INTRODUCTION

Adults typically integrate multisensory information in a statistically optimal fashion, that is, the information sources are weighted according to their reliabilities depending on the context, so that the combined estimates are more reliable than the best single source of information (Parise, Spence, & Ernst, 2012). Developmental studies have shown that optimal multisensory integration has a protracted time course of development, becoming mature only around the age of 8–10 years (Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Jones, Bedford, & Braddick, 2008; Petrini, Remark, Smith, & Nardini, 2014). For example, Gori et al. (2008) measured visuo-haptic integration abilities of form perception in children aged 5 to 10 years by asking them to discriminate the height of two blocks presented in sequence and found that only older children were able to optimally integrate visuo-haptic information, while younger children relied more on haptic cues. Similar findings were found in comparable age groups using different tasks (Gori, Sandini, & Burr, 2012; Nardini et al., 2008), thus suggesting that lack of multisensory integration may represent a distinctive (cognitive) feature before age 10. This late multisensory development has been interpreted as a trade-off between integration benefits and the child's body size and sensory systems continuously recalibrating and adjusting (Burr & Gori, 2012). This weighting of the sensory systems is indeed a signature of the heightened plasticity of development, learning, multisensory integration, plasticity, video games, visuo-spatial skills
the developing brain, but to date, no study has addressed to what extent it is possible to “push the boundaries” of such plasticity.

Here, we assessed this issue by training children aged 4–5 years with a short action-like mini game training for 2 weeks (45 min/day). Children were assigned to one of the following training groups: action video games (AVG), non-action video games (NAVG). A third group (CTRL) served as control group to assess re-test effects.

The use of video games was motivated by the growing literature documenting striking changes in adult action-video-game players (Bavelier, Green, Pouget, & Schrater, 2012; Green & Bavelier, 2003). In particular, these studies show behavioural improvements in several cognitive domains, mostly involving visuo-spatial attention, but also memory, and executive control (Bediou et al., 2018). These skills appear to improve in children too following action video games play, as shown in studies conducted in typically (Dye & Bavelier, 2004), as well as atypically developing populations (Franceschini et al., 2013; Gambacorta et al., 2014). In particular, the study of Franceschini et al. (2013) has shown that exposing children with dyslexia for only 12 hr to action-like mini games can dramatically improve their reading skills, suggesting a possible transfer of learning across very different skills that are mediated by attention.

Interestingly, there appears to be a bidirectional interplay between attention and multisensory processing (Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). For example, temporally and spatially aligned multisensory inputs have a higher probability to capture one individual’s attention, as shown in infant studies revealing that redundant and synchronous multisensory stimuli recruit attention and facilitate perceptual differentiation more than presenting the same information from a single modality (Bahrick & Licklider, 2000). However, also top-down attention facilitates the integration of multisensory stimuli, as shown by behavioural and neuroimaging studies that have identified several higher level factors, such as voluntarily oriented spatial attention and semantic congruency, which can influence whether and how integration across senses occurs (Welch, 1999).

In this study, by taking advantage of the benefits of video games on cognitive improvements (Bavelier et al., 2012; Franceschini et al., 2013), we investigated whether multisensory integration and visuo-attentional skills can be enhanced following a short-action-like mini game training in children aged 4–5 years using a simplified version of the paradigm adopted by Gori et al. (2008), and a multiple object tracking task (MOT) adapted from a study by Green and Bavelier (2006). Because video games are multisensory in nature (i.e. players use tactile, visual and auditory information to achieve their purpose during playing), we predicted improvements in multisensory integration and attention skills particularly in the group of children trained with action-like mini games. The reason is that in action video games, the feedback provided by the sensory modalities is more relevant to the aim of the game (i.e. winning), in comparison to non-action video games. In the latter, players are not pressured by time, thus paying attention to the feedback provided by the concurrent presentation of the sensory modalities is not strictly necessary to the aim of the game. That is, even though both action and non-action video games possess multisensory features, it is only in the action video games that multisensory feedback is functional to achieve a goal.

