the training program they had been allocated to. Therefore, the current study should be replicated using a larger sample size. Similarly, due to the specificity of the sample included and thus only small number of LA children could be recruited, the current study did not include a control group of LA children who did not receive any intervention. Although the results from the letter recognition task suggest that the training programmes affected mathematical abilities specifically, future studies should examine the extent to which general maturation affected the outcomes of the study. Finally, similar to most training studies that have focused on symbolic knowledge, none of the children in the current study caught up with the control group immediately after the training programs (see Mononen et al., 2014 for a review). It is thus unclear whether the children ever caught up with the control group and which children continued to perform low on mathematical ability tasks. Future studies should thus examine the impact of symbolic and non-symbolic training studies in older LA children who are receiving formal symbolic instruction, as it can be argued that improving ANS abilities is time sensitive and that older LA children will no longer benefit from non-symbolic training programs (Xenidou-Dervou et al., 2017). In addition, seeing the benefit of the training programs to overall mathematical abilities, it would worth examining whether children with other developmental disorders that impact mathematical abilities, can benefit to a similar extent.

5. Conclusions

Although further studies are required about the underlying mechanisms that caused the effects in the current study, the results showed that preschoolers with low mathematical abilities benefitted from training programs that target both symbolic and non-symbolic training programmes. In addition, the current study is the first intervention study to show that ordinality might play an important role in improving mathematical abilities. Traditionally, preschool instruction in the UK is informal and happens during play or in games with LA children receiving very little additional support. Seeing the bi-directional relationship between symbolic and non-symbolic abilities and that LA children benefitted both short-term as well as long-term from the DIGIT and PLUS training programs, preschool education should focus on games that target both symbolic and non-symbolic mathematical abilities as this may benefit children who are low achievers on mathematical ability tasks.

Acknowledgements:

We would like to thank all the preschool staff and children that have taken part in this study. This project was funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily those of the Foundation.

6. References

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*(4), 445-454. doi: 10.1177/1745691613491271

Carey, S. (2009). Where our number concepts come from. *Journal of Philosophy*, *106*(4), 220-254. doi: 10.5840/jphil2009106418

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, *148*, 163-172. doi: 10.1016/j.actpsy.2014.01.016

Costa, H. M., Nicholson, B., Donlan, C., & Van Herwegen, J. (2018). Low performance on mathematical tasks in preschoolers: the importance of domain-general and domain-specific abilities. *Journal of Intellectual Disability Research.* Advance online. doi: 10.1111/jir.12465

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). The relationship between symbolic and non-symbolic numerical magnitude processing and the typical and atypical development of mathematics: a review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*, 48-55. doi: 10.1016/j.tine.2013.06.001

Dillon, M. R., Pires, A. C., Hyde, D. C., & Spelke, E. S. (2015). Children's expectations about training the approximate number system. *British Journal of Developmental Psychology*, *33*(4), 411-418. doi: 10.1111/bjdp.12118

Eliot, C. D. & Smith, P. (2011). *British Ability Scales (Third edition): BAS3*. GL Assessment.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representation to each other and to mathematical achievement. *Journal of Experimental Child Psychology*, *123*, 53-72. doi: 10.1016/j.jecp.2014.01.013

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, *8*, 307–314. doi: 10.1016/j.tics.2004.05.002

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLoS ONE, 8*(1), e54651. doi:10.1371/journal.pone.0054651

Gebuis, T., & Reynvoet, B. (2011). Generating non-symbolic number stimuli. *Behavioural Research Methods*, *43*, 981–986. doi: 10.3758/s13428-011-0097-5

Ginsburg, H. P., & Baroody, A. J. (2003). T*est of Early Mathematics Ability: Third Edition.* Austin, TX: PRO-ED.

Honoré, N., & Noël, M-P. (2016). Improving Preschoolers' Arithmetic through Number Magnitude Training: The Impact of Non-Symbolic and Symbolic Training. *PLoS ONE, 11*(11), e0166685. doi:10.1371/journal.pone.0166685

Huck, S. W., & McLean, R. A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin*, *82*(4), 511-518.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131,* 92-107. doi: 10.1016/j.cognition.2013.12.007

Kuhn, J. T. & Holling, H. (2014). Number sense or working memory? The effect of two computer-based trainings on mathematical skills in elementary school. *Advanced Cognitive Psychology, 10*(2), 59-67. doi: 10.5709/acp-0157-2

Lipton, J.S. & Spelke, E.S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396-401. doi: 10.1111/1467-9280.01453

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*, 256–261. doi: 10.1016/j.cognition.2011.07.009

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L. & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, *17*, 714-726. doi: 10.1111/desc.12152

Lyons, I.M., Vogel, S.E., Ansari, D. (2016). Chapter 8 - On the ordinality of numbers: A review of neural and behavioral studies. In M. Cappelletti & W. Fias (Eds.), *Progress in Brain Research,* 227, 187-22. doi: 10.1016/bs.pbr.2016.04.010.

