nature human behaviour

ARTICLES

https://doi.org/10.1038/s41562-021-01254-x

Check for updates

Enhancing reading skills through a video game mixing action mechanics and cognitive training

Angela Pasqualotto [1,2] ✉, Irene Altarelli [3], Antonella De Angeli [4,5], Zeno Menestrina [4], Daphne Bavelier [2,6] ✉ and Paola Venuti [1,6]

In modern societies, training reading skills is fundamental since poor-reading children are at high risk of struggling both at school and in life. Reading relies not only on oral language abilities but also on several executive functions. Considering their importance for literacy, training executive functions—particularly, attentional control has been suggested as a promising way of improving reading skills. For this reason, we developed a video game-based cognitive intervention aimed at improving several facets of executive functions. This game is composed of mini-games that apply gamified versions of standard clinical exercises linked through a game environment with action video game dynamics. Here, in a study involving 151 typically reading children, we demonstrated that after this general-domain behavioural intervention reading abilities, as well as attentional and planning skills, were significantly improved. Our results showed that training attentional control can translate into better reading efficiency, maintained at a follow-up test 6 months later.

Literacy is an essential requisite in modern societies. Poor readers, that is children encountering difficulties with decoding and/or understanding print material, face a disproportionately higher risk of struggling at school and in life[1–3]. Thus, developing proper training strategies for reading is a goal of upmost importance.

Several models of reading acquisition have emerged in the course of the last decades, highlighting the importance of oral language skills and especially of phonological awareness, for the appropriate development of reading in childhood[4–7]. On top of phonological awareness skills—that is, the ability to identify and manipulate subunits within words, such as syllables and phonemes—other, more domain-general abilities have also been related to reading acquisition. Among these, attentional control—flexibly distributing attention on what is task-relevant and filtering out distractors—appears to be key at several levels. Attention needs to be correctly distributed on page to successfully decode text, as a printed page is typically quite crowded. In this context, attentional control is especially important to group letters into words and phrases as well as swiftly move from one line to another; this is especially so before perceptual expertise for printed material emerges[8–10]. Attentional control is also essential for suppressing close phonological competitors or orthographic competitors and flexibly switching between different types of information to be extracted from the same string of letters (that is, phonological, orthographic or semantic information)[11,12]. Interestingly, attentional skills are not only related to concurrent reading abilities during childhood but also early attentional abilities—measured before the acquisition of reading—have been identified as predictors of later reading skills, particularly for decoding[8].

Related aspects of executive functions (EFs) have also been associated with reading abilities. Among these, working memory appears to be essential for holding sequences of phonemes in memory as decoding proceeds and for keeping track of successive words as sentences unfold[13–19]. In addition, cognitive flexibility may be relevant for building cross-modal connections and switching between the multiple dimensions of written content, such as phonology, morphology, semantics and syntax[20–24]. Finally, it has been argued that planning—the ability to identify, monitor and, if needed, revise strategies—may be relevant for reading comprehension[25,26].

Considering the importance of attentional skills for reading, training attentional control has been suggested as a promising way of lifting one of the several potential brakes on achieving literacy. To this end, one possible approach is making use of action video games, which are known to change for the better several attentional components in typical adults[27–29](see ref. [30] for an extensive meta-analysis of action video game impact on perceptual, attentional and cognitive skills) and children[31]. Franceschini and collaborators[32] trained Italian dyslexic readers with child-appropriate commercially available action-based video games and showed increases in reading speed without significantly altering reading accuracy. This effect, first established in dyslexic readers of a transparent orthography, Italian[32], has recently been replicated in dyslexic readers of a more opaque written language, English[33]. Although promising, these findings remain controversial[34,35] (for a review, see ref. [36]). Also, these past experiments mainly focused on reading-impaired children leaving it unclear whether training attentional skills would benefit reading abilities in typically developing children in a way that is relevant for academic achievement.

The current study evaluates whether, in a relatively large sample of typically developing children, training attentional control will indeed result in benefits in reading abilities, as measured by not only standardized tests but also school grades. An aspect of this work is the development of an entertaining child-appropriate video game, Skies of Manawak (SOM). This game was designed to target the diversity of cognitive challenges that reading poses and to leverage the reported positive impact of action video game-based training on attentional control (Fig. 1). The game play, situated in a child-friendly adventure world, revolves around two main game mechanics. One, 'The Flight', bears key characteristics of action

[1]Department of Psychology and Cognitive Science, University of Trento, Trento, Italy. [2]Faculte de Psychologie et Science de l'Education, University of Geneva, Geneva, Switzerland. [3]Université de Paris, LaPsyDÉ, CNRS, Paris, France. [4]Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. [5]Present address: Faculty of Computer Science, Free University of Bozen-Bolzano, Bolzano, Italy. [6]These authors contributed equally: Daphne Bavelier, Paola Venuti. ✉e-mail: angela.pasqualotto@unige.ch; daphne.bavelier@unige.ch

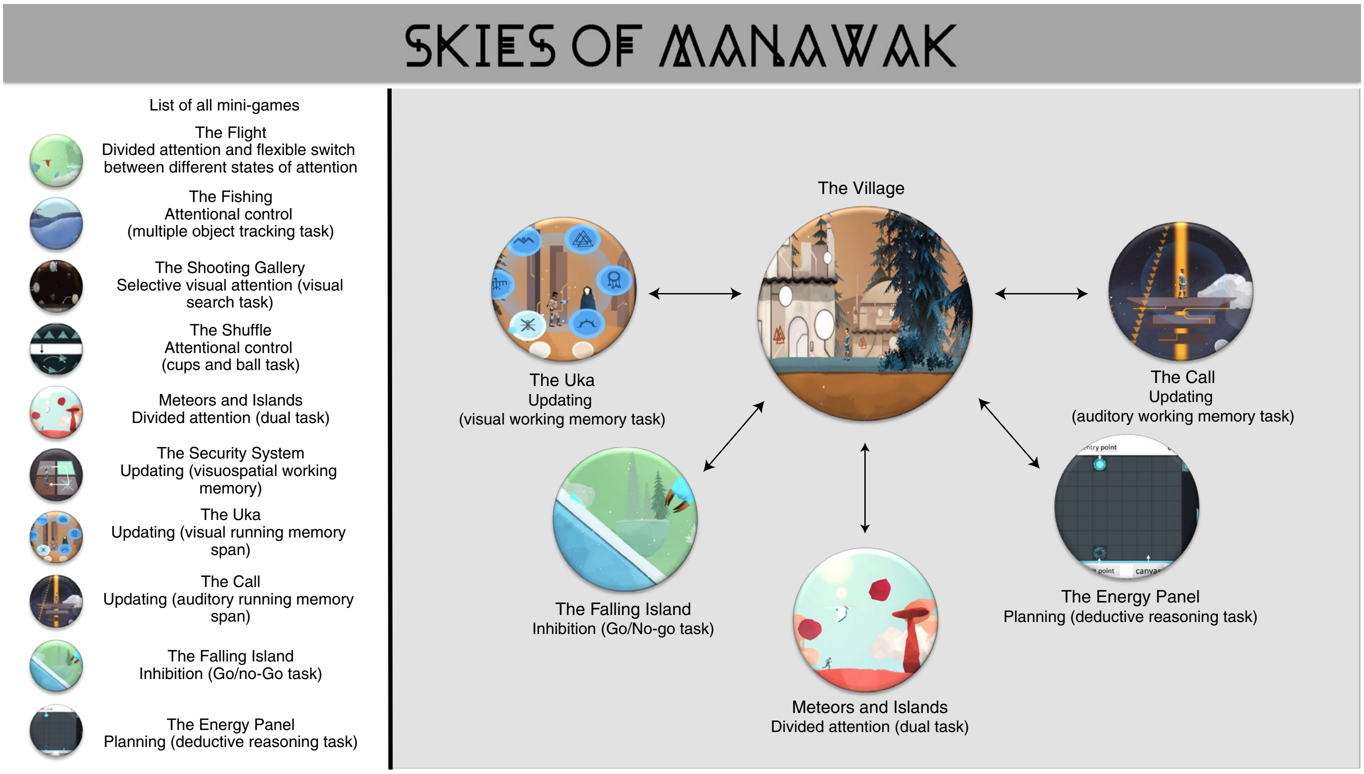

Fig. 1 | Skies of Manawak mini-games and targeted EF/attentional skills. We show here only the main executive subprocesses that are tapped by each of the mini-games.

video games and, in particular, a need for pacing, a load on distributed attention and the necessity to shift between divided and focused attention states at proper times[37]. The other, 'The Village', bears key characteristics of an incentive world, where the players through slow-paced, exploratory moves discover how points earned during the rest of the game play can be redeemed to personalize his/her world. In The Village, players also discover the next quest they are assigned to, sending them back into the action game. In addition to The Flight, the game consists of nine mini-games, each targeting aspects of EFs engaged during reading acquisition, such as visual and auditory working memory, inhibitory control and cognitive flexibility to cite a few. Best practice from cognitive training was naturally aligned with action video game mechanics in an effort to optimize impact on attentional control and its related EF[38,39]. Assigned mini-games, their in-game progression, as well as the action game difficulty was adaptively personalized. This game was developed over a period of 2 years through an iterative process that included game designers/artists, domain experts (cognitive and clinical scientists) and players (children between 8 and 14 years old who participated in story and aesthetics of the game through game ideation workshops, $n=60$, and helped refine the game through playtesting, $n=250$; Supplementary Information Section 5 on training regimens). We refer the reader to the Supplementary Table 1 for a synthesis of the attentional/executive skills targeted by SOM and their links to the various game elements in SOM.