### 2 | METHOD

#### 2.1 | Participants

Nineteen males and 22 females (mean age = 5.2, range 4.5–6.2 years) were recruited to take part in this study from kindergartens around the city of Milan (Italy). All children had no cognitive or neurological impairment. All but four children were right-handed, and five wore glasses. The parents of all children gave their informed consent before the experiment started. The study was approved by the ethical committee of the University of Milano-Bicocca, in accord with the Declaration of Helsinki.

Each child was assigned to one of the following three groups: action-like video games (AVG), non-action-like video game (NAVG) and controls (CTRL). In order to obtain homogeneous groups (i.e. groups of children with comparable experience playing video games), we administered a questionnaire to the parents assessing the familiarity of their child with video games. The questionnaire was developed by the experimenters in order to address the familiarity of the children with video games and the amount of time spent playing (see Supplementary Material). The questionnaire simply assessed whether and how integration across senses occurs.

We then labelled the children according to the overall time spent per day playing video games: “Intense Players” (N = 9, 30–60 min a day), “Medium Players” (N = 14, maximum 30 min a day) and “Non-players” (N = 16, no experience playing video games).
The training groups were then balanced, so that an equal number of intense, medium and non-players were assigned to each group: $N = 14$ children were assigned to the AVG group (4 intense players, 5 medium players and 5 non-players), $N = 13$ children were assigned to the NAVG group (3 intense players, 5 medium players and 5 non-players), $N = 14$ children were assigned to the CTRL group (4 intense players, 4 medium players and 6 non-players). Table 1 reports a summary about the number of children assigned to each group.

<table>
<thead>
<tr>
<th>ID</th>
<th>Training group</th>
<th>Age at testing (years)</th>
<th>Age first exposure to video games (years)</th>
<th>Average time spent playing (minutes a day)</th>
<th>Category assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>AVG</td>
<td>4.8</td>
<td>4</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>15</td>
<td>AVG</td>
<td>5.4</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>24</td>
<td>AVG</td>
<td>4.8</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>25</td>
<td>AVG</td>
<td>5.1</td>
<td>4</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>26</td>
<td>AVG</td>
<td>5.3</td>
<td>4</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>29</td>
<td>AVG</td>
<td>5.4</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>30</td>
<td>AVG</td>
<td>5.4</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>31</td>
<td>AVG</td>
<td>5.2</td>
<td>4</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>32</td>
<td>AVG</td>
<td>5.0</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>37</td>
<td>AVG</td>
<td>5.2</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>38</td>
<td>AVG</td>
<td>5.2</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>47</td>
<td>AVG</td>
<td>5.2</td>
<td>4</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>48</td>
<td>AVG</td>
<td>5.2</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>53</td>
<td>AVG</td>
<td>4.8</td>
<td>2</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>6</td>
<td>NAVG</td>
<td>4.8</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>10</td>
<td>NAVG</td>
<td>5.0</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>16</td>
<td>NAVG</td>
<td>4.7</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>19</td>
<td>NAVG</td>
<td>5.0</td>
<td>4.5</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>22</td>
<td>NAVG</td>
<td>5.4</td>
<td>4.5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>33</td>
<td>NAVG</td>
<td>5.2</td>
<td>4.5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>35</td>
<td>NAVG</td>
<td>5.3</td>
<td>4.5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>39</td>
<td>NAVG</td>
<td>5.4</td>
<td>5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>42</td>
<td>NAVG</td>
<td>4.7</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>45</td>
<td>NAVG</td>
<td>5.2</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>46</td>
<td>NAVG</td>
<td>4.8</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>50</td>
<td>NAVG</td>
<td>4.9</td>
<td>4</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>55</td>
<td>NAVG</td>
<td>4.5</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>5</td>
<td>Control</td>
<td>5.4</td>
<td>4.5</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>9</td>
<td>Control</td>
<td>5.1</td>
<td>4</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>7</td>
<td>Control</td>
<td>5.6</td>
<td>5</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>8</td>
<td>Control</td>
<td>5.3</td>
<td>5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>12</td>
<td>Control</td>
<td>5.6</td>
<td>5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>13</td>
<td>Control</td>
<td>5.7</td>
<td>5</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>14</td>
<td>Control</td>
<td>5.4</td>
<td>4</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
<tr>
<td>17</td>
<td>Control</td>
<td>5.6</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>18</td>
<td>Control</td>
<td>5.4</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>20</td>
<td>Control</td>
<td>5.7</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>21</td>
<td>Control</td>
<td>5.5</td>
<td>4</td>
<td>30–60</td>
<td>Medium player</td>
</tr>
<tr>
<td>23</td>
<td>Control</td>
<td>6.2</td>
<td>5</td>
<td>5–30</td>
<td>Medium player</td>
</tr>
<tr>
<td>27</td>
<td>Control</td>
<td>5.9</td>
<td>Never</td>
<td>0</td>
<td>Non-player</td>
</tr>
<tr>
<td>28</td>
<td>Control</td>
<td>6.0</td>
<td>4.5</td>
<td>30–60</td>
<td>Intense player</td>
</tr>
</tbody>
</table>
The sample size was determined based on previous studies using a Bayesian approach to the data analysis (see Gori et al., 2008). Note that, in comparison to Gori et al. (2008), each group in our study included over 10 children in order to achieve more power and a total of 41 children between 4 and 6 years of age.

