Research in Developmental Disabilities 112 (2021) 103919

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

Parent-based training of basic number skills in children with Down syndrome using an adaptive computer game

Silvia Lanfranchi [a,*], Sara Onnivello [a], Maristella Lunardon [a], Francesco Sella [b], Marco Zorzi [a,c,**]

[a] University of Padova, Italy
[b] Loughborough University, UK
[c] IRCCS San Camillo Hospital, Venice, Italy

ARTICLE INFO

Number of reviews completed is 2

Keywords:
Down syndrome
Basic number skills
Training
Number Race

ABSTRACT

Background: Numeracy is an area of difficulty for children with Down syndrome (DS). It has been demonstrated that *The Number Race*, a non-commercial adaptive computer game designed to foster basic mathematical abilities, represents a promising instrument to enhance these skills in children with DS when delivered by an expert in a clinical setting.
Aims: In the present study, we assessed the efficacy of *The Number Race* when administered at home by properly instructed and remotely supervised parents.
Methods and procedures: Basic numerical skills were assessed before and after training, as well as at three-months follow-up. Performance of children with DS who worked at home with the parent (PG) was compared with that of children who received the training by an expert (EG). For both groups, the training lasted ten weeks, with two weekly sessions of 20−30 min.
Outcomes and results: Results show that both groups improved across various measures of numerical proficiency, including the overall score of the numeracy assessment battery, while only the EG showed an improvement in a measure of mental calculation. The improvements were maintained three months after the end of the training.
Conclusions and implications: These findings confirm the efficacy of *The Number Race* and extend it to an home-based setting, whereby parents administer the training with external supervision.

What this paper adds?

In a previous study, the non-commercial adaptive computer game *The Number Race* was shown to be effective in fostering number sense in children with Down syndrome. In that case, training was administered in a clinical setting by a psychology researcher. Here we show that similar beneficial outcomes of *The Number Race* can be obtained when the training is conducted by parents at home. Information of this nature makes it possible to rethink administration modality and duration for this numerical training and for cognitive training in general. Our results pave the way for broader and no-cost access to the Down syndrome population, exploiting the familiar home-based setting and the possibility of prolonged training.

* Corresponding author at: Department of Developmental Psychology and Socialization, Via Venezia 8, 35131, Padova, Italy.
** Corresponding author at: Department of General Psychology, Via Venezia 8, 35131, Padova, Italy.
E-mail addresses: silvia.lanfranchi@unipd.it (S. Lanfranchi), marco.zorzi@unipd.it (M. Zorzi).

https://doi.org/10.1016/j.ridd.2021.103919
Received 1 February 2021; Accepted 22 February 2021
Available online 6 March 2021

1. Introduction

Down syndrome (DS), or trisomy 21, is the most common cause of intellectual disability of genetic origin, affecting about 1 in 700–1000 live births (e.g., McGrother & Marshall, 1990; Sherman, Allen, Bean, & Freeman, 2007). DS is characterized by intellectual disability in the great majority of cases, with a broad variability in terms of degrees of severity (e.g., Grieco, Pulsifer, Seligsohn, Skotko, & Schwartz, 2015).

Previous research has indicated that DS is associated with a specific behavioural phenotype, characterized by speech and language impairments (Chapman & Hesketh, 2000), with greater difficulties in expressive than in receptive language. Research has also shown that individuals with DS are extremely limited in terms of memory span, especially in auditory-verbal memory (Jarrold, Baddeley, & Hewes, 1999; Vicari, Marotta, & Carlesimo, 2004), as well as in executive functions, particularly working memory (Lanfranchi, Cornoldi, & Vianello, 2004; Lanfranchi, Baddeley, Gathercole, & Vianello, 2012), inhibition, planning, and cognitive flexibility (Lanfranchi, Jerman, Dal Pont, Alberti, & Vianello, 2010; Lee et al., 2011). Some aspects of motor functioning seem to be relatively impaired too (Block, 1991). On the other hand, non-verbal skills are reportedly less severely affected, though recently a variable picture has emerged depending on which aspect of visuospatial cognition is considered (Yang, Conners, & Merrill, 2014). Another area of relative strength is social functioning (Fidler, 2005).

Several studies have focused on basic number skills in individuals with DS (Onnivello, Lanfranchi, & Zorzi, 2019, for review). These abilities are fundamental not only to support math achievement but also for daily living, to adapt to the demands of the environment and thus have an autonomous life. Children are immersed in an environment rich with quantitative information and numerical experiences. They see numbers everywhere around them, and they see people using numbers to describe, count or measure. Infants represent small quantities in their mind and discriminate between them (Antell & Keating, 1983); moreover, they are able to discriminate between large quantities, although discrimination acuity is progressively refined during childhood (e.g. Halberda & Feigenson, 2008; Mou & Van Marle, 2014; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). Later on, children learn how to count arrays of objects and perform small calculations (for an overview see Sella, Hartwright, & Cohen Kadosh, 2018). Children with DS usually show poor basic number skills in comparison to children with the same chronological age. However, while some numerical abilities appear to be simply delayed and follow the typical developmental trajectory, for other skills there is evidence of specific impairments even with respect to mental age.

The ability of children with DS to represent and compare non-symbolic numerical quantities has been a major topic of investigation. Two core mechanisms represent non-symbolic numerical information (Feigenson, Dehaene, & Spelke, 2004): the Object Tracking System (OTS) provides an object-based exact representation of small sets (up to 4), whereas the Approximate Number System (ANS) encodes larger numerosities (i.e., more than 4) in an analogical and approximate way. Individuals with DS make more errors, even with respect to typically developing children with the same mental age, in comparing small numerosities (in particular 2vs3 and 3vs4; Paterson, Girelli, Butterworth, & Karmiloff-Smith, 2006; Sella, Lanfranchi, & Zorzi, 2013). This evidence suggests that subitizing, that is the accurate perception of small numerosities (up to 4) supported by the OTS, develops atypically in individuals with DS. In contrast, their ability to compare large numerical quantities, which relies on the ANS, seems to be functioning in line with mental age (Abreu-Mendoza & Arias-Trejo, 2015; Camos, 2009; Paterson et al., 2006; Sella, Lanfranchi, & Zorzi, 2013). Moreover, the performance of children with DS shows the typical ratio effect (the discriminability of two quantities becomes more difficult as the numerical ratio of the two sets gets closer to 1), supporting the idea that they use the same processes deployed by typically developing children (Abreu-Mendoza & Arias-Trejo, 2015; Camos, 2009).

Two studies investigated the ability to map numerical quantities onto a visual "number line" in DS (Lanfranchi, Berteletti, Torrisi, Vianello, & Zorzi, 2015; Simms, Karmiloff-Smith, Ranzato, & Van Herwegen, 2020). In these studies the performance of children with DS resulted aligned with that of typically developing children of the same mental age rather than chronological age. Therefore, numerical estimation skills in this population seems to be related to cognitive skills, rather than to experience or education.

Another area of difficulty in individuals with DS is counting, and there has been a broad debate on whether they have a superficial or a deep understanding of counting. Gelman and Cohen (1988), for example, found that children with DS performed less well than preschoolers matched for mental age in both counting and cardinality tests, and suggested that they merely count by rote, with no understanding of the cardinality principle. Cornwell (1974) also came to the conclusion that the counting performance in children with DS was not guided by principles, and the author added that a lack of understanding of counting principles made it impossible for these children to acquire higher-level arithmetical concepts. On the contrary, other studies found that individuals with DS could understand cardinality, though they might count shorter sequences and smaller arrays of objects. For instance, Caycho, Gunn, and Siegal (1991) replicated the Gelman and Cohen study, matching a group of children with DS with typically developing children on receptive vocabulary, concluding that children with DS can make use of counting principles and that their counting competence is related more to their receptive language than to DS per se. Bashash, Outhred, and Bochner (2003) likewise found that children with DS were able to apply the fundamental principles of counting in several counting situations. Similarly, Sella et al. (2013) found that, despite a slow and slightly less accurate performance, children with DS show a level of counting proficiency and cardinality understanding that is adequate for their mental age.