A total of 151 typically developing Italian children were recruited in this training study. Both the experimental group ($n=79$) and the active control group ($n=72$) were pre-tested for visuospatial attention, verbal and spatial working memory, planning and literacy skills (T1). All children then underwent 12 h of training, distributed over 6 weeks, either on the experimental video game, named SOM, or on a control video game, Scratch, a kid-tailored, engaging programming game. These two different interventions were presented as active to parents and teachers. Crucially, training was implemented during class hours under the supervision of dedicated staff, trained either in the experimental or the control game, ensuring balanced implementation across the two arms of our intervention. The same cognitive measures were collected at post-test, within a week of the end of training (T2), as well as 6 months later (T3). Importantly, among the three experimenters collecting data, two were blinded to the aims of the study, allowing us to control for possible experimenter effects. Academic achievement in Italian and in Mathematics was additionally tracked at 1.0 year follow-up (T4) and 1.5 years follow-up (T5). We predicted greater improvements in attention and in reading skills from before to after training in the experimental group as compared to the control group. In addition, standardized reading assessments were expected to show more robust group differences than did academic achievement scores, as the former are more targeted to the study design, while the latter represent further transfer. Still a key question concerned whether reading grades may show greater benefit in the experimental than the control group over the years, a test of the predicted transfer to everyday life activities.

Results

We first consider the impact of the experimental video game, SOM, as compared to that of the control training game, Scratch, from pre-test (T1) to post-test (T2) and possible long-lasting effects at 6 months (T3) on both reading skills (word, non-word and text reading speed and accuracy, as well as types of reading errors made in the text reading task) and on a measure of attentional control. We refer the reader to Supplementary information Section 2 giving additional training effects and Supplementary Table 3 for the more distal outcomes of short-term memory, working memory, writing and comprehension of a short text passage, for which we only have measurements at two time points. We also consider here planning, an exploratory outcome and the only other measure collected at all three time points. Second, we tested the possibility that the experimental video game training enhanced reading skills through attentional enhancements as proposed by Franceschini and colleagues[40]. Finally, the impact of our training in the academic domain was evaluated by analysing Italian grades and Mathematics grades collected at five time points (T1, T2 and T3 but also 12 months after the end of training, T4, and 18 months after the end of training, T5).

Analysis plan. Data were analysed using SPSS Statistics for Mac, v.16.0 (IBM Corp). Group differences were examined by means of analyses of covariance using repeated measures general linear model analyses with time as a within-subject factor and group (SOM versus control) as a between-subject factor. The time × group interaction was our primary effect of interest and was followed with post-hoc analyses when significant. To exclude possible confounding effects, chronological age and IQ scores (Raven's CPM[41]) at T1 were entered as covariates and sex as a between-subject factor for all multivariate analysis of covariance/analysis of covariance (MANCOVA/ANCOVA) analyses. For each of our four different hypotheses (1a, reading skills; 1b, types of reading errors; 2, attentional control and planning; 3, academic performance), we applied Bonferroni correction for multiple comparisons. In addition, effect sizes were presented as partial η^2 (SSeffect / (SSeffect + SSerror)), where SS is the sum of squares. See Supplementary information Section 2 for covariate effects.

All participants' performance measures were transformed into z-scores on the basis of age norms, while for IQ the standard score (mean, $M = 100$; s.d. = 15) was considered. Reading errors were analysed as raw scores. Importantly, the two groups were comparable at baseline in all demographic and neuropsychological variables except text comprehension and proportion of word-substitution errors (Extended Data Fig. 1). The descriptive statistics for the main variables at the first assessment time point (T1) are presented in Extended Data Fig. 1 (literacy skills performance, attention and planning and other EF skills) and in Extended Data Fig. 2 (demographic and IQ characteristics).

Second, to evaluate the proposal that video game training enhanced reading skills through attentional enhancements, we ran separate analyses for reading speed (Δ T2 – T1 general reading speed) and for reading accuracy (Δ T2 – T1 general reading accuracy). Stepwise multiple regression analyses were computed taking into consideration covariates effects (age, IQ and sex), pre-test reading performance (T1) and changes in attention and planning (Δ T2 – T1), noting that planning should be considered as an exploratory variable. Analogous regression models were also performed on the two main categories of reading errors observed in our sample.

Finally, to evaluate the generalization of potential improvements to the academic domain, we computed two separate ANCOVA on Italian grades and Mathematics grades respectively, obtained at all five time points.

Skies of Manawak enhances reading skills. Reading skills were assessed by measuring speed and accuracy on the reading of word lists, of non-word lists (DDE-2 battery) and of meaningful text (MT reading test[42,43]). Separate 3 × 3 × 2 MANCOVA with time (T1, T2 and T3) and task (word, non-word and text) as within-subject factors, group and sex as between-subject factors, and age and IQ as covariates were carried out for speed and accuracy.

For reading speed, a time × group interaction ($F(2,144) = 27.211$, $P < 0.001$, $\eta^2 P = 0.274$, 95% CI_$\eta^2 = 0.153$; 0.377) indicated greater improvement in reading speed in the SOM group than in the control group, that was largely maintained at T3 (Fig. 2a and Supplementary Table 2). The only other significant effect was a triple interaction time × group × task ($F(2,144) = 5.280$, $P = 0.002$, $\eta^2 P = 0.129$, 95% CI_$\eta^2 = 0.006$; 0.151) highlighting that the highest improvements in reading speed were achieved by the SOM group on the text reading fluency task, compared to word and non-word reading tasks. When considering the covariates, a significant interaction between time and age ($F(2,144) = 5.117$, $P = 0.014$, $\eta^2 P = 0.066$, 95% CI_$\eta^2 = 0.005$; 0.148) revealed that, regardless of the group, younger children improved more than older children did.

For reading accuracy, a time × group interaction ($F(2,144) = 15.630$, $P = 0.002$, $\eta^2 P = 0.178$, 95% CI_$\eta^2 = 0.073$; 0.280) confirmed greater improvement in reading skills in the SOM group than in the control group (Fig. 2b); no other effects were significant. For details, see Supplementary Section 2 on covariate effects.

Skies of Manawak differentially reduces reading errors. Reading errors made in the text reading fluency task[42,43] were categorized as either sounding-out behaviours (that is, sounding out parts of the word before producing the whole word) or word-substitution errors, following the proposal of Hendriks and Kolk[44], which has been applied to Italian by Trenta and collaborators[45]. The categorization of reading errors applied here aims at differentiating between decoding capabilities and word unitization (that is, the process by which single linguistic units are consolidated into whole-word units)[46] through the use of separate scores. Sounding-out behaviours are diagnostic of fluency weaknesses in the sublexical, orthographic-to-phoneme mapping processes, whereas word-substitution errors (in the presence of normal fluency) rather point to weaknesses in lexical, orthographic and phonological to semantic access processes[45–48].

As above, a 3 × 3 × 2 MANCOVA with time (T1, T2 and T3) and error type (sounding-out and word substitution) as within-subject factors, group and sex as between-subject factors, and age and IQ as covariates were carried out on error rates. A significant time × group interaction ($F(2, 144) = 15.940$, $P < 0.001$, $\eta^2 P = 0.181$, 95% CI_$\eta^2 = 0.075$; 0.283) indicated greater reduction of reading errors in SOM group compared to the control group, a reduction that was largely maintained at T3 (Fig. 3 and Supplementary Table 2). Moreover, a triple time × group × type of errors interaction ($F(2,144) = 11.635$, $P = 0.008$, $\eta^2 P = 0.074$, 95% CI_$\eta^2 = 0.045$; 0.238) highlighted a greater reduction in sounding-out behaviours compared to the word-substitution errors in the SOM group than in the control group.

Skies of Manawak enhances visuospatial attention. We examined attention after training by means of a classic barrage task[49], where children are asked to find as many target objects (bells) amidst distractors as possible. The mean accuracy score between the 'fast' score (number of targets found within 30 s) and the 'slow' score (total number of targets identified in 120 s) was estimated and used as a composite-dependent variable.

The 3 × 2 ANCOVA with time (T1, T2 and T3) and group (SOM versus control) and sex (age and IQ as covariates) revealed a significant time × group interaction ($F(2,144) = 63.764$, $P < 0.001$, $\eta^2 P = 0.47$, 95% CI_$\eta^2 = 0.349$; 0.557). Figure 4a highlights greater improvements in attention in SOM than in the control group from T1 to T2, with the effect being largely maintained at T3. No other significant effects were observed but for a main effect of group ($F(1,145) = 52.365$, $P = 0.001$, $\eta^2 P = 0.265$, 95% CI_$\eta^2 = 0.297$; 0.514). See Supplementary Section 2 on covariate effects for details.