2.2 | Description of the tasks

2.2.1 | Multisensory task

The multisensory size-discrimination task was adapted from Gori et al. (2008) and consisted in presenting children with pairs of wooden blocks of fixed width (97 mm) but variable height, ranging between 45 and 65 mm (with 2 mm increments, see Figure 1). The blocks were attached on a vertical wooden surface, which was designed to resemble the original set-up of Gori et al. (2008).

All children performed an overall 60 trials, divided into three conditions: unimodal visual, unimodal haptic and visuo-haptic. In the unimodal visual condition, the experimenter presented for about 3 s the first block in front of the child, sit at about 45 cm from the stimulus at eye-level. To prevent the child from seeing the experimenter changing the block, the experimenter covered the child’s eyes with the hand. Then, the child had another c. 3 s to look at the block. The experimenter then asked which of the two blocks was taller, and the response of the child (i.e. “first” or “second” responses) was manually recorded by the experimenter.

In the unimodal haptic condition, the blocks were attached behind the wooden surface facing the child, so that she/he could only touch without seeing them. The procedure was the same as for the unimodal visual condition, with the only difference that children could keep their eyes open throughout the testing.

In the visuo-haptic condition, children were presented with two blocks at the same time, one that they could see, and one that they could touch of same size.

The conditions were performed in separate sessions and were counterbalanced across children.

In all conditions, the standard block was always 55 mm high, and was presented either first or second with respect to the other block. The task took overall 30 min to complete.

2.2.2 | MOT task

In this task, children had to keep track of the spatial position of dynamic target stimuli, embedded among dynamic distractors. The target stimuli consisted in 0.7 cm diameter blue cartoon faces, with a sad facial expression, which appeared among 12 distractors, represented by yellow happy faces (see Figure 2). The stimuli appeared within a 14 cm diameter grey circle on a portable 15-inches HP laptop. The target and distractors moved randomly within the circle at a velocity of c. 5°/sec.

The trial was always started by the experimenter, who pressed a button on the keyboard. The 12 distractors appeared, followed shortly after by the target stimulus that remained visible for 2 s. During this time (Cueing Phase), children were instructed to keep track of the target stimulus until it turned into a yellow face too, thus leaving all faces indistinguishable (Tracking Phase). In the tracking phase, children were asked to track the previously marked blue target items for a duration of 5 s. Once the stimulus had changed into a distractor, a “?” appeared on the screen, and children had to tell whether in that specific spatial location there was a target before. Children had two buttons of the keyboard that could be pressed: one that corresponded to a “yes”, and one to a “no” response in relation to the target stimulus. The next trial would start only after the child’s response; thus, the experimenter encouraged the participants to enter a response even if they were uncertain. On average, children took a maximum of 5 s to respond.