Research has focused also on calculation, showing that this is a particularly difficult area for individuals with DS (Marotta, Viezzoli, & Vicari, 2006). They have difficulties with simple calculations within small ranges, such as 5 or 10 (Lanfranchi et al., 2015), so that school-age children with DS perform on this task even poorer than typically developing preschooler matched for mental age.

To sum up, basic number skills seem to be an area of particular impairment in individuals with DS. Some of these abilities, such as number estimation and the comparison of large quantities, display only a delayed developmental trajectory which is aligned with mental age, while other abilities, such as comparing small quantities, counting and mental calculation, seem to develop atypically and remain impaired even with respect to the overall cognitive level. The focus on basic number skills is strongly motivated by their relation to formal mathematical achievement, both in typical development (e.g., Halberda, Mazzocco, & Feigenson, 2008; Jordan, Kaplan, Oláh, & Locuniak, 2006) and in children with specific impairments, such as children with dyscalculia (Butterworth, Varma, & Laurillard, 2011; Piazza et al., 2010; Sella, Berteletti, Brazzolotto, Lucangeli, & Zorzi, 2014). Moreover, it has been demonstrated that training non-symbolic arithmetic leads to an improvement in symbolic arithmetic (Hyde, Khanum, & Spelke, 2014; Park & Brannon, 2013). This suggests that supporting numeracy development in DS children with specific training programs is of paramount importance.

In recent years, a non-commercial software developed by leading experts of numerical cognition in the form of an interactive and adaptive computer game has been successfully used to foster basic number skills both in typically developing children and in children with dyscalculia: *The Number Race* (Wilson, Dehaene et al., 2006). The game is based on four principles: enhancing number sense, cementing the links between different ways of representing numbers, conceptualizing and automatizing arithmetic, and maximizing motivation. Previous studies have demonstrated the efficacy of this training both in typically developing children (Sella, Tressoldi, Lucangeli, & Zorzi, 2016) and in children with dyscalculia (Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). The Number Race has also been recently used with children with DS (Sella, Onnivello, Lunardon, Lanfranchi, & Zorzi, 2021). In the latter study, training was administered by an expert and lasted for ten weeks, with two weekly sessions of 20−30 min. The study design also included a control group of children with DS who worked with a software aimed to foster reading skills (also delivered by an expert). Results showed that children who were trained with *The Number Race*, compared to the control group, improved in specific numerical skills and in mental calculation; moreover, these gains were maintained three months after the end of the training.

The aim of the present study is to explore the possibility of training basic number skills at home under the supervision of parents. In the study by Sella et al. (2021), each child played with *The Number Race* under the supervision of an expert. However, taking a child twice a week to a specialized center can be challenging for many parents, due to the great number of weekly activities in which these children are usually involved. Implementing the same training at home can be a great advantage for parents. Previous studies have already demonstrated the feasibility of a cognitive training administered by parents. For example, Conners, Rosenquist, Arnett, Moore and Hume (2008) found that a rehearsal training administered at home by parents was effective in improving memory span in children with DS. Moreover, Pulina, Carretti, Lanfranchi, and Mammarella (2015) demonstrated that, with adequate support, parents could effectively train working memory in their children. Although parent-implemented intervention may entail intervening variables, we believe that it can have greater ecological validity and, if successful, improvements could be sustained by means of regular maintenance exercises. Therefore, it is worthwhile to assess whether giving parents guidance on how to work with *The Number Race* on basic number skills could produce similar results to those achievable by an expert.

2. Materials and methods

2.1. Participants

Children with DS were recruited from local associations, which offer support to families of children with intellectual disabilities, or through acquaintances. All children lived and were included in mainstream schools in north-eastern Italy. Informed written consent was obtained from parents. The study was approved by the Ethics Committee for Psychological Research of the University of Padova (Italy) and it conformed to the standards of the Declaration of Helsinki. Overall, forty-one children with DS received training through the computer game *The Number Race* (Wilson, Dehaene et al., 2006) in its Italian version (*La gara dei Numeri*; Berteletti, De Filippo, De Grazia & Zorzi, 2012). Training was administered either at home by a parent (Parent Group, hereafter PG; N = 21) or by an expert in psychology (Expert Group, hereafter EG; N = 20).

The present work is part of a wider project, in which as a first step the EG was compared with an active control group (CG) of children with DS who practised literacy skills (with a specific software and under expert supervision) to assess the effectiveness of training with *The Number Race* (Sella et al., 2021). As a second step of the project, in the present paper the same EG data are used as a comparison group to statistically assess the effect of the parent-guided home-based training delivered to PG children. In the present study performance of CG children is reported in the graphs to appreciate the specific effect of numerical training (as it controls for the effect of time and individual work with children) and it is used as a baseline to compute effect sizes for both PG and EG.

Table 1
EG and PG group characteristics.

	EG (n = 20)	PG (n = 21)	$t(39)$	p	BF_{10}
Chronological age (months)	118.10 (24.59) [70−149]	128.43 (25.12) [81−174]	−1.33	0.19	0.62
PPVT-R age equivalent (months)	55.90 (21.09) [27−103]	56.71 (32.02) [27−181]	0.14	0.88	0.31
CPM age equivalent (months)	68.50 (18.61) [39−111]	54.86 (16.06) [30−92]	2.52	0.02[a]	3.45

Note. Mean (SD) and group comparisons are reported. EG = Expert group; PG = Parent group; BF=Bayes Factor; BF

[a] In light of the significant difference between PG and EG in the non-verbal mental age estimates, analyses were repeated considering only a subgroup of PG participants matched to the EG group for this measure. The results are reported in the Supplementary Information and are fully aligned with those reported in the main text.

General cognitive level was assessed in all groups using a measure of non-verbal ability – *Raven's Colored Progressive Matrices* (CPM; Belacchi, Scalisi, Cannoni, & Cornoldi, 2008; Raven, Raven, & Court, 1998), and a measure of verbal ability – the *Peabody Picture Vocabulary Test-Revised* (PPVT-R; Dunn & Dunn, 1997; Stella, Pizzioli, & Tressoldi, 2000). Participants' characteristics are presented in Table 1.

2.2. Apparatus and procedure

2.2.1. Pre-training, post-training and follow-up measures

Each test phase consisted of 4 meetings of around 45 min each. The three test phases were set immediately before the training (pre-test), immediately after the training (post-test) and the last one carried out three months after the end of the training (follow-up). The researcher met each child individually in a comfortable and quiet room. A series of paper and pencil tasks was administered in order to assess different aspects of numerical cognition. Moreover, several computerized tasks were administered for the same purpose. Considering that the majority of computerized tasks assess the same aspects of numerical cognition of the paper and pencil tasks, and that for these tasks the follow up data of 6 children are missing for the EG and 1 chid for the PG due to a computer failure, we report description and results of the computerized tasks as supplementary material in the Supporting Information section. Only description and results for the *Mental Calculation* task are reported in the main text since this is the only measure of arithmetic proficiency in the experimental protocol.

2.2.1.1. Numerical Intelligence Battery (BIN: Batteria Intelligenza Numerica). The *BIN* test (Molin, Poli, & Lucangeli, 2007) assesses several numerical skills, which are the precursors of later mathematical and arithmetical learning. It is composed of four subscales: *Lexical, Semantic, Counting* and *Pre-syntactic*. The battery demonstrated good psychometric properties, with all subscales presenting high reliability (Lexical subscale: r = 0.89, Semantic subscale: r = 0.69, Counting subscale: r = 0.74, Pre-syntactic subscale: r = 0.79). The *Lexical* subscale assesses the ability to read and write Arabic numbers as well as the ability to connect number-words to the correct digits. The *Semantic* subscale measures the ability to compare non-symbolic (i.e., dots) and *Symbolic* (i.e., Arabic digits) numerical quantities. The *Counting* subscale assesses the ability to recite the number–words sequence forward and backward as well as the knowledge of the order of Arabic digits from 1 to 5. The *Pre-syntactic* subscale evaluates the ability to link numbers to their quantity representation and to order multiple quantities. For each subscale, the number of correct responses was calculated. The scores in the four subscales are summed to obtain a total score, which is here considered as an overall index of basic numerical abilities.