In addition, we carried out exploratory analyses on planning, which we had assessed at all three time points through the Tower of London task[50]. A similar ANCOVA with the accuracy of planning as the dependent variable revealed a significant time × group interaction ($F(2,144) = 15.546$, $P = 0.002$, $\eta^2 P = 0.178$, 95% CI_$\eta^2 = 0.072$; 0.280), indicating greater improvement in the SOM than in the control group from T1 to T2, with this effect being sustained at T3 (Fig. 4b and Supplementary Table 2). The only other significant effect was a triple interaction time × group × sex ($F(1, 145) = 7.852$, $P = 0.012$, $\eta^2 P = 0.051$, 95% CI_$\eta^2 = 0.004$; 0.135) discussed in Supplementary Section 2 on covariate effects).

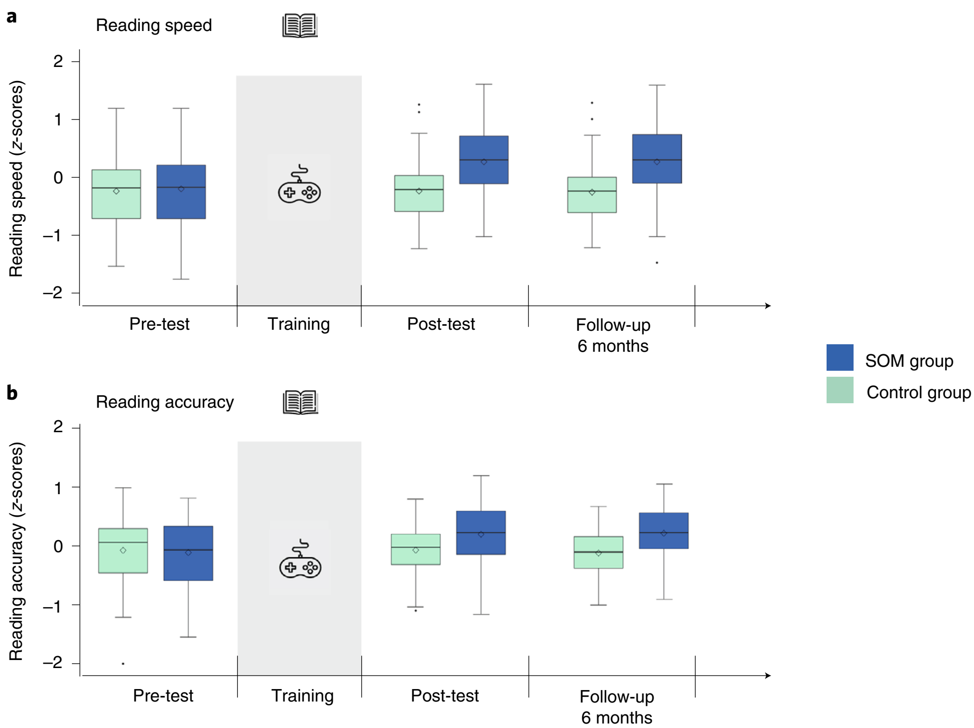

Fig. 2 | Combined performance on the word, non-word and text reading assessments: speed and accuracy. a,**b**, Distribution of word, non-word and text reading speed (**a**) and accuracy (**b**), as measured before training (T1), after training (T2) and at 6 months follow-up (T3) in the SOM-trained ($n=79$) and the Scratch-trained (control) ($n=72$) groups. Reading performance is expressed in *z*-scores. The solid horizontal line shows the median, the diamond shows the mean, the box is bounded by the upper and lower quartile values; the whiskers are set to the maximum observed value just below 1.5× the upper quartile for the top one and the minimum observed value just above 1.5× the lower quartile for the bottom one. Relevant statistical information and notes are provided in Supplementary Table 2.

Thus, as expected given the specific focus of SOM in training attention and higher executive function skills, 12 h of SOM play led to greater improvement in these skills, as compared to playing the same amount of the control game, Scratch.

In summary, the SOM-trained group exhibited greater improvement in reading speed and in reading accuracy, as well as in our measures of attentional control and planning, than did the control-trained group. These differential gains were largely maintained 6 months later at follow-up.

Controlling for possible test–retest learning confounds. The same tests (namely, all reading tests, as well as the barrage and planning tasks) were used through the three assessment sessions, meaning that improvements could be due to better memorization of the test items per se rather than enhanced reading or attentional skills. To check for such a possible test–retest learning confound, two tests were administered at the 6-months follow-up: a new text reading task[51] and a barrage task[51] (Supplementary Section 3). A Pearson correlation was run to determine the relationship between these tests and the original text reading task and barrage task as measured at follow-up. We then checked the differences in these tests at follow-up between SOM and control group.

Strong, positive correlations were observed between the two reading tests scores ($r=0.653$, $n=151$, $P<0.001$, 95% CI_$\eta^2=0.551$; 0.735) as well as between the two attention tasks scores ($r=0.970$, $n=151$, $P<0.001$, 95% CI_$\eta^2=0.959$; 0.978), establishing that the tasks capture similar constructs as the original ones.

Separate one-way ANCOVAs were conducted to compare the effect of training regimens (SOM versus control) on the text reading and the attentional control tasks at T3. For each of the tasks, a main group effect confirmed an advantage in reading and in attention for the SOM-trained group over the control-trained group: new text reading (SOM: $M=0.46$, s.d. = 0.50; control: $M=-0.025$, s.d. = 0.49; main effect of group: $F(1, 150)=36.575$, $P<0.001$, $\eta^2P=0.201$, 95% CI_$\eta^2=0.093$; 0.302) and new barrage task (SOM: $M=1.25$, s.d. = 0.85; control: $M=-0.01$, s.d. = 0.63; main effect of group: $F(1, 150)=112.629$, $P<0.001$, $\eta^2P=0.437$, 95% CI_$\eta^2=0.312$; 0.523). Thus, the advantages of the SOM-trained group in reading and in attention are not limited to the tests used at T1–T2–T3 but rather generalize to similar tasks with stimuli at T3.

Role of attentional control in fostering literacy skills. A link has been proposed between training attention and enhancing reading abilities, especially as regards reading speed. A stepwise multiple regression analysis was performed on the entire sample of children ($n=151$). The dependent variables were reading speed or reading accuracy improvements between T1 and T2 and the predictors were: (1) age, IQ and sex; (2) general reading speed respectively reading accuracy at T1, (3) attentional control changes (T2 – T1) and as an exploratory variable, changes in planning (T2 – T1).

The first stepwise multiple regression (for details, see Supplementary Table 4) concerned reading speed. About the demographic variables, only age significantly explained some of the variance in reading speed changes (4.2%, $F(1, 149)=6.53$, $P<0.01$, $\Delta R^2=0.042$). A model, which not only included age but also sex and IQ, did not provide a significantly better fit to the data compared to the model with age only ($F(1,147)=0.001$, $P=0.92$, Bayes factor (BF) = 0.082). Reading speed at T1 explained an additional 11% of variance ($F(1, 146)=18.47$, $P<0.001$, $\Delta R^2=0.107$). In particular, children who improved most were those who started with lower levels of reading speed at T1. Finally, changes in planning between T1 and T2 did not provide a better fit (($F(1,145)=2.289$,

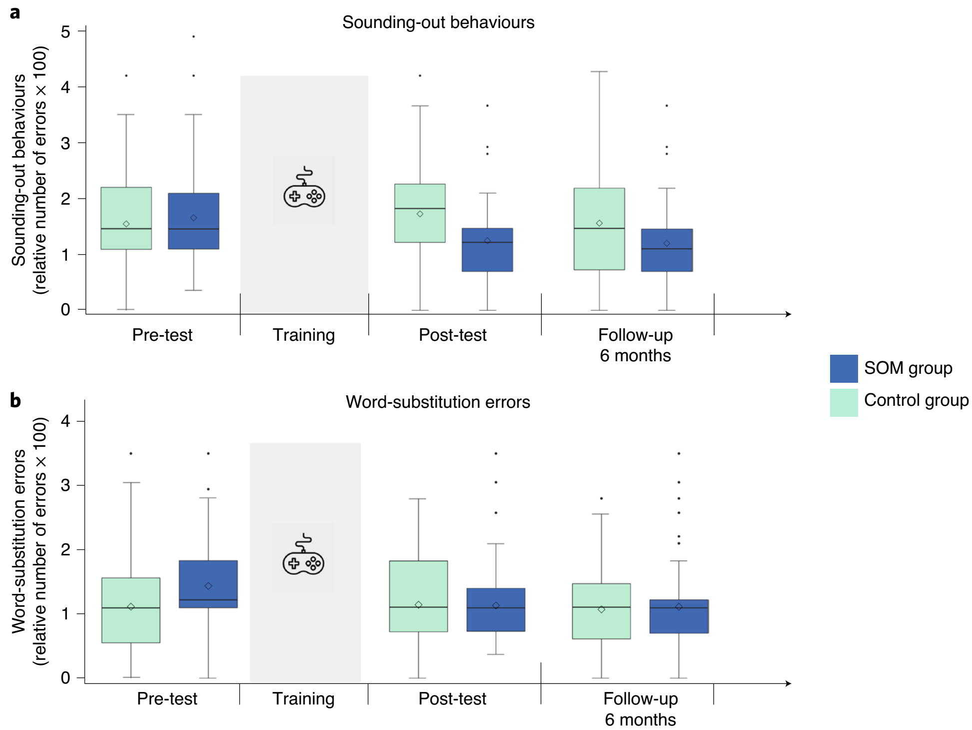

Fig. 3 | Reading errors: sounding-out behaviours and word substitutions. a,**b**, Distribution of reading errors in terms of sounding-out behaviours (**a**) and word-substitution errors (**b**), as measured before training (T1), after training (T2) and at 6 months follow-up (T3) in the SOM-trained ($n = 79$) and the Scratch-trained (control) ($n = 72$) groups. It should be noted that compared to children reading opaque languages (for example, English), Italian-speaking children make considerably fewer errors. Thus, errors—expressed as the number of errors over the total number of enunciated words (×100)—are relatively rare. Box plot conventions as in Fig. 2. Relevant statistical information and notes are provided in Supplementary Table 2.