The task followed a staircase method, by which following three consecutive correct responses, the number of target stimuli would increase by one stimulus, up to a maximum of seven stimuli. In other words, following three consecutive correct responses provided by the child, the number of target stimuli would increase to two; after another three correct responses the number of target stimuli would increase to four, and so on. This staircase converges to 65% in terms of performance (see Brown, 1996, Table 1).

If the child gave two incorrect responses in a row, the number of target stimuli would decrease by one stimulus. The threshold would be reached after five reversal (i.e. the change in number of target stimuli presented), or after a maximum number of 40 trials. The task lasted about 15 min, and the overall experimental session for each child lasted approximately 1 hr, including at least one break between the two tasks.
2.3 | General procedure

The tasks were administered before and after the 2-weeks training. Note that the children only played during week-days, so overall 10 days. Both tasks were administered to each child and they were counterbalanced among children, so that half of the children started with the multisensory task, and the other half with the MOT task.

The testing was conducted by two expert experimenters, who administered them inside the kindergartens which provided a quiet room. The first testing lasted one week, after which the training immediately started.

The training consisted in the AVG and NAVG participants playing for two consecutive weeks with video games of different types (see below for a list of video games used in both trainings) for 45 min/day. The training was supervised by the two expert experimenters, who made sure that the children played with the video games for the exact amount of time, on the specific video games assigned to each group (see below the list of video games used).

The control group only served to control for re-test effects and thus did not make any training between the two experimental sessions.

Following the training, the two tasks were administered again, and again it took ca. one week to test all children.

2.3.1 | List of video games played

The video games were classified as action or non-action mini games according to the characteristics provided by Green Li and Bavelier (2010), who classified action video as all sharing a set of qualitative features, including “extraordinary speed (both in terms of very transient events and in terms of the velocity of moving objects), a high degree of perceptual, cognitive and motor load in the service of an accurate motor plan (multiple items that need to be tracked and/or kept in memory, multiple action plans that need to be considered and quickly executed typically through precise and timely aiming at a target), unpredictability (both temporal and spatial), and an emphasis on peripheral processing (with important items most often appearing away from the centre of the screen)”. It is important to note that all the video games we used provided auditory, visual and tactile feedback; that is, children actively touched the tablet, watched the scenes and heard auditory tones.

On the contrary, non-action video games are those that do not require speed (i.e. objects do not move fast) and are not necessarily presented in the periphery. Non-action video games are more accuracy-based and time does not pressure the participants. Most importantly though, we selected the video games so that none contained violent scenes or actions. In other words, in none of the selected video game there was any shooting, blood, death or anything that could remind the hurting of individuals. All video games were played on an iPad during the training.

For the action-like mini games training we used the following games:

FRUIT NINJA, SMASH MASTER, NINJUMP, RAIL RUSH, BRICKS BREAKER RE, BANANA KONG, PJ. MASKS: MOONLIGHT HEROES, FRUIT SMASHER, LEP’S WORLD 3, LIGHTNING McQUEEN, TRAIN CONDUCTOR WORLD, ROLLING SKY, MINION RUSH, JUMP FEVER.

For the non-action-like mini games training, we used the following games:

PUZZINGO, BUBBLE SHOOTER, DRAWING for KIDS, KIDS MAZES, MEMORY GAME, DOCTOR KIDS, ANGRY BIRDS, ANIMALS PUZZLE, FROSTY SHOT, DIRTY FARM, PLAYING CARDS for KIDS, FIND THE DIFFERENCE, COLOURING BOOKS, MASHA and the BEAR.

3 | RESULTS

3.1 | Multisensory task

Several studies have shown that adults integrate visual and haptic information (and information from other modalities) in a statistically optimal fashion, weighting each sense according to its reliability (Deneve & Pouget, 2004; Ernst & Banks, 2002; Parise et al., 2012). The maximum likelihood estimate (MLE) predicts as main signature for cross-modal integration that the thresholds for dual-modality presentation is lower than either visual or haptic thresholds. This model, described below, is a perfect way to quantify the level of multisensory integration in children. Since Gori et al. (2008) demonstrated that MLE predicts visual-haptic multisensory integration only after 8–10 years of age, we used this model to quantify the improvement of multisensory integration after the training with video games. To make the MLE calculation, we first extracted the threshold for each unimodal condition (visual and haptic alone) by
calculating the proportion of trials in which the child responded that
the standard stimulus was taller than the other variable stimuli and
by fitting for each child the data into a cumulative Gaussian function,
in which the obtained standard deviation estimated the discrimina-
tion threshold.