2.2.1.2. Number Comparison. In this task, children were asked to choose the larger between two number words. In each trial, the experimenter said: "Which is 'more' between *x* candies and *y* candies?", where *x* and *y* were two numbers from one to nine. The to-be-compares pairs of numbers were: 4-2, 2-7, 3-8, 2-1, 8-7, 5-4, 3-6, 7-6, 1-5, 9-3, 1-4. We calculated the percentage of errors as the outcome measure.

2.2.1.3. Number-To-Position (NTP). In the *Number-to-Position* (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Siegler & Opfer, 2003) children were presented with a 20-cm line on a white landscape sheet. The ends of the lines were labeled with 1 on the left and with either 10 or 20 on the right. Two intervals were administered, one at the time: 1−10 and 1−20. The child had to place a given number on the line. The number to be positioned was shown on the left upper corner of the sheet. For each interval, there were eight randomly presented numbers (i.e., 2, 3, 4, 5, 6, 7, 8, 9 for the 1–10 interval; 2, 4, 6, 7, 13, 15, 16, 18 for the 1–20 interval). For each number to place, a separate sheet with the number line was presented. The following instructions were given: "Now we are going to play a game with number lines. You can see that this line goes from 1 to 10 (or 20). I will tell you a number and you have to indicate which is the place of this number on the line, as precise as you can." The instructions were repeated as many times as needed but no feedback was given. As training trials, children had to place 1 and 10 in the 1−10 line, whereas in the 1−20 line the training trials were 1 and 20. The experimenter named the target number on every trial. In most cases, children were given a pencil and they drew a vertical mark on the line where they thought the target number should be; some children had difficulties in holding the pencil, so they were asked to point with their finger the place of the target number on the line, as precise as they could, and the experimenter made the mark. An index of accuracy on this task was obtained by computing the percentage of absolute error, i.e. (|Estimate-Target Number|/Numerical Interval)*100. For each child the mean percentage of error in both intervals was computed.

2.2.1.4. Mental Calculation. In the *Mental Calculation* task, 18 operations (8 additions and 10 subtractions) were individually shown on a computer screen and read aloud by the experimenter. Children had to give the result as fast as they could and were given no support to solve the problem; however, they were allowed to use their own fingers. Verbal answers were recorded by the experimenter on the computer using the keyboard. For each task, the percentage of correct responses was computed.

2.2.2. Training

Training lasted for ten weeks, with two weekly sessions of 20−30 min each. EG children worked individually with the support of a research assistant in a clinical context, whereas PG children worked at home with the support of a parent. An expert in psychology met all parents who agreed to deliver the training at home to provide information about basic number skills, instructions on how to conduct the training program, how to work with their children and how to sustain motivation. Moreover, the expert supervised parents through weekly phone calls: feedback was given and questions were answered. Finally, the expert was available for parents' phone calls whenever needed for additional advice or troubleshooting. All participants completed the training.

2.2.2.1. The Number Race. Children played with the Italian version of *The Number Race* game (Berteletti, De Filippo De Grazia, & Zorzi, 2012; Wilson, Dehaene et al., 2006). Players compete against the software in a numerical comparison task, choosing the larger between pairs of numerical quantities that range from 1 to 9, which may be sets of objects or digits; moreover, as the difficulty of the game increases, children are also required to solve additions or subtractions to perform the comparison with numbers in the 1−10 range. The players choose "which is more", while the other quantity is given to the opponent (the software). The difficulty of the comparison is increased or decreased by changing the time the stimuli remain on the screen, the size of the dots in the non-symbolic comparisons, or the numerical distance between the to-be-compared quantities. An adaptive algorithm modulates the difficulty of the game to keep it optimally challenging, thus working on the "zone of proximal learning" (Vygotsky, 1976). After each numerical comparison, players move their character forward along a number line on a board, and the race ends when the first player reaches the end of the board. Verbal feedback is constantly provided to maintain motivation.

2.3. Data analysis

In order to evaluate the efficacy of the training run by a parent compared to the training run by an expert, we analyzed results on numerical tasks before, immediately after and three months after the end of the training. We ran a mixed ANOVA for each measure with Session (Pre-test vs. Post-test vs. Follow-up) as within-subject factor and Group (PG vs. EG) as between-subject factor.

Mauchly's Test of Sphericity was run to check the equality of the variances of the differences between all the levels. When the assumption of Sphericity was violated, the Greenhouse-Geisser adjustment was applied to *p*-values (reported as $p_{[gg]}$). Post-hoc *t*-tests were two-tailed and the *p*-values were corrected for multiple comparisons using the Bonferroni criterion (i.e., alpha value divided by the number of tests).

Bayesian analyses were run to quantify the evidence for both the alternative hypothesis (H1) and the null hypothesis (H0). We reported Bayes factors (BF_{10}) expressing the probability of the data given H1 relative to H0, where values larger than 1 are in favour of H1 whereas values smaller than 1 are in favour of H0 (for example, a $BF_{10} = 10$ indicates that the data are 10 times more likely to occur under H1 compared to H0). In the results we reported the evidence associated with BFs as "anecdotal" (BF < 3), "moderate" (BF > 3), "strong" (BF > 10), "very strong" (BF > 30), or "extreme" (BF > 100) (Jeffreys, 1961). Bayes factors are also used to indicate the relative evidence between competing models: in particular, the ratio of BF_{10} for the model including the interaction Group X Session to the BF_{10} for the model only including the main effects (Group and Session) is important to ascertain whether the effect of intervention differs between groups. Here a ratio smaller than 1 is evidence for a lack of interaction, thereby suggesting similar training effects in

Table 2
Descriptive Statistics for EG and PG.

	EG ($n = 20$)			PG ($n = 21$)		
	Pre-test	Post-test	Follow-up	Pre-test	Post-test	Follow-up
BIN Total score	67.64	79.95	80.71	63.66	75.25	75.65
Correct answers (%)	(21.94)	(16.93)	(16.94)	(25.13)	(23.75)	(23.78)
	[29.25−99.06]	[50−100]	[52.83−99.06]	[18.87−99.06]	[21.70−98.11]	[30.19−100]
BIN Lexical Subscale	82.39	90.65	90.87	87.79	90.89	90.68
	(20.07)	(12.40)	(12.21)	(19.86)	(17.33)	(17.30)
	[43.48−100]	[69.57−100]	[65.22−100]	[21.74−100]	[30.44−100]	[34.68−100]
BIN Semantic Subscale	73.81	86.19	86.67	71.20	87.53	86.40
	(15.26)	(9.88)	(11.62)	(17.59)	(15.54)	(17.58)
	[42.86−100]	[66.67−100]	[66.67−100]	[38.10−100]	[57.14−100]	[38.10−100]
BIN Counting Subscale	59.88	73.75	74.13	54.64	67.86	68.57
	(31.47)	(26.81)	(25.54)	(38.63)	(35.58)	(32.62)
	[20.00−100]	[32.50−100]	[37.50−100]	[0−100]	[0−100]	[7.50−100]
BIN Pre-syntactic Subscale	60.46	74.09	76.36	47.62	60.61	62.55
	(23.28)	(20.76)	(19.26)	(26.95)	(25.40)	(27.57)
	[18.18−100]	[31.82−100]	[31.82−95.46]	[9.09−95.46]	[13.64−90.91]	[22.73−100]
Number Comparison	29.09	14.55	14.09	43.29	28.14	19.48
Errors (%)	(18.56)	(15.16)	(14.59)	(21.69)	(22.44)	(24.94)
	[0−63.64]	[0−45.46]	[0−45.46]	[0−100]	[0−72.73]	[0−100]
NTP 0−10	20.43	17.13	15.54	26.58	22.50	25.00
Errors (%)	(9.46)	(10.55)	(11.16)	(16.28)	(17.59)	(15.79)
	[7.53−44.44]	[4.33−42.44]	[3.06−50.88]	[7.51−59.60]	[5.73−63.89]	[5.67−50.00]
NTP 0−20	22.33	16.56	14.65	28.05	24.60	25.91
Errors (%)	(11.48)	(11.09)	(9.76)	(16.13)	(15.65)	(16.02)
	[2.44−48.19]	[3.28−44.21]	[5.88−47.64]	[7.26−51.97]	[4.87−51.97]	[3.65−51.97]
Mental Calculation	23.81[a]	47.22[a]	50.00[a]	27.49[a]	28.95[a]	30.99[a]
Accuracy (%)	(32.79)	(43.19)	(42.98)	(32.58)	(29.78)	(31.50)
	[0−94.44]	[0−100]	[0−100]	[0−94.44]	[0−88.89]	[0−94.44]

Note. Means (SD) for each outcome variable are reported. EG = Expert group; PG = Parent group.
[a] Due to a computer failure, the sample size was reduced for *Mental Calculation* ($N_{EG} = 14$; $N_{PG} = 19$).

expert- and parent-led groups.