$P < 0.133$, $\Delta R^2 = 0.011$, BF = 0.268); yet the addition of change in attention between T1 and T2 explained an additional 14% of variance (F(1,144) = 29.143 $P < 0.001$, $\Delta R^2 = 0.142$).

A similar stepwise multiple regression with reading accuracy as dependent variable showed that none of the demographic variables provided a significantly better fit to the data ($F(1,147) = 1.784$, $P = 0.18$, BF = 0.202). Reading accuracy score at T1 uniquely explained 38% of the variance in reading accuracy changes ($F(1,146) = 90.58$, $P < 0.001$, $\Delta R^2 = 0.378$). Again, children with worse accuracy improved more overall. Adding the improvements in planning to the regression model did not provide a better fit to the data compared to the previous model (($F(1,145) = 3.273$, $P < 0.72$, $\Delta R^2 = 0.013$, BF = 0.439), whereas the addition of the improvements in attention explained an additional 2.6% of the variance ($F(1,144) = 6.56$, $P = 0.01$, $\Delta R^2 = 0.026$). For details, see Supplementary Table 5.

Furthermore, in a further, more exploratory analysis, we investigated the link between the reduction in reading errors and the attentional improvements. We performed a stepwise multiple regression analysis on the entire sample of children. The dependent variable was the changes in sounding-out behaviours and, as above, the predictors were: (1) age, IQ and sex; (2) sounding-out behaviours at T1, (3) attentional control changes (Δ T2 – T1) and, as an exploratory variable, planning changes (Δ T2 – T1). The delta scores in sounding-out behaviours were calculated by taking the difference between sounding-out behaviours at T2 and T1. Analogously, to measure the unique effect of attentional improvements on word-substitution errors reduction, we performed a similar separate analysis, this time using the changes in word-substitution errors as dependent variable and pre-test word-substitution errors as one of the predictors of the regression model. For details, see Supplementary Table 6.

The results showed that none of the demographics variable contributed variance in either sounding-out (SO) or word-substitution (WS) errors changes (SO: $F(1,147) = 0.897$, $P = 0.35$, $\Delta R^2 = 0.006$; BF = 0.129; WS: $F(1,147) = 0.614$, $P = 0.43$, $\Delta R^2 = 0.004$; BF = 0.112). Errors at T1 uniquely explained 28% of the variance in sounding-out behaviours (SO: $F(1,146) = 56.942$, $P < 0.001$, $\Delta R^2 = 0.278$) and 27% in word substitution (WS: $F(4,146) = 55.905$, $P < 0.001$, $\Delta R^2 = 0.271$). Specifically, children expressing such behaviours more frequently at T1 were those who improved the most overall. Adding improvements in planning and attention accounted for, respectively, 3% and 6% of the variance in sounding-out behaviours changes (planning—$F(1,145) = 6.38$, $P < 0.013$, $\Delta R^2 = 0.03$; attention—$F(1,144) = 13.703$, $P < 0.001$, $\Delta R^2 = 0.059$); yet, no such effect was observed for word-substitution errors (planning BF = 0.098; attention BF = 0.18). These results link attentional control to the reduction in sounding-out behaviours and thus more efficient orthographic-to-phoneme mapping processes (that is, sequencing skills). For details, see Supplementary Tables 6 and 7.

Links with academic performance. In this study, eight classes (two third grade, two fourth grade, two sixth grade and two seventh grade) were involved. For each grade, one class was randomly allocated to the SOM group and the other one to the control group. The two pairs of classes in elementary school (third and fourth grades) shared the exact same teams of teachers, while the middle school classes had in common Italian, Mathematics & Science and English teachers (for a total of 19 teaching hours out of 30 h per week). Such matching at the teacher level was preferred to limit possible confounds due to teaching styles when considering the impact of training on academic achievements.

Grades in Italian and in Mathematics were entered into two separate 5 × 2 ANCOVA with time (T1, T2, T3, T4 and T5) and groups

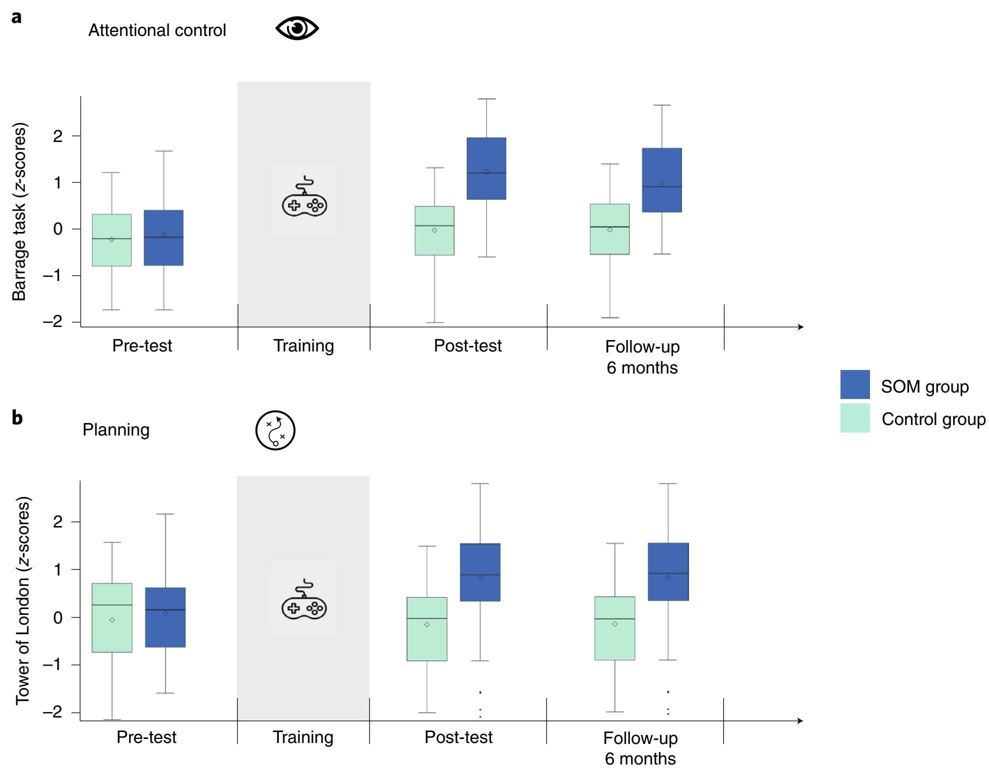

Fig. 4 | Performance in the barrage task (attentional control) and the Tower of London task (planning). a,**b**, Distribution of the performance (accuracy score) in the barrage task (**a**) and the Tower of London task (**b**), expressed in *z*-scores, as measured before training (T1), after training (T2) and at 6 months follow-up (T3) in the SOM-trained ($n=79$) and the control-trained ($n=72$) groups. Box plot conventions as in Fig. 2. Relevant statistical information and notes are provided in Supplementary Table 2.

(SOM and control). For Italian, a significant time × group interaction ($F(4,\ 142)=3.448$, $P=0.016$, $\eta^2P=0.023$, 95% CI_$\eta^2=0.006$; 0.165) was observed, highlighting growing group differences over time, due to a small advantage of the SOM group over the control group over time (Fig. 5).

For Mathematics, no significant time × group interaction ($F(4,\ 142)=0.778$, $P=1.08$, $\eta^2P=0.005$, 95% CI_$\eta^2=-0.004$; 0.160) was observed. See Supplementary Sections 2 for covariate effects.

Discussion

The current study presents a custom-designed video game for children targeting attentional control training and EFs relevant to reading, with the goal of facilitating some aspects of reading acquisition. The SOM game includes several mini-games loading on EFs, each accessed through a central game encompassing action video game mechanics, hypothesized to enhance attentional control[37]. Overall, our results demonstrate both near- and far-transfer benefits following such training as compared to training on Scratch, a block-based visual programming game. Children in the experimental group showed greater enhancements in attentional control as compared to the control-trained group, as well as greater improvements in reading speed and accuracy. These enhancements were maintained 6 months after the end of training. Importantly, none of the training activities required reading nor searching for static target objects among distractors, as in the paper-and-pencil barrage tasks we used in the assessments.