To assess optimal integration, we compared the discrimination
thresholds with the threshold predicted by the Maximum Likelihood
Estimation model before and after the training. The MLE calculation
assumes that for size judgments, the optimal bimodal MLE pre-
diction for the visuo-haptic threshold ($\sigma_{VH}$) is given by:

$$\sigma_{VH}^2 = \frac{\sigma_V^2 \sigma_H^2}{\sigma_V^2 + \sigma_H^2} \leq \min \left( \sigma_V^2, \sigma_H^2 \right)$$

where $\sigma_V$ and $\sigma_H$ are the visual and haptic unimodal thresholds ex-
tracted from the psychometric function. The improvement is greatest ($\sqrt{2}$) when $\sigma_V = \sigma_H$. The prediction of the model is good when the
bimodal threshold predicted from the unimodal estimates as per the
equation above is not different from the measured bimodal thresh-
old. The critical test for optimal multi-sensory integration is that the
measured bimodal threshold be as good as the MLE estimate of the
bi-modal threshold, derived from optimally combining the unimodal
performances.

Following this model, we compared the thresholds in all modal-
ties to the MLE prediction by means of paired $t$ tests, and found that
in the visuo-haptic modality, before starting the training, all groups
of children presented higher thresholds than the MLE prediction
(all $p < 0.003$), suggesting no optimal multisensory integration ca-
pabilities. However, after the training, a specific improvement was
observed in the group of children that performed the action-like
mini-game training. Indeed, while the NAVG $t(12) = 2.39, p = 0.03$
and CTRL group $t(13) = 4.45, p < 0.001$ still showed higher mea-
sured threshold than the MLE prediction, the AVG group showed
optimal multisensory integration, as their measured thresholds did
not differ from the MLE prediction ($p = 0.10$).

The comparison of the haptic and visual modalities alone with the
MLE prediction showed that all groups presented higher thresholds
than the MLE prediction before and after the training (all $p < 0.003$).

We also conducted a repeated measures ANOVA on the discrimi-
nation thresholds of the three groups of children, with Modality
(haptic alone, visual alone and visuo-haptic) and Session (pre- and
post-training) as within-subjects factors, and Group (AVG, NAVG and
controls) as between-subjects factor.

The analysis revealed a main effect of Session [$F(1, 38) = 11.94, p = 0.001$] and Modality [$F(2, 76) = 14.13, p < 0.001$],
a Session × Group interaction [$F(2, 38) = 4.02, p = 0.03$], due to
the AVG group performing better after the training ($p = 0.003$),
while the other two groups did not show any difference between
sessions (both $p > 0.99$). There was also a Modality × Group [$F(4, 76) = 3.46, p = 0.01$], and a Session × Modality interaction [$F(2, 76) = 3.48, p = 0.04$]. To further explore whether the performance
on each modality changed across groups, we conducted other
ANOVAs on each modality. The analysis conducted on the visuo-
haptic thresholds showed a main effect of Group [$F(2, 38) = 6.96, p = 0.003$], a main effect of Session, [$F(1, 38) = 27.97, p < 0.001$],
and a Group × Session interaction [$F(2, 38) = 5.81, p = 0.006$], the
latter caused by children only in the AVG group improving their
visuo-haptic discrimination abilities following training ($p < 0.001$).
on Bonferroni post-hoc test). On the contrary, the performance of the other two groups did not vary following either the non-action video game training ($p = 0.38$) or the re-test ($p = 0.99$, see Figure 3a, upper left panel for a summary of results, and Figure 4 for psychometric functions of the whole three groups).

Furthermore, analysis on both the haptic and visual modality alone did not reveal any main effect or interaction in any group and in any session (all $p > 0.14$). That is, the training specifically improved multisensory integration abilities, but not the performance in the single modality.