Effect size values were computed to analyse and compare improvements of both EG and PG, in comparison to the CG. These values were calculated as follows: the pre–post change in EG or PG means minus the pre–post change in the CG means, divided by the pooled pre-test standard deviation with a bias adjustment (Carlson & Schmidt, 1999; Morris, 2008). The same computation was applied to estimate effect size of pre–follow change and post–follow change (in the latter case, pooled post-test standard deviation was used)

Table 3
ANOVA results for each task are reported in terms of main effects of Session (Pre-test, Post-test and Follow up), Group (Expert Group vs Parent Group) and Session x Group interaction. The last column shows the ratio of Bayes Factor (BF) for the model including the interaction to the BF for the model only including the main effects; values smaller than 1 indicate evidence for a lack of interaction (i.e., similar effects in the two groups).

Task	Main Effect of Session	Main Effect of Group	Interaction Session*Group	BF_{10} Group+Session+Session*Group / BF_{10} Group+Session
BIN Total Score	F(1.35, 52.65)=57.72 $p_{[gg]}$<0.001 η^2_p=0.60 BF_{10}= 2.03x10^{16}	F(1, 39)=0.48 p=0.49 η^2_p=0.01 BF_{10}= 0.60	F(1.35, 52.65)=0.09 $p_{[gg]}$=0.84, η^2_p=0.002 BF_{10}=2.01x10^{12}	BF_{10}=0.14
	Group+Session: BF_{10}=1.43x10^{13}			
BIN - Lexical	F(1.58, 61.66)=16.48 $p_{[gg]}$<0.001 η^2_p=0.30 BF_{10}= 5748.40	F(1, 39)=0.13 p=0.72 η^2_p=0.003 BF_{10}= 0.57	F(1.58, 61.66)=3.69 $p_{[gg]}$=0.04 η^2_p=0.09 BF_{10}=6822.20	BF_{10}= 1.93
	Group+Session: BF_{10}=3526.43			
BIN - Semantic	F(1.68, 65.46)=31.03 $p_{[gg]}$<0.001 η^2_p=0.44 BF_{10}= 1.10x10^{8}	F(1, 39)=0.20 p=0.90 η^2_p<0.001 BF_{10}= 1.42	F(1.68, 65.46)=0.46 $p_{[gg]}$=0.60 η^2_p=0.01 BF_{10}=7.68x10^{6}	BF_{10}=0.18
	Group+Session: BF_{10}=4.23x10^{7}			
BIN - Counting	F(1.34, 52.25)=23.24 $p_{[gg]}$<0.001 η^2_p=0.37 BF_{10}= 1.29x10^{6}	F(1, 39)=0.33 p=0.57 η^2_p=0.008 BF_{10}= 0.58	F(1.34, 52.25)=0.01 $p_{[gg]}$=0.96 η^2_p<0.001 BF_{10}=1.12 x10^{5}	BF_{10}=0.14
	Group+Session: BF_{10}=8.25x10^{5}			
BIN - Pre-syntactic	F(1.72, 67.23)=29.23 $p_{[gg]}$<0.001 η^2_p=0.43 BF_{10}= 3.94x10^{7}	F(1, 39)=3.54 p=0.07 η^2_p=0.08 BF_{10}= 1.91	F(1.72, 67.23)=0.03 $p_{[gg]}$=0.96 η^2_p=0.001 BF_{10}=7.10x10^{6}	BF_{10}=0.13
	Group+Session: BF_{10}=5.32x10^{7}			
Number Comparison	F(2, 78 = 23.57) p < 0.001 η^2_p= 0.38 BF_{10}= 1.23x10^{6}	F(1,39) = 4.46 p = 0.04, η^2_p = 0.10 BF_{10}= 1.72	F(2, 78) = 1.39 p = 0.26 η^2_p = 0.03 BF_{10}=8.43x10^{5}	BF_{10}=0.37
	Group+Session: BF_{10}=2.29x10^{6}			
NTP 1-10	F(2,78) = 2.23 p =0.11 η^2_p= 0.05 BF_{10}= 3.73	F(1,39) = 3.31 p = 0.08 η^2_p = 0.08 BF_{10}= 2.21	F(2,78) = 0.31 p = 0.73 η^2_p = 0.01 BF_{10}=3.52	BF_{10}=0.43
	Group+Session: BF_{10}=8.24			
NTP 1-20	F(2,78) = 4.97 p =0.01 η^2_p= 0.11 BF_{10}= 2.98	F(1,39) = 4.91 p = 0.03 η^2_p = 0.11 BF_{10}= 1.90	F(2,78) = 1.27 p = 0.29 η^2_p = 0.03 BF_{10}=2.24	BF_{10}=0.38
	Group+Session: BF_{10}=5.95			
Mental Calculation	F(1.57, 48.55) = 7.23 $p_{[gg]}$ = 0.004 η^2_p= 0.20 BF_{10}= 4.64	F(1,31) = 0.96 p = 0.33 η^2_p = 0.03 BF_{10}= 0.64	F(1.57, 48.55) = 4.73 $p_{[gg]}$ = 0.02, η^2_p = 0.13 BF_{10}= 10.97	BF_{10}=3.56
	Group+Session: BF_{10}= 3.08			

(Fig. 1). This approach, which compares changes across groups between times (i.e. pre vs. post, pre vs. follow, post vs. follow), was chosen because it gives increased precision on estimates of treatment effects and statistically accounts for any pre-intervention differences between groups (Morris, 2008).

3. Results

Descriptive statistics are reported in Table 2. The sample size for the analysis on *Mental Calculation* task was 19 for PG and 14 for EG due to a computer failure. In Table 3 ANOVA results for all the tasks are reported. Finally, effect size values comparing gains of EG and PG against the CG for each task are reported in Fig. 1.