Moreover, benefits—though of a small effect size—were also observed in school grades in Italian—a core school subject—18 months after training, thus showing further transfer to academic performance. Crucially, no comparable improvements were observed as an outcome of the active control intervention, using Scratch.

These results extend effects previously reported in the context of developmental dyslexia remediation to a large sample of typically developing children. A handful of studies have used non-conventional training tools such as a child-friendly, commercial video game (Rayman's Raving Rabbids) to yield improvements in both attentional control and reading speed in Italian-[32,52] and English-speaking[33] dyslexic children (but see ref. [35] for a lack of replication). Importantly, our results expand on those previous findings by highlighting benefits in not only reading speed but also reading accuracy, the latter being crucially aligned with the general aim of increasing literacy outcomes as broadly as possible and potentially supporting the development of text comprehension later in time[53]. Moreover, the gains in reading accuracy can be further appreciated by considering changes in reading behaviours. Indeed, the scoring performed on reading behaviours during text reading allowed us to separately assess orthographic-to-phonological decoding from recognizing words as whole units. Results showed that children who trained with the experimental video game displayed overall fewer errors after training than those trained with the control intervention, an effect driven possibly by an even greater reduction in sounding-out behaviours at post-test in the experimentally trained group. Importantly, this advantage was maintained at a follow-up test 6 months later.

Additionally, our results add to previous literature by showing that the obtained benefits on attentional control and reading abilities are lasting, as they were found to be maintained 6 months after the end of training.

A distinctive feature of the present study was to simultaneously track academic achievement in Italian and to continue tracking performance in Italian classes for another full academic year. Italian grades comprise both written and oral comprehension; speaking and writing capacities (that is, oral and written production); as well

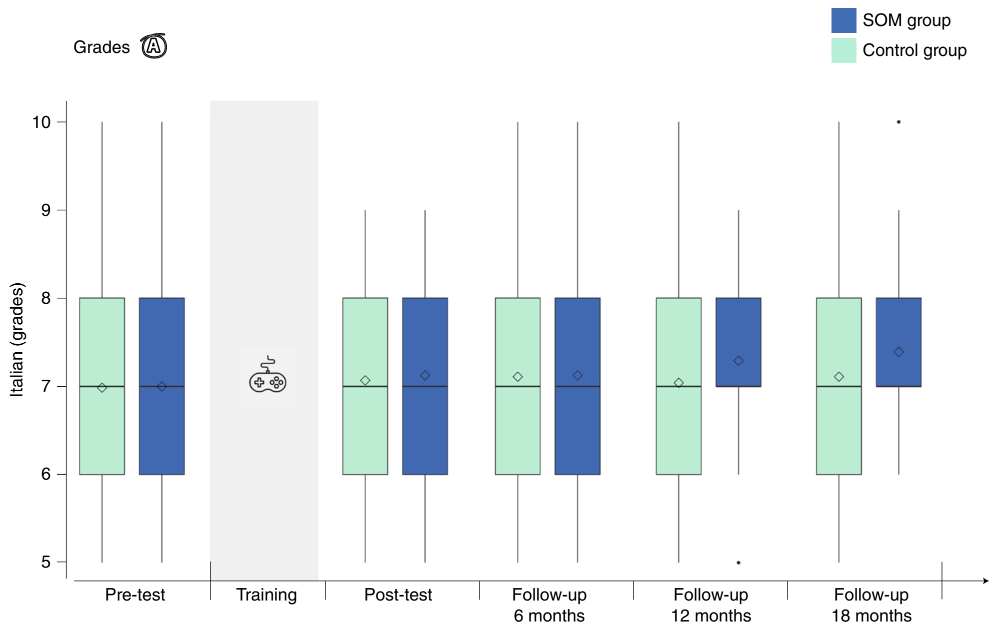

Fig. 5 | Distribution of academic grades (out of a maximum of ten) in Italian for children of the two groups (SOM and control) at five time points. The time points used are: pre-test (T1, at the beginning of the first school year), post-test (T2, at the end of the first semester) and follow-ups at 6 months after the end of the training (T3, at the end of the first school year), 12 months after the end of the training (T4, at the end of the first semester of the second school year) and 18 months after the end of the training (T5, at the end of the second school year). Box plot conventions as in Fig. 2. Relevant statistical information and notes are provided in Supplementary Table 2.

as metalinguistic skills (that is, knowledge and application of grammar rules in Italian, which is the participants' mother tongue). We show that our video game intervention provides transfer beyond laboratory-based assessment of reading, namely by positively affecting school grades in Italian.

Importantly, children, parents and all but one experimenter were blinded as to the aim of the study. Note that an analysis of only the children assessed by blind experimenters confirms all main results (see Supplementary Information Section 2 on controlling for possible testing confounds). Although teachers could not be blinded to their class assignment, the finding that differences in Italian grades were only seen 18 months after the training speaks against expectation effects. A potential weakness of this work is that the cognitive tests administered at pre-tests, post-tests and follow-up tests used the exact same items. At follow-up, however, two additional measures of reading ability and attention were added to ensure that greater gains in the SOM group as compared to the control-trained group were not due to greater memory of the very test items administered at all time points. This was indeed the case. By showing generalization to test items, these results provide further support for the observation that SOM brings benefits to reading ability and attentional control skills per se.

A critical issue concerns the link between reading performance and attentional control skills. As for Franceschini and collaborators[32], we tested for a link between reading improvement from pre- to post-test and attention improvements. We did confirm such a relationship. This is in line with a number of previous investigations that have suggested the existence of a link between attention and reading development. For instance, in typically developing children, a longitudinal study[8] showed that attentional skills evaluated at pre-reading stages are predictive of reading abilities later in grade 1 and grade 2. Indeed, several attentional components have been related to reading proficiency. Among these, efficient attentional control appears to be essential for moving the spotlight of attention over the letters of a text, while inhibiting the processing of close phonological or orthographic competitors and flexibly shifting from one line to the other[10,54]. Attentional control is also fundamental for visual word recognition, which consists of the analysis of different linguistic attributes of printed words (that is, the orthographic, phonological and semantic). Indeed, it has been shown that attention has the capacity of enhancing the processing of printed material, flexibly switching between different levels of information processing depending on task goals[12].

In addition, greater attentional control and more generally better ability to attend to task-relevant materials and ignore sources of noise or distractions have been linked to enhanced reading skills in typically developing children[8,12] and in dyslexic children[54–56].

In line with those previous results, our study also highlights a putative link between attentional control and reading abilities, as a multiple regression on the entire sample of children has shown that gains in attention explained up to 15% of the variance in reading speed.

This finding should be taken with a note of caution, however. We acknowledge that the existence of a causal link between gains in attentional control and gains in reading abilities remains to be established. Indeed, while changes in attentional control could explain part of the variance in reading gains across the whole sample of children ($n=151$), this relation was not confirmed when splitting the whole population into the two training groups (see Supplementary Section 2 on the role of attentional control on fostering reading skills). Differences in attentional control gains and in reading gains between the groups are thus driving the observed relation, not allowing us to firmly conclude the existence of a causal link between the two.

An interesting and important question concerns the impact of any training regimen on real-life competence and, in the case of school-aged children, on academic achievement. By collecting grades in both Italian and Mathematics at multiple time points and up to 18 months after the end of training, we uncovered increasing group differences over time in Italian grades, with an advantage of the SOM-trained group over the control-trained group. Few studies have managed to collect this type of information. Among those that have, Goldin and collaborators[57] did uncover the impact of multiple EFs training in 6–7-year-old children when considering a composite grade for school mathematics and literacy/language. It remains unclear, given the use of such a composite, whether both literacy and mathematics improved. While several experiments indicate enhanced school mathematics after EF training, grades for mathematics did not differ across training in our experiment. This null effect is difficult to interpret as the control game Scratch may have positively impacted school mathematics[58]. The greater improvement in school language grades with SOM underlines the far-transfer effects of our intervention. Interestingly, the uncovered effects on school grades in Italian were delayed in time, being statistically significant only 18 months after the end of the intervention. This is in line with studies indicating that positive effects of training might be detectable in everyday life activities only after a long delay (that is, 6 months to 1 year)[59,60].

Overall, this study highlights how training with SOM not only improved attentional control and reading skills but also other fundamental components of the executive system. Taking stock of previous studies (refs. [14,19,60–63] and for two extensive meta-analyses of the benefits of working memory programmes see ref. [64,65]), SOM was purposefully designed to train attentional control, in particular working memory and cognitive flexibility[39]. Interestingly, SOM consists of a variety of training tasks, to obtain overall improvement in domain-general abilities and higher chances of generalization to academic performance. Another design feature of SOM is to encompass challenging and variable activities integrated within a main story frame, requiring children to continuously adapt their behaviour, while creating a fair sense of appropriation and competition. This departs from many cognitive training programmes that tend to use just a few stand-alone mini-tasks[66,67]. In this regard, the overall appreciation for the training activities was high in terms of entertainment, engagement and motivation for the children. Importantly, SOM and the control game were comparable in that regard (see the Supplementary Information Section 2 feedback questionnaire). Finally, thanks to the high adaptability of the video game, each player pursues a different training path: for example, the children who performed worse at the beginning of the training in a specific mini-game that trains inhibition will be exposed to more exercises that tap into that specific cognitive function. Vice versa, children who already possessed adequate levels of a specific skill will be trained less on that specific skill in favour of others.