Mononen, R., Aunio, P., Kaponen, T., & Aro, M. (2014). A review of early numeracy interventions for children at risk in mathematics. *International Journal of Early Childhood Special Education*, *6*(1), 25-54.

Mussolin, C., Nys, J., Content, A. & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS ONE*, *9*(3), e91839. doi: 10.1371/journal.pone.0091839

Negen, J., & Sarnecka, B. W. (2015). Is there really a link between exact-number knowledge and approximate number system acuity in young children? *British Journal of Developmental Psychology*, *33*(1), 92-105. doi: 10.1111/bjdp.12071.

Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General*, *143*(1), 255-265. doi: 10.1037/a0030825

Park, J., & Brannon, E. M. (2013). Training the Approximate Number Sense improves math proficiency. *Psychological Science, 24*(10), 2013-2019. doi: 10.1177/0956797613482944

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition*, *133*, 188-200. doi: 10.1177/0956797613482944

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources, 27,* 313-328.

Reynvoet B., & Sasanguie D. (2016). The Symbol Grounding Problem Revisited: A Thorough Evaluation of the ANS Mapping Account and the Proposal of an Alternative Account Based on Symbol-Symbol associations. *Frontiers in Psychology*, *7*, 1581. doi: 10.3389/fpsyg.2016.01581

Sasanguie D., De Smedt B., & Reynvoet B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychological Research*, *81*(1), 231-242. doi: 10.1007/s0042

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S.S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, *20*(3). doi: 10.1111/desc.12372

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, *75*(3), 417–453. doi:10.3102/00346543075003417

Starr, A., Libertus. M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Science of the Unites States of America, 110*(45), 18116-18120. *doi/10.1073/pnas.1302751110*

Van de Walle, J. A., & Lovin L. A. H. (2006). *Teaching Student-Centered Mathematics, Grades K–3.* Boston, MA: Pearson.

Van Herwegen, J., Costa, H. M., Passolunghi, C. M. (2017). Improving number abilities in preschoolers: the preschool number learning scheme PLUS. *School Psychology Quarterly*. Ahead of print.

Van Herwegen, J. & Karmiloff-Smith, A. (2015). Genetic developmental disorders and numerical competence across the lifespan. In R. Cohen Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition*. Oxford: Oxford University Press.

Wang, J. J., Odic, D., Halberda, J., & Feigenson, L. (2016). Changing the precision of preschoolers' approximate number system representations changes their symbolic math performance. *Journal of Experimental Child Psychology*, *147*, 82-99. doi: 10.1016/j.jecp.2016.03.002

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220-251.

Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., & van Lieshout, E.C.D.M. (2017). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learning and Instruction*, 50, 1-13. doi: 10.1016/j.learninstruc.2016.11.001

Xu, F., & Spelke, E.S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 10(74), B1-B11.

Table 1

Descriptive statistics for the training and control groups: Chronological Age (CA), General Cognitive Ability (GCA) scores and amount of training sessions

	PLUS Training group N=19 (6 Female)		DIGIT Training group N=19 (11 Female)		Control group N=20 (10 Female)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
CA in months	43.42	(3.99)	44.21	(4.52)	45.20	(3.91)
GCA scores	90.79	(9.67)	91.89	(12.95)	108.90	(13.11)
Training sessions	13.11	(2.18)	14.05	(2.07)	N/A	N/A

Table 2

Socio-economic status information for participant groups

	PLUS Training group	DIGIT Training group	Control Group
No qualifications	20%	33%	18%
Secondary school	27%	17%	9%
Vocational Qualifications	27%	0%	18%
Undergraduate degree	6%	8%	9%
Postgraduate degree	20%	42%	46%