3.2 | MOT task

Because less than 50% of the children reached the threshold, for each child we calculated the average number of target stimuli she/he could track on each trial. This average was calculated on the correct responses only. Thus, the mean number of tracked objects was entered as dependent variable in a repeated measure ANOVA, with the within subject factor Session and the between subject Group as main factors. The analysis revealed a main effect of Session [$F(1, 38) = 24.99, p < 0.001$], Group [$F(2, 38) = 3.25, p = 0.04$], and a Session × Group interaction [$F(2, 38) = 9.24, p < 0.001$], caused by the AVG group increasing the average number of object tracked following the training (pre-training: $M = 1.86, SD = 0.58$; post-training: $M = 3.04, SD = 0.59, p < 0.001$) in comparison to the NAVG (pre-training: $M = 1.88, SD = 0.53$; post-training: $M = 2.34, SD = 0.72, p = 0.39$) and the control group (pre-training: $M = 2.08, SD = 0.40, p = 0.99$). Importantly, no difference across the three groups emerged in the testing conducted before the training (all $p = 0.99$, see Figure 3b, upper panel for a summary of results).

To observe whether changes in visuo-spatial attentional skills could subserve changes observed in multisensory integration, we performed correlations between the two tasks. This was done by subtracting the mean performance in the post-training phase to the pre-training phase performance separately for group. However, in all groups, we did not find any correlation between the two tasks (all $p > 0.22$ on two-tailed Pearson correlation).

4 | THREE MONTH-FOLLOW-UP

4.1 | Participants, tasks and procedure

To assess the stability of the changes observed selectively in the AVG group, we conducted a follow-up in this same group of children after 3 months from the end of the training.

Only $N = 13$ children of the AVG group took part, because one child could not participate. This child was removed from the analyses aimed at comparing the three experimental sessions (pre-training, post-training and follow-up). Note that the questionnaire regarding the familiarity with video game use was also administered again to the families of the children, to make sure that their playing pattern did not vary within these 3 months. Indeed, they confirmed that the pattern did not vary in any child.

The stimuli, tasks and procedure were identical to the other sessions. The children were again tested in the same kindergarten they were tested in the previous sessions.

5 | RESULTS OF THE 3-MONTH FOLLOW-UP

The analyses conducted in the follow-up revealed that the measured threshold of the AVG in the follow-up did not differ from the MLE prediction ($p = 0.34$), thus showing optimal multisensory integration. Furthermore, to observe differences across sessions, we performed a repeated measure ANOVA with Session (pre-training, post-training and 3-month follow-up) as only factor. This analysis revealed a main effect of Session [$F(2, 24) = 13.95, p < 0.001$], caused by differences in visuo-tactile integration between the pre-training and the post-training ($p < 0.001$), the pre-training and the follow-up ($p = 0.01$), but not between the post-training and the follow-up ($p = 0.13$). That is, the improvement in multisensory integration was maintained after 3 months following the end of the training (see Figure 3a, lower panel for a summary of results).

Stability of changes were also observed in the MOT, though less robust than in the multisensory task. That is, we found a main effect of Session [$F(2, 24) = 19.47, p < 0.001$], caused by a difference in
performance between the pre- and post-training \((p < 0.001)\), and between the pre-training and follow-up \((p = 0.01)\). Note however, that the follow-up differed from the post-training \((p = 0.02)\), suggesting that after 3 months, visuo-spatial enhancement was slightly decaying \((\text{follow-up: } M = 2.47, SD = 0.54, p = 0.01, \text{ see Figure 3b, lower panel for a summary of results})\).

6 | DISCUSSION

The results of this study show that a short training with action-like mini games can promote optimal integration of multisensory signals during a particular developmental phase in which this ability has been reported to be sub-optimal.

This suggests that selectively training multisensory abilities aids the developing brain to better combine redundant multisensory estimates into a coherent multimodal percept.