There are a few limitations of our study we need to acknowledge. In the current version of SOM, data from the participants' game play were not recorded. A few studies report that game play performance during training correlates with performance in standardized neuropsychological tests (for example, NEPSY-II) assessing visuospatial, attentional and language skills[68]. In addition, Franceschini and Bertoni[52] showed that only the children who improved through the game play obtained significant transfer effects in untrained, cognitively related tasks such as reading skills. In the present experimental game, in-game data have the potential to disentangle the effects of the more action-like video game play (that is, timed activities that involve response under time pressure, divided attention, timely shifts between focused and divided) from the effects of the mini-games that do not have any specific action game constraints. Further studies are thus required to fully exploit video game play as a way of assessing how changes in various cognitive skills relate to reading gains. These further investigations will shed critical light on the factors that most successfully predict the outcome of training for reading skills. It is also of note that, while the present results form a promising basis for more individually based training, future research should more carefully address the cognitive profile of each child beyond language and into the executive domain, with the aim of reaching the best training algorithm on the basis of the pattern of strengths and weaknesses of each child[69]. Finally, it will be important to assess SOM efficacy in children speaking other languages, such as English or French, to understand whether benefits are seen in an opaque orthography, in which phonological components are expected to play a larger role in reading acquisition and automatization[70,71].

Methods

Participants. A total of 151 typically developing Italian children were recruited from eight classes of third, fourth, sixth or seventh grades (8–12 years old) within a northern Italy public school. One of the two classes included at each grade were randomly assigned to either the training or the control group. Although all children within each class were treated similarly, criteria for data analyses inclusion were: (1) normal or corrected to normal visual acuity; (2) no attention deficit hyperactivity disorder diagnosis; (3) no diagnosis of learning disorder or no reading delay in word and non-word reading tests; and (4) no diagnosis of intellectual disability. Children with learning disorders ($n=21$) or with borderline intellectual functioning ($n=10$) participated in the study like any other classmate; however, their data were not included in the analysis sample leading to a final sample of $n=79$, 39 females, in the experimental group and $n=72$, 41 females, in the control group.

All the children's parents gave written informed consent after a description of the research study, in accordance with the principles of the Declaration of Helsinki; the University of Trento ethics committee approved the research protocol (protocol no. 2019/048). At the end of the school year, participants were thanked for their time and effort with a small gift.

Study design. The study included five phases, a first data and grade collection time point, followed by the intervention and then two other data and grades collection time points; finally, grades were collected at two later time points.

Procedures. *Testing.* Children were tested at school. Baseline (T1) assessments were completed 1 week before the start of treatment and outcome assessments (T2 and T3) were carried out 1 week and 6 months after the end of the treatment. Tasks were administered in a pseudorandom manner; IQ (Raven's CPM[41]) and comprehension abilities (MT reading task[42,43]) were always assessed in a collective session, whereas the other tests were administered individually in a quiet room. Other tasks included attentional control (modified Bells test[49]), verbal and visual working memory (digit and visuospatial span[72], planning (Tower of London[50]) and literacy skills (word, non-word and text reading[42,43]; word and non-word writing). A detailed description of the measures is presented in Supplementary Information Section 3 on tools. Note that IQ was only assessed once at T1; otherwise, T1 and T2 assessment used identical test materials. The assessment battery at T3 consisted of a selection of tests previously used (for reading, attention and planning) and two additional tests (text reading task that measures reading abilities of meaningful material and a barrage task that evaluates visuospatial attention[51]). Testing was conducted by three different clinicians, two of them blinded to childrens' assignments and one of them (A.P.) being the experimental game expert. Note that although we report data over the whole sample, results from the sample tested by blinded experimenters only did not differ from the whole sample (see Supplementary Section 2 on controlling for possible testing confounds).

Classification of the reading errors. Reading performance was further analysed adopting the error classification proposed by Hendricks and Kolk[44] and applied to Italian by Trenta and colleagues[45]. Reading behaviours of each student at the three time points when reading the MT text passage were analysed by focusing on the syllabic level. The passages are normed for each grade and thus contain different numbers of syllables (third grade, 305 syllables; fourth grade, 297 syllables; sixth grade, 592 syllables; seventh grade, 596 syllables).

The classification focused on the contrast between two types of reading behaviours: sounding-out and word-substitution behaviours. The former refers to a particular behaviour in which the child tries to read the target word through several attempts (for example, the child has to read 'crocodile' and instead he/she reads 'cr.cro.crocodile'). In contrast, the latter indicates a substitution of the target word with another word (for example, the child has to read 'mood' and he/she reads 'moon'). The classification includes a third category, that is, residuals, combining reading errors or dysfluencies that are not comprised in the other two main categories. A brief description of each subcategory of the classification used from Trenta and colleagues[45] (along with examples of errors made in our sample of Italian typically reading children) can be found in the Supplementary Section 4 on errors classification. For examples in English and Dutch, please refer to Hendriks and Kolk[44].

At pre-test, the proportion of reading errors were rather equally distributed between sounding-out behaviours (SOM 48.6% and control 44.6%) and word substitutions (SOM 41.9% and control 34.1%) with <1% of residual errors (given this small percentage, this type of errors was not analysed further). Fleiss' kappa procedure[73] was used to determine reliability. Scoring was carried out by A.P. and two additional independent raters, who analysed 20% of the data, randomly selected. All the raters were blind to group allocation and to the time points in which data were collected. The reading errors classification showed very high inter-rater reliability, with kappa values (95% CI) of 0.81 (0.80, 0.81), 0.85 (0.84, 0.85), 0.45 (0.44, 0.45) for sounding-out, word substitution and residuals, respectively.

Grade collection. School grades were obtained directly from school teachers. The same teacher provided grades for both the experimental and the control class they had been assigned. Maths and Italian grades corresponded to those in the report cards containing the average grade for each semester.

Intervention treatment. Training was administered at school during school time with duration and frequency (1 h per session, twice a week for 6 weeks) matched across the two intervention groups. All children trained for at least 11 h out of the 12 h of game play ($M = 11.40$; s.d. $= 0.27$) to be included in the analyses. The experimental group trained on SOM; those in the control group received a computer-based training on coding using Scratch, a programming language developed by the Lifelong Kindergarten (MIT) Group, which allows children to create interactive stories, games, animations and simulations[74]. Each game was administered under the supervision of an expert experimenter ensuring that participants were focused and that they remained motivated, encouraging them whenever needed. The SOM expert was A.P.; the Scratch expert was a computer science teacher also familiar with the school. It has to be noted that none of the activities required reading: in SOM, text-to-speech audio clips were created for each dialogue, providing a full dubbing of the game, while in Scratch all the explanations regarding the software's functionalities and the different task assignments were provided orally by the teacher. Furthermore, at the end of the training all children completed a questionnaire aimed at investigating their motivation, the perceived difficulty and appreciation of the video game, as well as their personal evaluation of the training. No statistically significant differences in the level of appreciation between the two training activities were found ($P = 0.243$; for further details, see the Supplementary Information Section 2 feedback questionnaire).

Additional information regarding SOM and the control activity, Scratch, can be found in the Supplementary Information 5 on training regimens.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Behavioural data, analysis code and study materials have been deposited in Open Science Framework with the identifier: https://doi.org/10.17605/osf.io/4rzgc (https://osf.io/4rzgc/).

Code availability

Code for figures and analyses of behavioural data are available from the corresponding author on request.

Received: 26 March 2020; Accepted: 1 November 2021;
Published online: 17 January 2022