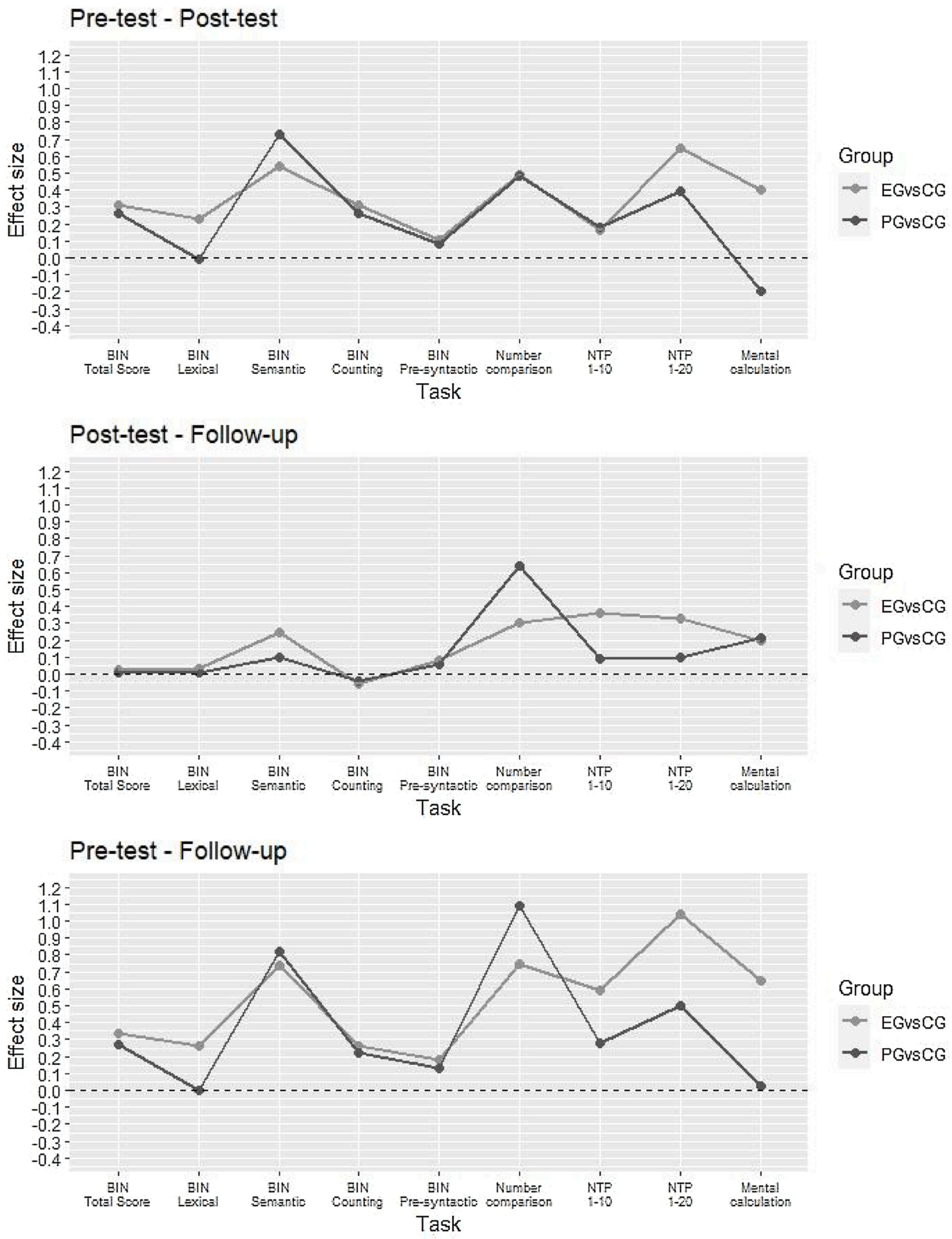
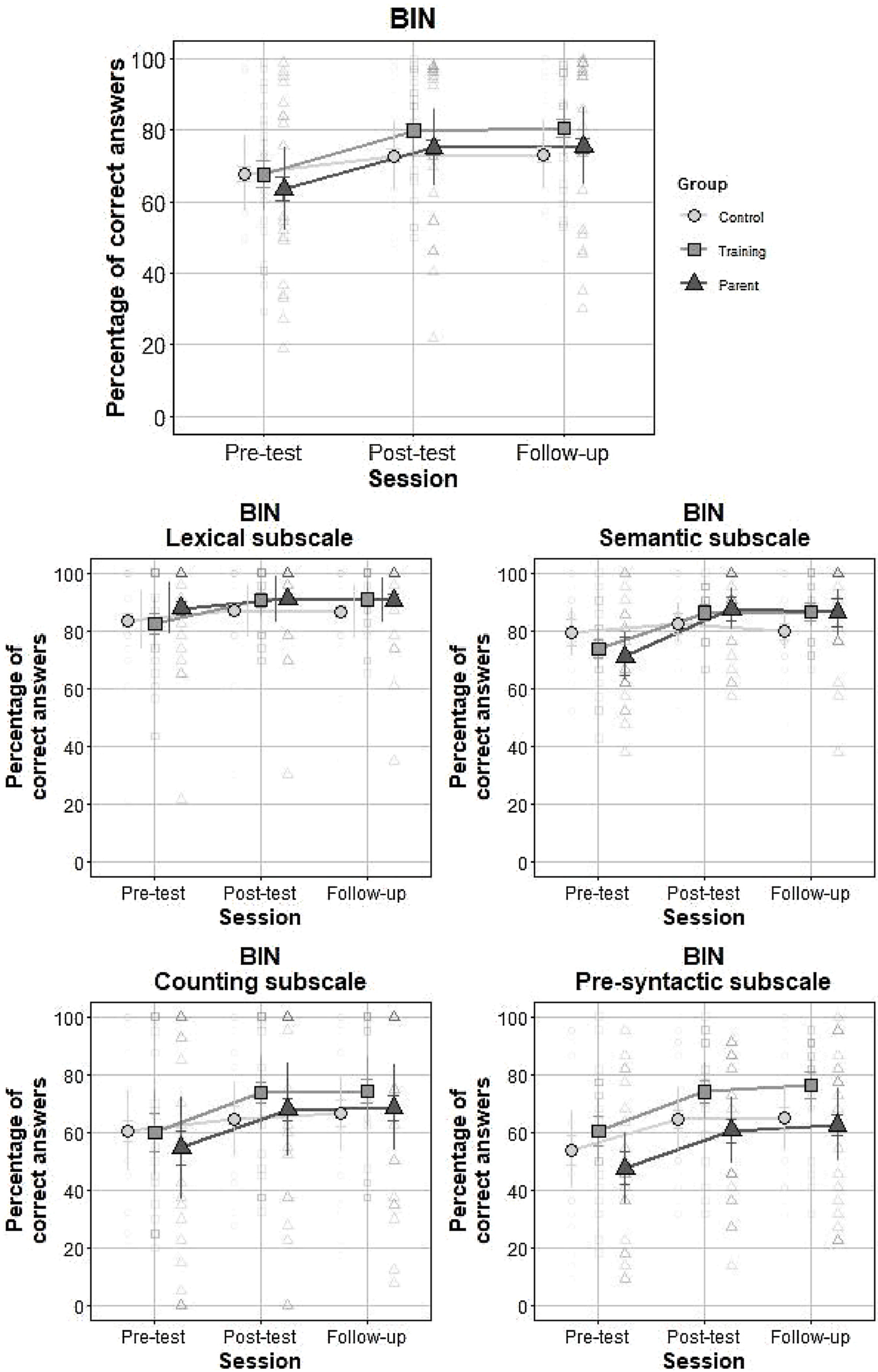


Fig. 1. Effect size values comparing gains of EG and PG against CG for each task. In order to simplify graph reading, effect size values of tasks assessing percentage of errors have been reversed, so that for all the tasks positive values represent improvement of EG and PG against CG.



(caption on next page)

Fig. 2. Mean BIN total and subscales percentage scores of PG, EG and CG and distribution of individual scores before, immediately after and three months after the end of the training. Error bars represent between-groups and within-groups (horizontal segments) 95 % confidence intervals. Transparent points represent individual scores.