In line with previous studies (Green & Bavelier, 2012; Green, Pouget, & Bavelier, 2010), our findings suggest that the enhanced visuo-spatial attention and multisensory processes seen in children in the AVG group may reflect a change in learning strategies. It should be discussed which mechanisms may be responsible for the changes observed. That is, did the training promote better multisensory binding, or did the children learn to better estimate the sensory weights? While our data cannot provide a conclusive answer to these questions, two aspects should be noted. First, it might be that children learned (or improved) the correspondence between the sensory signals. That is, the perceptual system has to determine whether two signals belong to the same object or event before deciding whether to integrate them or not. Thus, it could be that children at this age are yet not able to causally infer that there is a connection between the two senses, and the training may have improved such correspondence.

It could also be claimed that instead of being a purely binding problem, children may have learned to better estimate the sensory weights; this, in turn, might have favoured optimal integration. Optimal integration, indeed, means that signals from unimodal cues are integrated in proportion to their reliability as predicted by Bayes's rule. Estimating the sensory weights may however be a very difficult task for children, as the sensory systems involved with spatial perception must recalibrate continuously during development because of the body rapidly changing in size. As suggested by previous studies (Burr & Gori, 2012; Gori, 2015; Gori et al., 2008), it may be that for the developing child, calibration is more important than optimizing perception by integration; and if sensory information is integrated, one sense cannot be used to calibrate the other. The training with the action-like mini games may have facilitated the binding of information between haptic and visual information but the effect does not appear to be related to a change in the precision of unimodal inputs. Indeed, we did not observe any difference in the precision of the haptic and visual systems before and after the training but only a specific change in the bimodal condition after the training in the group trained with the action video games.

Furthermore, even though our findings may suggest that multisensory integration improvements go hand in hand with visual attention enhancement, it should be noted that we did not find any correlation between the two tasks, which in turn suggests that the improvements observed in the two tasks may reflect different mechanisms underlying attention and multisensory learning skills. In particular, it could be that learning to optimally integrate the senses could occur in a very automatic fashion, without any attentional requirements.

Most importantly, we also found that these plastic changes have long-term effects, at least lasting up to 3 months, and are observed both at a multisensory integration and visuo-spatial level. Importantly, the fact that performance in the MOT after 3 months was less robust than in the multisensory task corroborates the idea that, while multisensory processes are fully automatic and once the association is optimally established it cannot be reversed, attentional processes are more top-down and likely need to be continuously trained in order to be maintained.

It should be acknowledged that, for practical issues, the experimenters who conducted the training were not blind to the purpose of the study. Even though the experimenters were highly trained and were explicitly told that any outcome would have contributed to the knowledge about brain plasticity in children, we cannot completely exclude a slight bias.

In conclusion, these training-induced, long-term plastic changes in the developing brain suggest that, first, multisensory integration is not under maturational constrains, that is, it is not the growing body—which has to recalibrate all the sensory signals in the first years of life—that imposes limits to multisensory development. It rather appears that multisensory integration is protracted throughout late childhood because the environment does not typically impose high demands on the child's brain. However, if the child's brain is given the opportunity to systematically train an ability, it rapidly shows adult-like patterns of multimodal cue integration.

The fact that children can rapidly learn to recalibrate their sensory systems has dramatic implications for all those conditions in which multisensory integration appears pathologically impaired. This is particularly true for autistic (Brandwein et al., 2012), but also for dyslexic children, who have recently been documented to have impaired multisensory integration capabilities (Harrar, Tammam, Pérez-Bellido, Stein & Spence, 2014). Thus, training with action-like mini games could prove a useful, fast and fun way to promote long-lasting plastic multisensory and attentive changes in these populations.

ACKNOWLEDGEMENTS

We thank Martina Solbiati and Federica Turazzi for their amazing help in recruiting and testing the children involved in the study. We also thank all children and their families for their willing participation in this study.

CONFLICT OF INTEREST

All authors declare no conflict of interest.
DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author (elena.nava@unimib.it). Please note that in the Figures S1–S3, all data from each participant in the Multisensory Task are plotted as a function of Time (Pre- and Post-Training) and Condition (Visual, Tactile and Bimodal).

REFERENCES


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Nava E, Föcker J, Gori M. Children can optimally integrate multisensory information after a short action‐like mini game training. Dev Sci. 2019;e12840. https://doi.org/10.1111/desc.12840.