References

1. Parsons, S. & Bynner, J. *Illuminating Disadvantage: Profiling the Experiences of Adults with Entry level Literacy or Numeracy Over the Lifecourse* (NRDC, 2007).
2. Pape, K., Bjørngaard, J. H., Westin, S., Holmen, T. L. & Krokstad, S. Reading and writing difficulties in adolescence and later risk of welfare dependence. A ten year follow-up, the HUNT Study, Norway. *BMC Public Health* **11**, 718 (2011).
3. Carneiro, R. & Gordon, J. Warranting our future: literacy and literacies. *Eur. J. Educ.* **48**, 476–497 (2013).
4. Snowling, M. Dyslexia: a language learning impairment. *J. Br. Acad.* **2**, 43–58 (2014).
5. Ziegler, J. C. & Goswami, U. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* **131**, 3–29 (2005).
6. Elhassan, Z., Crewther, S. G. & Bavin, E. L. The contribution of phonological awareness to reading fluency and its individual sub-skills in readers aged 9- to 12-years. *Front. Psychol.* **8**, 533 (2017).
7. Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. Specific reading disability (dyslexia): what have we learned in the past four decades? *J. Child Psychol. Psychiatry* **45**, 2–40 (2004).
8. Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K. & Facoetti, A. A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* **22**, 814–819 (2012).
9. Grainger, J., Tydgat, I. & Isselé, J. Crowding affects letters and symbols differently. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 673–688 (2010).
10. Grainger, J., Dufau, S. & Ziegler, J. C. A vision of reading. *Trends Cogn. Sci.* **20**, 171–179 (2016).
11. Strijkers, K., Bertrand, D. & Grainger, J. Seeing the same words differently: the time course of automaticity and top-down intention in reading. *J. Cogn. Neurosci.* **27**, 1542–1551 (2015).
12. Ruz, M. & Nobre, A. C. Attention modulates initial stages of visual word processing. *J. Cogn. Neurosci.* **20**, 1727–1736 (2008).
13. Verhoeven, L., Reitsma, P. & Siegel, L. S. Cognitive and linguistic factors in reading acquisition. *Read. Writ.* **24**, 387–394 (2011).
14. Karbach, J., Strobach, T. & Schubert, T. Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychol.* **21**, 285–301 (2015).
15. Loosli, S. V., Buschkuehl, M., Perrig, W. J. & Jaeggi, S. M. Working memory training improves reading processes in typically developing children. *Child Neuropsychol.* **18**, 62–78 (2012).
16. Siegel, L. S. & Ryan, E. B. The development of working memory in normally achieving and subtypes of learning disabled children. *Child Dev.* **60**, 973–980 (1989).
17. Gathercole, S. E., Willis, C. S., Baddeley, A. D. & Emslie, H. The children's test of nonword repetition: a test of phonological working memory. *Memory* **2**, 103–127 (1994).
18. Baddeley, A. Working memory and language: an overview. *J. Commun. Disord.* **36**, 189–208 (2003).
19. Henry, L. A., Messer, D. J. & Nash, G. Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. *Infant Child Dev.* **23**, 84–103 (2014).
20. Blair, C. & Razza, R. P. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* **78**, 647–663 (2007).
21. Colè, P., Duncan, L. G. & Blaye, A. Cognitive flexibility predicts early reading skills. *Front. Psychol.* **5**, 565 (2014).
22. Cartwright, K. B. et al. Cognitive flexibility deficits in children with specific reading comprehension difficulties. *Contemp. Educ. Psychol.* **50**, 33–44 (2017).
23. Cartwright, K. B., Marshall, T. R., Huemer, C. M. & Payne, J. B. Executive function in the classroom: cognitive flexibility supports reading fluency for typical readers and teacher-identified low-achieving readers. *Res. Dev. Disabil.* **88**, 42–52 (2019).
24. Christopher, M. E. et al. Predicting word reading and comprehension with executive function and speed measures across development: a latent variable analysis. *J. Exp. Psychol. Gen.* **141**, 470–488 (2012).
25. Reiter, A., Tucha, O. & Lange, K. W. Executive functions in children with dyslexia. *Dyslexia* **11**, 116–131 (2005).
26. Locascio, G., Mahone, M. E., Eason, S. & Cutting, L. E. Executive dysfunction among children with reading comprehension deficits. *J. Learn. Disabil.* **43**, 441–454 (2010).
27. Bejjanki, V. R. et al. Action video game play facilitates the development of better perceptual templates. *Proc. Natl Acad. Sci. USA* **111**, 16961–16966 (2014).
28. Cardoso-Leite, P. & Bavelier, D. Video game play, attention, and learning. *Curr. Opin. Neurol.* **27**, 185–191 (2014).
29. Bavelier, D., Green, C. S., Pouget, A. & Schrater, P. Brain plasticity through the life span: learning to learn and action video games. *Annu. Rev. Neurosci.* **35**, 391–416 (2012).
30. Bediou, B. et al. Meta-analysis of action video game impact on perceptual, attentional, and meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychol. Bull.* https://doi.org/10.1037/bul0000130 (2018).
31. Dye, M. W. G., Green, C. S. & Bavelier, D. The development of attention skills in action video game players. *Neuropsychologia* **47**, 1780–1789 (2009).
32. Franceschini, S. et al. Action video games make dyslexic children read better. *Curr. Biol.* **23**, 462–466 (2013).
33. Franceschini, S. et al. Action video games improve reading abilities and visual-to-auditory attentional shifting in English-speaking children with dyslexia. *Sci. Rep.* **7**, 5863 (2017).
34. Bavelier, D., Green, C. S. & Seidenberg, M. S. Cognitive development: gaming your way out of dyslexia? *Curr. Biol.* **23**, R282–R283 (2013).
35. Łuniewska, M. et al. Neither action nor phonological video games make dyslexic children read better. *Sci. Rep.* **8**, 549 (2018).
36. Peters, J. L., Losa, L. D., Bavin, E. L. & Crewther, S. G. Neuroscience and biobehavioral reviews efficacy of dynamic visuo-attentional interventions for reading in dyslexic and neurotypical children: a systematic review. *Neurosci. Biobehav. Rev.* **100**, 58–76 (2019).
37. Cardoso-Leite, P., Joessel, A. & Bavelier, D. in *Handbook of Game-based Learning* (eds Plass, J. et al.) Ch. 18 (MIT Press, 2020).
38. Anguera, J. A. & Gazzaley, A. Video games, cognitive exercises, and the enhancement of cognitive abilities. *Curr. Opin. Behav. Sci.* **4**, 160–165 (2015).
39. Bavelier, D. & Green, C. S. Enhancing attentional control: lessons from action video games. *Neuron* **104**, 147–163 (2019).
40. Franceschini, S. et al. Shall we play a game? Improving reading through action video games in developmental dyslexia. *Curr. Dev. Disord. Rep.* **2**, 318–329 (2015). https://doi.org/10.1007/s40474-015-0064-4

41. Raven, J. J. in *Handbook of Nonverbal Assessment* (ed. McCallum, R. S.) 223–237 (Springer, 2003).
42. Cornoldi, C. & Colpo, G. *Prove di Lettura MT per la Scuola Elementare-2* (Edizioni OS, 1998).
43. Cornoldi, C. & Colpo, G. *Prove di Lettura MT per la Scuola Media Inferiore-2* (Edizioni OS, 1995).
44. Hendriks, A. W. & Kolk, H. H. J. Strategic control in developmental dyslexia. *Cogn. Neuropsychol.* **14**, 321–366 (1997).
45. Trenta, M., Benassi, M., Di Filippo, G., Pontillo, M. & Zoccolotti, P. Developmental dyslexia in a regular orthography: can the reading profile be reduced to strategic control? *Cogn. Neuropsychol.* **30**, 147–171 (2013).
46. Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* **108**, 204–256 (2001).
47. Perry, C., Ziegler, J. C. & Zorzi, M. CDP++.Italian: modelling sublexical and supralexical inconsistency in a shallow orthography. *PLoS ONE* **9**, e94291 (2014).
48. Rakhlin, N. V., Mourgues, C., Cardoso-Martins, C., Kornev, A. N. & Grigorenko, E. L. Orthographic processing is a key predictor of reading fluency in good and poor readers in a transparent orthography. *Contemp. Educ. Psychol.* **56**, 250–261 (2019).
49. Biancardi, A. & Stoppa, E. Il test delle campanelle modificato (TCM): una proposta di studio per l'attenzione in età evolutiva. *Psichiatr. Infanz. Adolesc.* **64**, 73–84 (1997).
50. Sannio Fancello, G., Vio, C. & Cianchetti, C. *Tower of London Test* (Erickson, 2006).
51. Bisiacchi, P., Cendron, M., Gugliotta, M., Tressoldi, P. E. & Vio, C. *BVN 5-11—Batteria di Valutazione Neuropsicologica per l'Età Evolutiva* (Erickson, 2005).
52. Franceschini, S. & Bertoni, S. Improving action video games abilities increases the phonological decoding speed and phonological short-term memory in children with developmental dyslexia. *Neuropsychologia* **130**, 100–106 (2019).
53. Double, K. S., Mcgrane, J. A., Stiff, J. C. & Hopfenbeck, T. N. The importance of early phonics improvements for predicting later reading comprehension. *Br. Educ. Res. J.* **45**, 1220–1234 (2019).
54. Vidyasagar, T. R. & Pammer, K. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn. Sci.* **14**, 57–63 (2010).
55. Facoetti, A. et al. The relationship between visuo-spatial attention and nonword reading in developmental dyslexia. *Cogn. Neuropsychol.* **23**, 841–855 (2006).
56. Gabrieli, J. D. E. & Norton, E. S. Reading abilities: importance of visual-spatial attention. *Curr. Biol.* **22**, R298–R299 (2012).
57. Goldin, A. P. et al. Far transfer to language and math of a short software-based gaming intervention. *Proc. Natl Acad. Sci. USA* **111**, 6443–6448 (2014).
58. Calao, L. A., Moreno-León, J., Correa, H. E. & Robles, G. in *Design for Teaching and Learning in a Networked World* (eds Conole, G. et al.) 17–27 (Springer, 2015).
59. Blair, C. & Raver, C. C. Closing the achievement gap through modification of neurocognitive and neuroendocrine function: results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLoS ONE* **9**, e112393 (2014).
60. Holmes, J., Gathercole, S. E. & Dunning, D. L. Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* **12**, 9–15 (2009).
61. Bergman-Nutley, S. & Klingberg, T. Effect of working memory training on working memory, arithmetic and following instructions. *Psychol. Res.* **78**, 869–877 (2014).
62. Judd, N. & Klingberg, T. Training spatial cognition enhances mathematical learning in a randomized study of 17,000 children. *Nat. Hum. Behav*. **5**, 1548–1554 (2021).
63. Johann, V. E. & Karbach, J. Effects of game-based and standard executive control training on cognitive and academic abilities in elementary school children. *Dev. Sci.* **23**, e12866 (2020).
64. Spencer-Smith, M. & Klingberg, T. Benefits of a working memory training program for inattention in daily life: a systematic review and meta-analysis. *PLoS ONE* **10**, e0119522 (2015).
65. Berger, E. M., Fehr, E., Hermes, H., Schunk, D. & Winkel, K. *The Impact of Working Memory Training on Children's Cognitive and Noncognitive Skills* NHH Department of Economics Discussion Paper No. 09/2020 (SSRN, 2020).
66. Burgers, C., Eden, A., Van Engelenburg, M. D. & Buningh, S. How feedback boosts motivation and play in a brain-training game. *Comput. Hum. Behav.* **48**, 94–103 (2015).
67. Eichenbaum, A., Bavelier, D. & Green, C. S. Video games: play that can do serious good. *Am. J. Play* **7**, 50–72 (2014).
68. Martinovic, D., Burgess, G. H., Pomerleau, C. M. & Marin, C. Comparison of children's gaming scores to NEPSY-II scores: validation of computer games as cognitive tools. *Comput. Hum. Behav.* **49**, 487–498 (2015).
69. Mishra, J., Anguera, J. A. & Gazzaley, A. Video games for neuro-cognitive optimization. *Neuron* **90**, 214–218 (2016).
70. Goswami, U., Ziegler, J. C. & Richardson, U. The effects of spelling consistency on phonological awareness: a comparison of English and German. *J. Exp. Child Psychol*. **92**, 345–365 (2005).
71. Defior, S. in *Handbook of Children's Literacy* (eds Nunes, T. & Bryant, P.) 631–649 (Springer, 2004).
72. Pazzaglia, F., Mammarella, I. C., Toso, C. & Cornoldi, C. *BVS Corsi—Batteria per la Valutazione della Memoria Visiva e Spaziale* (Erickson, 2008).
73. Fleiss, J. L. Measuring nominal scale agreement among many raters. *Psychol. Bull.* **76**, 378–382 (1971).
74. Resnick, M., Maloney, J. & Hernández, A. M. Scratch: programming for everyone. *Commun. ACM* **52**, 60–67 (2009).