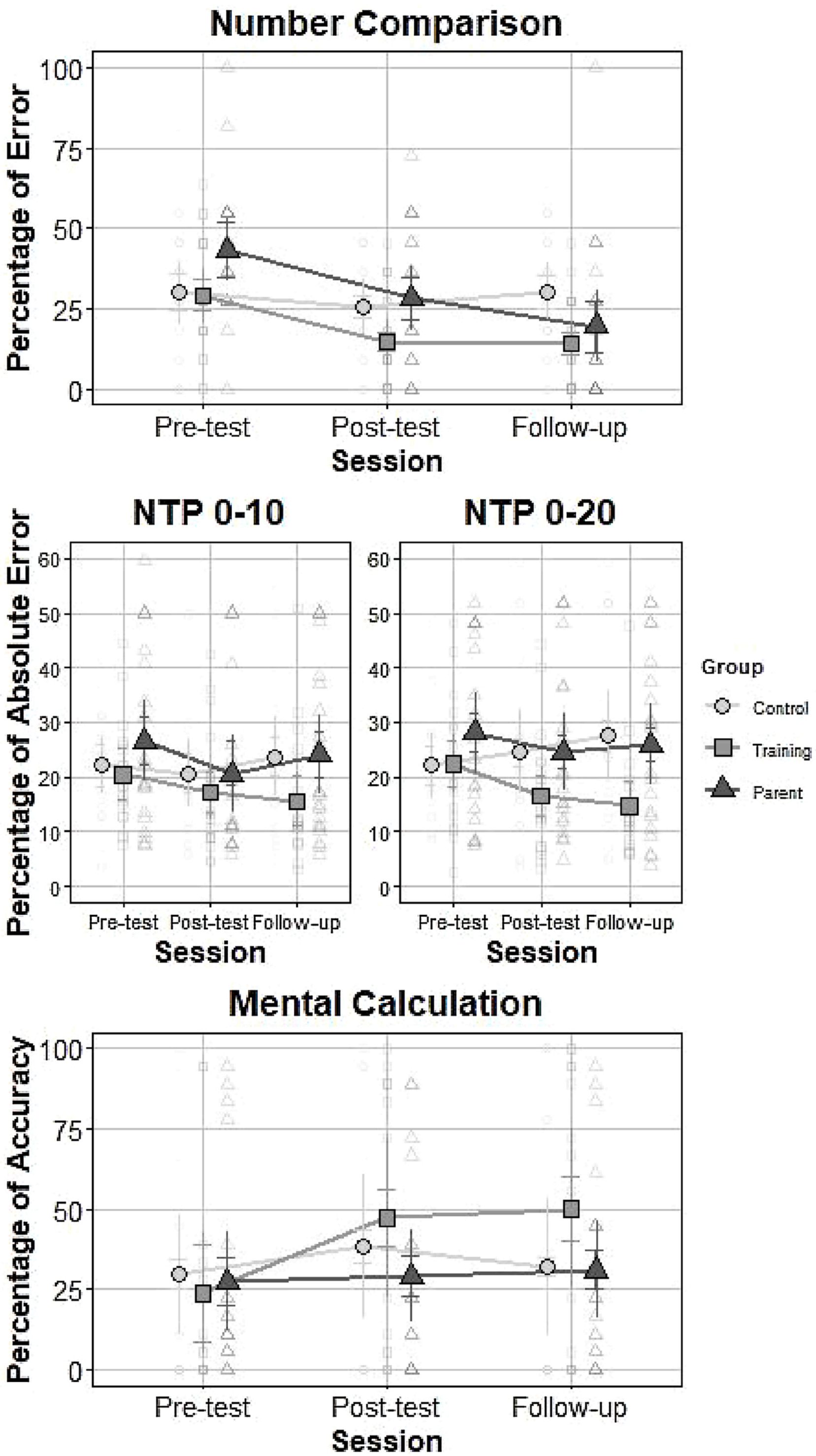


Fig. 3. Mean scores in Number Comparison, Number-to-position and Mental Calculation tasks of EG and CG across sessions and distribution of individual scores before, immediately after and three months after the end of the training. Error bars represent between- and within- groups (horizontal segments) 95 % confidence intervals. Transparent points represent individual scores.

3.1. BIN

The 3 × 2 ANOVA on the total numerical intelligence score showed a main effect of Session (see Fig. 2). Overall, participants' scores improved from pre-test to post-test ($MDiff_{pre\text{-}post}$ =-11.94, $p < 0.001$, $BF_{10} = 2.04 \times 10^7$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =-12.52, $p < 0.001$, $BF_{10} = 1.87 \times 10^7$) but no significant difference was present between post-test and follow-up, indicating a maintenance effect. No main effect of Group was found. Moreover, there was no evidence ($BF_{10} = 0.14$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. Effect size values comparing gains of EG and PG against CG showed for both groups a medium effect of the training (respectively 0.32 for EG and 0.27 for PG) between pre-test and post-test that was maintained at follow-up.

We then conducted a series of 3 × 2 ANOVAs considering the percentage of correct answers of each *BIN* subscale (see Fig. 2). In the *Lexical* subscale, participants' performance improved from pre-test to post-test ($MDiff_{pre\text{-}post}$ =-5.62, $p < 0.001$, $BF_{10} = 225.51$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =-5.62, $p < 0.001$, $BF_{10} = 143.01$) but no significant difference was present between post-test and follow-up. There was no main effect of Group. An anecdotal evidence ($BF_{10} = 1.93$) emerged in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. *T*-test between-groups comparisons, with Bonferroni adjustment to alpha level of 0.05/3 = 0.016, revealed that PG did not differ from EG at pre-test, post-test and follow-up. The *t*-test within-group comparisons showed that, considering Bonferroni adjustment to alpha level of 0.008 (i.e., 0.05/6), EG improved from pre-test to post-test ($t(19) = 3.67$, $p = 0.002$, $g = 0.49$, $BF_{10} = 24.01$) and from pre-test to follow-up $t(19) = 4.17$, $p < 0.001$, $g = 0.50$, $BF_{10} = 65.26$), while PG did not show improvements between sessions. Accordingly, the effect size values showed a medium effect of the training for EG (0.23) with respect to CG that was maintained at follow-up. No effect was found for the PG (-0.01).

For the *Semantic* subscale, there was a main effect of Session: participants' performance improved from pre-test to post-test ($MDiff_{pre\text{-}post}$ =−14.40, $p < 0.001$, $BF_{10} = 2.79 \times 10^5$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =-14.05, $p < 0.001$, $BF_{10} = 2.55 \times 10^4$) but no significant difference was present between post-test and follow-up. No main effect of Group was found and there was no evidence ($BF_{10} = 0.18$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. Effect size values showed for both groups a greater gain with respect to CG, with a medium effect (0.55) for the EG and a large effect (0.73) for the PG. The gains were maintained at follow up.

The *Counting* subscale presented again a main effect of Session. Participants' performance improved from pre-test to post-test ($MDiff_{pre\text{-}post} = -13.54$, $p < 0.001$, $BF_{10} = 3496.91$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow} = -14.09$, $p < 0.001$, $BF_{10} = 2100.54$) but no significant difference was present between post-test and follow-up. There was no main effect of Group. There was no evidence ($BF_{10} = 0.14$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. For both groups effects sizes showed medium effect of the training (respectively 0.31 for EG and 0.26 for PG) with respect to the CG at post-test that was maintained at follow-up.

For the *Pre-syntactic* subscale, a main effect of Session emerged: participants' performance improved from pre-test to post-test ($MDiff_{pre\text{-}post} = -13.30$, $p < 0.001$, $BF_{10} = 1.27 \times 10^4$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow} = -15.41$, $p < 0.001$, $BF_{10} = 1.38 \times 10^5$) but no significant difference emerged between post-test and follow-up. There was no main effect of Group. There was no evidence ($BF_{10} = 0.13$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. For both groups the training had a small effect (respectively 0.11 for EG and 0.08 for PG) at post test with respect to the CG that was maintained at follow-up.

3.2. Number Comparison

A 3 × 2 ANOVA was conducted on the percentage of errors (see Fig. 3). A main effect of Session was found. Participants' errors decreased in all the session: from pre-test to post-test ($MDiff_{pre\text{-}post}$ =14.86, $p < 0.001$, $BF_{10} = 2616.14$), from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =19.51, $p < 0.001$, $BF_{10} = 1.35 \times 10^4$) but no significant difference emerged between post-test and follow-up. The main effect of Group was also significant: the percentage of errors of EG was lower than the percentage of errors of PG ($M_{EG} = 19.38$, $M_{PG} = 30.44$, $p = 0.005$, $BF_{10} = 7.13$). There was no evidence ($BF_{10} = 0.37$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. In both groups effect sizes showed a medium effect (0.49 for both EG and PG) of the training at post test. A medium effect size with respect to the CG was also found for the EG (0.30) from post-test to follow-up, while the effect size was large for the PG (0.64).

3.3. Number-To-Position (NTP)

To investigate the effect of training on number estimation, the percentage of absolute error in the *NTP* task was analyzed with a 3 × 2 ANOVA (see Fig. 3). Analyses were run separately for the 0−10 and 0−20 interval.

For the 0−10 interval the ANOVA showed neither effect of Session nor of Group. Moreover, there was no evidence ($BF_{10} = 0.43$) in favor of the model with the two mains effects and the interaction compared to the model with only the two main effects.

For the 0−20 interval the ANOVA showed a main effect of Session. Participants' accuracy improved from pre-test to post-test ($MDiff_{pre\text{-}post}$=-4.58, $p = 0.01$, $BF_{10} = 4.30$), and from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =-4.84, $p = 0.02$, $BF_{10} = 2.67$) but no significant difference emerged between post-test and follow-up. A main effect of Group was also found: the percentage of absolute error of EG was lower than the percentage of absolute error of PG ($M_{EG} = 17.95$, $M_{PG} = 26.29$, $p < 0.001$, $BF_{10} = 29.31$). There was no evidence ($BF_{10} = 0.38$) in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. A medium effect of the training with respect to the CG was found in both groups at post-test (respectively 0.65 for EG and

0.39 for PG). For the EG a medium effect (0.33) indicated a further improvement from post-test to follow-up; in the PG the effect was small (0.10).