Table 3

Description of the PLUS and DIGIT training games

PLUS games: matching & estimation games	**DIGIT games: digit recognition & counting games**
In a line: matching game. Child needs to organize cards with amounts of objects in the right order from not so many to many objects and from many to not so many.	*Digit snake*: digit game. Child was given some digits and asked to sort them in the correct order from left to right.
Matching amounts: Matching game. Using cards with 5 to 20 different sized coloured objects children were asked to match the quantities on the card in their hand to cards presented on the table (lots with lots or few with few objects).	*Match digits to numbers*: digit game. Children are shown cards with digits on and need to say the correct digit. Child who responds the first wins the card. Child with the most cards wins.
Action game: matching game. One child/ experimenter performs an action (i.e. claps her hands a lot or only a few times) and the other child has to copy it approximately.	*Counting actions*: counting game. Child one or researcher performs an action with certain number of repetitions and child is asked to count the actions.
Domino game: matching game. Special cards were used that display a scattered number of dots on the left and right side of the card of varied ratios (e.g. 8 dots on left of card versus 16 dots on right of card). The child was asked to identify which side of the card contained more dots and to match the correct side (large or small number) with the card presented on the table.	*Show me the number*: The researcher said a number name and the child had to identify the correct digit from a group of printed coloured digits cards.
Grab and Guess: estimation game. Children were asked to grab some uncooked pasta/blocks of different shapes and sizes from a box. The researcher then grabbed a different amount and the child was asked who had more (or less).	*Grab and Count*: counting game. Grab a number of same shape blocks and count how many shapes there are. The child who has counted their pile correctly first wins
In the sock: estimation game. The researcher hid two quantities of different sized beads in two different socks. The child was then asked to feel both of the socks with each hand and to guess which sock contained more (or fewer) beads.	*Hidden digits*: digit game. The child was shown a book and asked to spot any digits on the pages as fast as they could.
Play that number: estimation game. Children had to guess which instrument played more/less sounds.	*Number Rhymes: Counting game.* Children sang English number rhymes to learn the number names.
Guess!: estimation game. Children were shown two cards with different sized objects on and had to say which card had more/less object as fast as possible.	*Counting cards game*: counting game. Children were shown a card with objects on and asked to count the objects.

Table 4

Overall mathematical abilities as measured by Test or Early Mathematical Abilities (TEMA) Raw scores and Ability scores, ANS raw scores, Counting abilities as measured by highest number counted correctly in verbal counting and Enumeration task, total number of digits (Digit Recognition task) and total number of letters recognised (Letter Recognition task).

			Pre		Post		Follow-up	
Task	Group	N[7]	Mean	SD	Mean	SD	Mean	SD
TEMA (Ability)	PLUS	15	90.20	5.36	95.60	5.24	102	7.88
	DIGIT	17	88.82	4.43	96.88	7.27	98.71	9.96
	Control	19	116.63	8.22	119.37	7.73	117.89	8.74
ANS	PLUS	19	24.05	3.69	28.05	4.48	32.42	5.18
	DIGIT	17	26.47	2.21	30.65	4.74	31.35	5.11
	Control	18	34.83	2.68	34.56	3.36	34.00	4.75
Counting	PLUS	18	8.78	4.49	10.61	4.49	15.2	6.54
	DIGIT	19	10.63	5.45	14.00	5.45	20.16	10.00
	Control	20	20.65	9.12	21.80	9.12	31.20	9.46
Enumeration	PLUS	13	10.00	4.83	13.00	1.63	13.69	2.63
	DIGIT	11	13.00	1.61	14.09	2.95	15.27	4.34
	Control	17	16.65	3.46	17.38	3.33	20.00	0.00
Digit recognition	PLUS	13	4.77	2.83	5.92	3.15	8.38	2.02
	DIGIT	19	3.58	3.00	5.16	4.75	7.32	4.51
	Control	18	9.33	2.81	10.33	4.34	13.06	3.90
Letter Recognition	PLUS	15	1.27	1.33	1.40	1.45	3.00	2.14
	DIGIT	16	1.81	3.19	1.69	2.44	2.94	2.74
	Control	20	10.65	7.18	12.20	7.64	14.55	6.96

[7] Number of participants per group after excluding outliers.

Table 5

Performance on Cardinality task at Pre-, Post-, and at Follow-up (FU) assessment for the PLUS, DIGIT, and Control group

Group	*PLUS*			*DIGIT*			*Control*		
Time	*Pre*	*Post*	*FU*	*Pre*	*Post*	*FU*	*Pre*	*Post*	*FU*
0-knower	*1*	*0*	*1*	*1*	*0*	*0*	*0*	*0*	*0*
1-knower	*3*	*4*	*0*	*5*	*3*	*1*	*0*	*0*	*0*
2-knower	*6*	*2*	*0*	*5*	*4*	*5*	*1*	*0*	*0*
3-knower	*4*	*5*	*3*	*3*	*4*	*1*	*0*	*0*	*0*
4-knower	*4*	*5*	*2*	*3*	*3*	*1*	*2*	*2*	*0*
5-knower	*1*	*2*	*12*	*2*	*5*	*11*	*17*	*18*	*20*