Acknowledgements

We thank A. Siesser for his invaluable contribution to the design and development of SOM, as well as B. Meuleman and S. Cekic for their invaluable help with statistical issues. We thank the Italian Ministry for University and Research for their support of A.d.A. through the project 'La Città Educante' PON 2007-2013; we thank the Marie Curie Grant (no. 661667, Learning Determinants) for their support of I.A., the NCCR Evolving Language as well as the Klaus J. Jacobs Foundation for their support of D.B. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Z.M., A.P., A.d.A. and P.V. contributed to the design of the video game. A.P. and P.V. designed the experiment. I.A. and D.B. provided advice on the experimental design. A.P. contributed and supervised the data collection. A.P., I.A. and D.B. performed the data analyses and wrote the manuscript. All authors read and approved the final manuscript. This study constitutes part of A.P.'s PhD thesis under the supervision of P.V.

Competing interests

D.B. is founding member and on the scientific board of Akili Interactive, Boston. A.P., Z.M. and D.B. have filled a patent which protects some of the mechanics of the video game SOM. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-021-01254-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-021-01254-x.

Correspondence and requests for materials should be addressed to Angela Pasqualotto or Daphne Bavelier.

Peer review information *Nature Human Behaviour* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

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

		SOM (n=79)		CONTROL group (n=72)		Test statistic	p
		Mean	SD	Mean	SD		
LITERACY SKILLS							
Word Reading	Speed	-0.30	0.81	-0.31	0.86	0.30[b]	0.94
	Accuracy	-0.41	0.84	-0.36	0.89	0.29[b]	0.74
Non-word Reading	Speed	-0.24	0.71	-0.28	0.76	0.28[b]	0.77
	Accuracy	-0.12	0.85	-0.06	0.74	2.05[b]	0.67
Text Reading	Speed	-0.02	0.72	-0.10	0.70	0.43[b]	0.53
	Accuracy	0.12	0.57	0.13	0.64	0.81[b]	0.92
Writing	Word	-0.70	1.98	-0.65	1.69	2.16[b]	0.86
	Non-word	-0.11	1.25	-0.11	1.17	0.06[b]	0.98
Comprehension (reading comprehension passage task)		0.23	0.73	-0,05	0.95	4.53[b]	0.04*
READING ERRORS							
Sounding-out behaviors		1.66	0.84	1.55	1.03	0.512[b]	0.48
Word-substitutions errors		1.43	0.82	1.11	0.85	5.650[b]	0.02*
ATTENTION AND PLANNING							
Barrage task	Speed	-0.15	1.03	-0.18	0.91	0.29[b]	0.84
	Accuracy	-0.14	0.86	-0.22	0.95	0.59[b]	0.59
Tower of London		0.10	0.94	-0.04	0.96	0.24[b]	0.35
OTHER EF SKILLS							
Digit span	Forward	-0.63	0.89	-0.71	0.73	1.13[b]	0.57
	Backward	-0.24	0.88	-0.22	0.87	0.06[b]	0.83
Visuo-spatial span	Forward	-0.36	1.39	-0.44	0.91	15.07[b]	0.67
	Backward	-0.26	1.15	-0.25	1.06	1.03[b]	0.95

Extended Data Fig. 1 | Means and standard deviations at T1 of neuropsychological measures for the two groups (SOM and Control). All the variables are expressed in z-scores with the exception of reading errors [that is, number of errors over the total number of enunciated words (*100)] aχ2-score, bF-score; *sig. $\alpha = 0.05$.

	SOM		CONTROL group			
	(n=79)		(n=72)		Test statistic	p
	Mean	SD	Mean	SD		
Sex (f-m)	39-40		41-31		.87[a]	.35
Age (years)	10.39	1.48	10.30	1.45	.004[b]	.67
IQ (Raven's matrices)	107.32	9.73	106.85	9.69	.06[b]	.77

Extended Data Fig. 2 | Demographic and IQ characteristics (Means, SD) of SOM group and Control group prior to intervention (T1). The two groups were carefully matched for Sex, Chronological Age, full IQ. aχ2-score,bF-score; *sig. $\alpha = 0.05$.

nature portfolio

Corresponding author(s): Angela Pasqualotto
Daphne Bavelier

Last updated by author(s): Nov 1, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No specific software was used during the data collection phase since all the tests administered in the different time points were paper-and-pencil.
Data analysis	We used Excel 2016 (Microsoft) and SPSS Statistics 16.0 (IBM) for data analysis and R 3.3.1. (https://www.r-project.org/) for plotting and for the regressions analyses. Illustrations were adjusted in Illustrator CC (Adobe).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The authors declare that all data supporting the findings of this study are available within the paper and its supplementary information. Moreover, the behavioural data that support the findings of this study, along with analysis code and study materials, have been deposited in Open Science Framework with the identifier: https://doi.org/10.17605/osf.io/4rzgc (https://osf.io/4rzgc/).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study included only behavioural data regarding children's performance in standardized paper-and-pencil tests assessing attentional control, verbal and visuo-spatial working memory, planning and literacy skills. In addition, we measured the school grades in Maths and Italian.
Research sample	151 typically developing Italian children were recruited from eight classes from 3rd, 4th, 6th or 7th grades (8-12 years old) within a northern Italy public school. The experimental group (N=79) was composed by 39 females, while 41 in the control group (N = 72).
Sampling strategy	Sample sizes were chosen on the basis of preliminary experiments so as to provide sufficient power for statistical comparison. Power analysis was not performed a priori.
Data collection	Participants were tested at school through paper-and-pencil standardized tasks. Two tasks (IQ and text comprehension) were administered in a collective session in which – in addition to the experiment – there was a teacher, whereas the other tests were administered individually in a quiet room (with only the experimenter). Testing was conducted by three different clinicians, two of them blinded to children assignments and one of them (AP) being the experimental game expert. All the researchers were aware of the study hypothesis.
Timing	Baseline (T1) assessments were completed one week prior to the start of treatment (i.e., beginning of October), and outcome assessments (T2 and T3) one week and 6-months after the end of the treatment (December and beginning of June). Grades were collected at two later time points (12-months and 18-months after the end of the training).
Data exclusions	Children with learning disorders (n=21) or with borderline intellectual functioning (n=10) participated in the study like any other classmate, however their data were not included in the analysis sample leading to a final sample of N=79, 39 females in the experimental group and N = 72, 41 females in the control group.
Non-participation	No participants dropped out from the study.
Randomization	In this study, we involved 8 classes from 3rd, 4th, 6th or 7th grades. One of the two classes included at each grade were randomly assigned to either the training or the control group.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology and archaeology
☒	☐	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data
☒	☐	Dual use research of concern

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	The participants were recruited through school newsletters or the association of the parents of the school.

Ethics oversight | University of Trento

Note that full information on the approval of the study protocol must also be provided in the manuscript.