3.4. Mental Calculation

To examine the effect of training on arithmetic proficiency, the percentage of accuracy on the *Mental Calculation* task was considered in a 3 × 2 ANOVA (see Fig. 3) that showed no main effect of Group, while a main effect of Session was present. Participants' performance increased from pre-test to post-test ($MDiff_{pre\text{-}post}$=-10.77, $p = 0.04$, $BF_{10} = 1.49$) and from pre-test to follow-up ($MDiff_{pre\text{-}follow}$ =-13.13, $p = 0.01$, $BF_{10} = 3.35$).

Moderate evidence ($BF_{10} = 3.56$) emerged in favor of the model with the two main effects and the interaction compared to the model with only the two main effects. The *T*-test between groups comparisons, with Bonferroni adjustment to alpha level of 0.05/3 = 0.016, revealed that PG did not differ from EG at pre-test, post-test and follow-up

The *t*-test within group comparisons showed that, considering Bonferroni adjustment to alpha level of 0.008 (i.e., 0.05/6), improvements between sessions of both groups did not reach statistical significance (all $p > 0.008$). However, effect size values comparing gains of the two training groups against the CG showed a medium effect of the training from pre to post test in the EG (0.41) that was maintained at follow up, while a negligible effect was found in the PG (-0.20).

4. Discussion

The aim of the present work was to assess the feasibility of improving basic number skills in individuals with DS through an intervention program delivered at home by parents. Findings from the literature suggest that cognitive training of individuals with intellectual disabilities needs several sessions per week (e.g. Buckley, 2008), and should last longer (e.g. Ottersen & Grill, 2015) to be effective. However, it is often the case that parents find it difficult to involve their children in specific training programs due to lack of time or opportunity to regularly visit a specialised center. In a previous study, we demonstrated the efficacy of *The Number Race*, an adaptive computer game, in improving basic number skills in individuals with DS, working under the supervision of an expert in psychology (Sella et al., 2021). In the present work, we compared the data collected in the previous study with those of a group of children with DS that has been trained at home by parents with the same schedule.

First of all, the results of the present study confirmed the efficacy of *The Number Race* in improving basic number skills in atypically developing children (Wilson, Dehaene, Dubois, & Fayol, 2009). Indeed, both groups (i.e., the one trained by an expert and the one trained by parents) improved their performance after the training not only in specific tasks directly trained by the software, such as number comparison, but also in a more general measure of numerical intelligence (the *BIN* total score) that taps areas not directly exercised with the software. In both groups the gain was greater than the one seen in an active control group working on reading skills, whose results were described in the previous study. Moreover, in both groups the improvement was largely maintained after three months, thereby showing a long-lasting effect.

It is interesting to note that in the ability directly trained by the software, *Number Comparison,* both groups improved with respect to the CG not only from pre-test to post-test, but also from post-test to follow-up. It seems that playing with the software helps the child to pay attention to quantities, and this attitude continues to develop even outside the context of the training activities.

Concerning the *Number-to-Position* tasks, a decrement of the absolute percentage of error was seen only for the 1−20 number line, but not for the 1−10 one. This is probably due to the fact that the performance of individuals with DS aged between 6 and 16 years on the 1−10 line is already in line with mental age, if not better, with a linear mapping of numbers, as previously shown by Lanfranchi et al. (2015).

Previous studies demonstrated the link between basic number skills and later math achievement in typically developing children (e.g. Halberda et al., 2008; Jordan et al., 2006), and in children with dyscalculia (Piazza et al., 2010; Sella et al., 2014). Moreover, the link between basic number skills and many activities of everyday life, fundamental for adaptive behavior, such as handling money or time, is well known. For these reasons, the results of our study, suggesting the feasibility of training basic number skills in individuals with DS, are particularly important considering the specific deficit showed by this population in small number comparison, counting and mental calculation (see Onnivello et al., 2019, for a review).

Secondly, our results indicate that parents were able to effectively administer the training to their children, who improved their performance in numerical skills, number estimation and number comparison after playing with *The Number Race*. This seems to be in line with previous studies that supported the possibility for parents of directly training their children, under the supervision of an expert, in other cognitive domains, such as verbal (Conners et al., 2008) or visuo-spatial working memory (Pulina et al., 2015).

The only area where the improvement was negligible in children trained by parents, compared to a moderate effect for the group trained by the expert, was *Mental Calculation*. In this regard, it should be noted that *The Number Race* is an adaptive videogame and most of the children in the PG, differently from the EG, did not reach the most difficult game levels that engage mental calculation. Accordingly, our results suggest that a longer period of intervention would better suit the needs of children with DS and it would increase the probability that every child reaches the game levels that exercise also the ability to perform mental calculation.

In summary, our results confirm the efficacy of *The Number Race* to improve basic number skills in individuals with DS. Notwithstanding the methodological limitations (e.g., non-random allocation and lack of blinding; see Sella et al., 2021 for a thorough discussion) the present study suggests the possibility to use this software not only under the supervision of an expert, but also at home under the supervision of parents (with adequate support and instruction). We believe that the slightly lower efficacy of the program when administered under the supervision of parents might be compensated by more focused instructions and supervision to parents together with a longer duration of the training (which is much easier to pursue in a home-based setting). Moreover, a home-based intervention may change caregivers' attitude toward mathematics and the confidence they have in teaching the subject to their children.

Our results have important practical significance because they open the possibility of a broader use of *The Number Race* to train basic number skills in children with DS. On one side, working at home permits a large number of children to have access to the training program; on the other side, a more frequent or a periodic use of the program would be possible in order to enhance the effect of the training.

Funding

This work was supported by a grant from Fondation Jérôme Lejeune (#1595) to M.Z.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

We are grateful to all the children who took part in this study and their families.

We thank A.Gen.Do onlus Vicenza, AIPD Venezia Mestre and Associazione Down Dadi Padova for their support in contacting families. We thank Benedetta De Santi for her help in the data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ridd.2021.103919.

References

Abreu-Mendoza, R. A., & Arias-Trejo, N. (2015). Numerical and area comparison abilities in Down syndrome. *Research in Developmental Disabilities, 41–42*, 58–65. https://doi.org/10.1016/j.ridd.2015.05.008.

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development, 54*, 695–701. https://doi.org/10.2307/1130057.

Bashash, L., Outhred, L., & Bochner, S. (2003). Counting skills and number concepts of students with moderate intellectual disabilities. *International Journal of Disability, Development, and Education, 50*(3), 325–345. https://doi.org/10.1080/1034912032000120480.

Belacchi, C., Scalisi, T. G., Cannoni, E., & Cornoldi, C. (2008). *CPM - coloured progressive matrices. Standardizzazione italiana*. Firenze: Giunti O.S.

Berteletti, I., De Filippo De Grazia, F., & Zorzi, M. (2012). "La Gara dei Numeri": Un videogioco educativo per il potenziamento delle abilità numeriche ed il trattamento della discalculia. *Difficoltà in Matematica, 9*.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*(2), 545–551. https://doi.org/10.1037/a0017887.

Block, M. E. (1991). Motor development in children with Down syndrome: A review of the literature. *Adapted Physical Activity Quarterly, 8*, 179–209. https://doi.org/10.1123/apaq.8.3.179.

Buckley, S. (2008). It is time to take memory training seriously. *Down Syndrome Research and Practice, 12*(2), 105–106.

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*, 6033. https://doi.org/10.1126/science.1201536, 1049-60331053.

Camos, V. (2009). Numerosity discrimination in children with Down syndrome. *Developmental Neuropsychology, 34*(4), 435–447. https://doi.org/10.1080/87565640902964557.

Carlson, K. D., & Schmidt, F. L. (1999). Impact of experimental design on effect size: Findings from the research literature on training. *The Journal of Applied Psychology, 84*, 851–862. https://doi.org/10.1037/0021-9010.84.6.851.

Caycho, L., Gunn, P., & Siegal, M. (1991). Counting by children with Down syndrome. *American Journal on Mental Retardation, 95*(5), 575–583.

Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 84–95. https://doi.org/10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P.

Conners, F. A., Rosenquist, C. J., Arnett, L., Moore, M. S., & Hume, L. E. (2008). Improving memory span in children with Down syndrome. *Journal of Intellectual Disability Research, 52*(3), 244–255. https://doi.org/10.1111/j.1365-2788.2007.01015.x.

Cornwell, A. C. (1974). Development of language, abstraction, and numerical concept formation in Down's syndrome children. *American Journal of Mental Deficiency, 79*(2), 179–190.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test–III*. Circle Pines, MN: American Guidance Service.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002.

Fidler, D. J. (2005). The emerging Down syndrome behavioral phenotype in early childhood. *Infants and Young Children, 18*(2), 86–103.

Gelman, R., & Cohen, M. (1988). Qualitative differences in the way Down syndrome and normal children solve a novel counting problem. In L. Nadel (Ed.), *Issues in the biology of language and cognition. The psychobiology of Down syndrome* (pp. 51–99). Cambridge, Massachusetts: The MIT Press.

Grieco, J., Pulsifer, M., Seligsohn, K., Skotko, B., & Schwartz, A. (2015). Down syndrome: Cognitive and behavioral functioning across the lifespan. *American Journal of Medical Genetics Part 6 (Seminar in Medical Genetics), 169C*, 135–149. https://doi.org/10.1002/ajmg.c.31439.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*(5), 1457. https://doi.org/10.1037/a0012682.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665–668. https://doi.org/10.1038/nature07246.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*(1), 92–107. https://doi.org/10.1016/j.cognition.2013.12.007.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (1999). Genetically dissociated components of working memory: Evidence from Down's and Williams syndrome. *Neuropsychologia, 37*, 637–651. https://doi.org/10.1016/S0028-3932(98)00128-6.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Jordan, N. C., Kaplan, D., Oláh, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*(1), 153–175. https://doi.org/10.1111/j.1467-8624.2006.00862.x.

Lanfranchi, S., Baddeley, A., Gathercole, S., & Vianello, R. (2012). Working memory in Down syndrome: Is there a dual task deficit? *Journal of Intellectual Disability Research, 56*(2), 157–166. https://doi.org/10.1111/j.1365-2788.2011.01444.x.

Lanfranchi, S., Berteletti, I., Torrisi, E., Vianello, R., & Zorzi, M. (2015). Numerical estimation in individuals with Down syndrome. *Research in Developmental Disabilities, 36*, 222–229. https://doi.org/10.1016/j.ridd.2014.10.010.

Lanfranchi, S., Cornoldi, C., & Vianello, R. (2004). Verbal and visuospatial working memory deficits in children with Down syndrome. *American Journal on Mental Retardation, 109*(6), 456–466. https://doi.org/10.1352/0895-8017(2004)109<456:VAVWMD>2.0.CO;2.

Lanfranchi, S., Jerman, O., Dal Pont, E., Alberti, A., & Vianello, R. (2010). Executive function in adolescents with Down syndrome. *Journal of Intellectual Disability Research, 54*(4), 308–319. https://doi.org/10.1111/j.1365-2788.2010.01262.x.

Lee, N. R., Fidler, D. J., Blakeley-Smith, A., Daunhauer, L., Robinson, C., & Hepburn, S. L. (2011). Caregiver report of executive functioning in a population-based sample of young children with Down syndrome. *American Association on Intellectual and Developmental Disabilities, 116*(4), 290–304. https://doi.org/10.1352/1944-7558-116.4.290.

Marotta, L., Viezzoli, D., & Vicari, S. (2006). Le abilità numeriche e di calcolo nella Disabilità Intellettiva. *Età Evolutiva, 83*, 71–78.

McGrother, C. W., & Marshall, B. (1990). Recent trends in incidence, morbidity and survival in Down's syndrome. *Journal of Mental Deficiency Research, 34*, 49–57. https://doi.org/10.1111/j.1365-2788.1990.tb01514.x.

Molin, A., Poli, S., & Lucangeli, D. (2007). *BIN 4-6: Batteria per la valutazione dell'intelligenza numerica in bambini dai 4 ai 6 anni*. Trento: Erickson.

Morris, S. B. (2008). Estimating effect size from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059.

Mou, Y., & Van Marle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review, 34*(1), 1–25. https://doi.org/10.1016/j.dr.2013.11.001.

Onnivello, S., Lanfranchi, S., & Zorzi, M. (2019). Mathematical abilities in Down syndrome. In S. Lanfranchi (Ed.), *International review of research in developmental disabilities. State of the art of research on Down syndrome* (pp. 257–291). London, UK: Academic Press. https://doi.org/10.1016/bs.irrdd.2019.06.005.

Ottersen, J., & Grill, K. M. (2015). Benefits of extending and adjusting the level of difficulty on computerized cognitive training for children with intellectual disabilities. *Frontiers in Psychology, 6*, 1233. https://doi.org/10.3389/fpsyg.2015.01233.

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science, 24*, 2013–2019. https://doi.org/10.1177/0956797613482944.

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47*(2), 190–204. https://doi.org/10.1111/j.1469-7610.2005.01460.x.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012.

Pulina, F., Carretti, B., Lanfranchi, S., & Mammarella, I. (2015). Improving spatial-simultaneous working memory in Down syndrome: Effect of a training program led by parents instead of an expert. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01265.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Coloured progressive matrices*. Oxford: Oxford University Press.

Sella, F., Hartwright, C., & Cohen Kadosh, R. (2018). The neurocognitive bases of numerical cognition. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, 3*, 1–47. https://doi.org/10.1002/9781119170174.epcn316.

Sella, F., Berteletti, I., Brazzolotto, M., Lucangeli, D., & Zorzi, M. (2014). Number line estimation in children with developmental dyscalculia. *Learning Disabilities: A Contemporary Journal, 11*(2), 41–49.

Sella, F., Lanfranchi, S., & Zorzi, M. (2013). Enumeration skills in Down syndrome. *Research in Developmental Disabilities, 34*(11), 3798–3806. https://doi.org/10.1016/j.ridd.2013.07.038.

Sella, F., Onnivello, S., Lunardon, M., Lanfranchi, S., & Zorzi, M. (2021). Training basic numerical skills in children with Down syndrome: a randomised controlled trial using the computerized game "The Number Race". *Scientific Reports, 11*, 2087. https://doi.org/10.1038/s41598-020-78801-5.

Sella, F., Tressoldi, P., Lucangeli, D., & Zorzi, M. (2016). Training numerical skills with the adaptive videogame "The Number Race": A randomized controlled trial on preschoolers. *Trends in Neuroscience and Education, 5*(1), 20–29. https://doi.org/10.1016/j.tine.2016.02.002.

Sherman, S. L., Allen, E. G., Bean, L. H., & Freeman, S. B. (2007). Epidemiology of Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 13*, 221–227. https://doi.org/10.1002/mrdd.20157.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–243. https://doi.org/10.1111/1467-9280.02438.

Simms, V., Karmiloff-Smith, A., Ranzato, E., & Van Herwegen, J. (2020). Understanding number line estimation in Williams syndrome and Down syndrome. *Journal of Autism and Developmental Disorders, 50*(2), 583–591. https://doi.org/10.1007/s10803-019-04268-7.

Stella, G., Pizzioli, C., & Tressoldi, P. E. (2000). *PPVT-R, peabody picture vocabulary test-revised*. Torino: Omega Edizioni.

Vicari, S., Marotta, L., & Carlesimo, G. A. (2004). Verbal short-term memory in Down's syndrome: an articulatory loop deficit? *Journal of Intellectual Disability Research, 48*(2), 80–92. https://doi.org/10.1111/j.1365-2788.2004.00478.x.

Vygotsky, L. S. (1976). Play and its role in the development of the child. In J. S. Bruner, A. Jolly, & K. Sylva (Eds.), *Play—Its role in the development of evolution* (pp. 537–554). Middlesex: Penguin Books.

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-Status kindergarten children. *Mind Brain and Education, 3*(4). https://doi.org/10.1111/j.1751-228X.2009.01075.x.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*(19). https://doi.org/10.1186/1744-9081-2-19.

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "the Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*(20). https://doi.org/10.1186/1744-9081-2-20.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*, 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x.

Yang, Y., Conners, F. A., & Merrill, E. C. (2014). Visuo-spatial ability in individuals with Down syndrome: Is it really a strength? *Research in Developmental Disability, 35*, 1473–1500. https://doi.org/10.1016/j.ridd.2014.04.002.