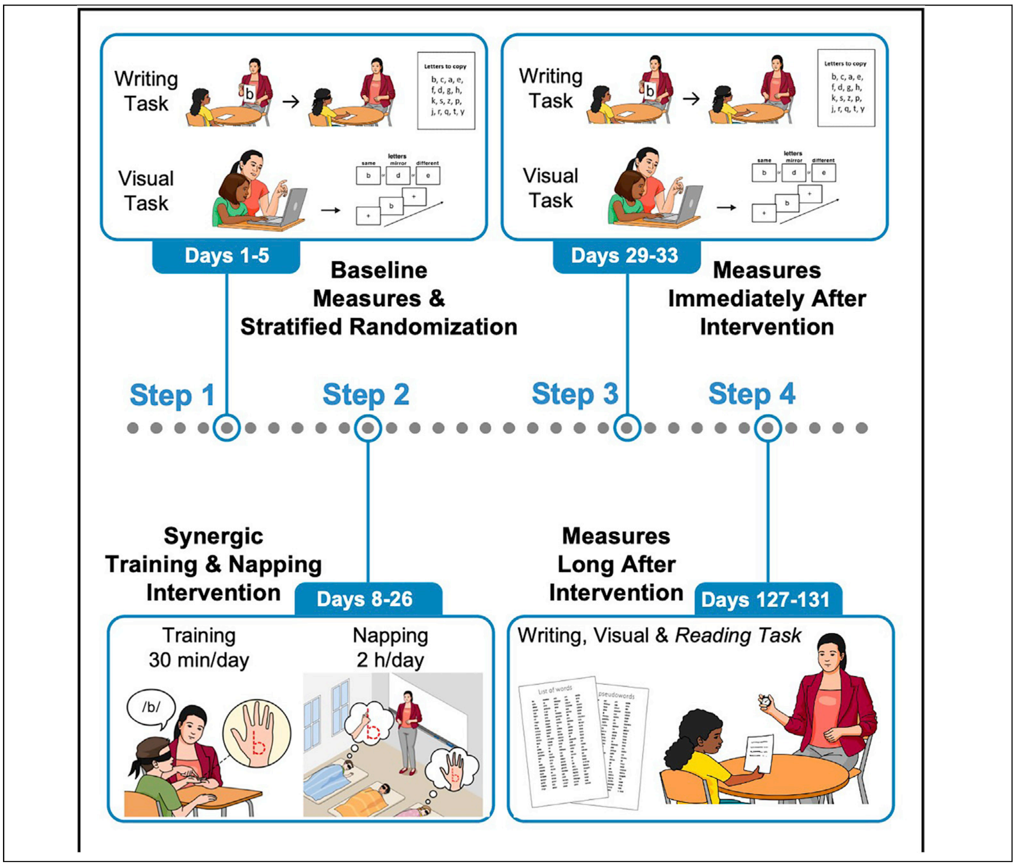
STAR Protocols

CellPress OPEN ACCESS

Protocol

A protocol to examine the learning effects of 'multisystem mapping' training combined with post-training sleep consolidation in beginning readers



We have recently used randomized controlled trials to examine the impact of a short neuroscience-informed targeted training combined with post-training sleep. Using this training protocol, we have shown unprecedented improvements in visual perception of letters, writing, and reading fluency in first graders. Here, we describe this ecologically valid school-based intervention protocol to probe inhibition of mirror invariance for letters, including the detailed training instructions, post-training sleep consolidation, as well as practical tips and potential adaptations to different school sizes.

Felipe Pegado, Ana Raquel Torres, Janaina Weissheimer, Sidarta Ribeiro

felipe.pegado@univ-amu.fr (F.P.)
torres.ar@neuro.ufrn.br (A.R.T.)
janaina.weissheimer@gmail.com (J.W.)
sidartaribeiro@neuro.ufrn.br (S.R.)

Highlights

School-based protocol to probe inhibition of mirror invariance for letters

Gains in visual perception of letters, writing and reading fluency in first graders

Learning effects sustained for 4 months when sleep followed targeted training

Simple, ecological and cost-effective 3 weeks intervention to double reading fluency

Pegado et al., STAR Protocols 2, 100712
September 17, 2021
https://doi.org/10.1016/j.xpro.2021.100712

STAR Protocols

CellPress
OPEN ACCESS

Protocol

A protocol to examine the learning effects of 'multisystem mapping' training combined with post-training sleep consolidation in beginning readers

Felipe Pegado,[1,5,*] Ana Raquel Torres,[2,*] Janaina Weissheimer,[2,3,*] and Sidarta Ribeiro Ph.D.[2,4,*]

[1]Laboratory of Cognitive Psychology, Institute for Language Communication and the Brain, CNRS and Aix-Marseille University- 3 place Victor Hugo, 13331 Marseille, France

[2]Laboratory of Memory, Sleep and Dreams, Brain Institute, Federal University of Rio Grande do Norte, Av. Senador Salgado Filho, 3000, Campus Universitário, Lagoa Nova - 59078-970

[3]Department of Modern Foreign Languages and Literatures, Federal University of Rio Grande do Norte, Av. Sen. Salgado Filho s/n, 59078-970 Natal, Brazil

[4]Lead contact

[5]Technical contact

*Correspondence: felipe.pegado@univ-amu.fr (F.P.), torres.ar@neuro.ufrn.br (A.R.T.), janaina.weissheimer@gmail.com (J.W.), sidartaribeiro@neuro.ufrn.br (S.R.)
https://doi.org/10.1016/j.xpro.2021.100712

SUMMARY

We have recently used randomized controlled trials to examine the impact of a short neuroscience-informed causal intervention using a targeted training to inhibit a deeply rooted visual mechanism (mirror invariance) that hinders literacy acquisition, combined with post-training sleep (for learning consolidation). Using this training protocol, we have shown unprecedented improvements in visual perception of letters, writing, and a two-fold increase in reading fluency in first graders. Here, we describe this ecologically valid school-based intervention protocol to probe inhibition of mirror invariance for letters, including the detailed training instructions, post-training sleep consolidation, as well as practical tips and potential adaptations to different school sizes.
For complete details on the use and execution of this protocol, please refer to Torres et al., (2021).

BEFORE YOU BEGIN

Here we present a school-based randomized controlled trial protocol that was able to consistently reveal the impact of a synergic neuroscience-based intervention to selectively inhibit the expression of mirror invariance for letters (i.e., confusion with letter's orientation, e.g., b = d), with a consequent two-fold increase in reading speed in first graders (Figure 1). The learning effects were sustained for at least 4 months, but only when sleep followed the training (Torres et al., 2021). Below we describe the specificities of the 3-week intervention, including the targeted training (30 min per weekday of multisensory-motor activities aiming to create 'multisystem mappings' for letters), post-training sleep consolidation (for up to 2h post-lunch naps), and design choices.

Experimental and control groups

To causally measure the impact of a given intervention, it is important to have experimental groups that receive the treatment (i.e., intervention) and the control groups that do not receive it. Note that even two (or eventually more) intervention strategies can be tested in *combination*, including to probe *synergic* effects (in our case the two treatements were: mirror discrimination training + naps consolidation). If the proposed intervention has a significant effect on performance measures, then changes in performance are expected only in the experimental groups, not in the control groups.

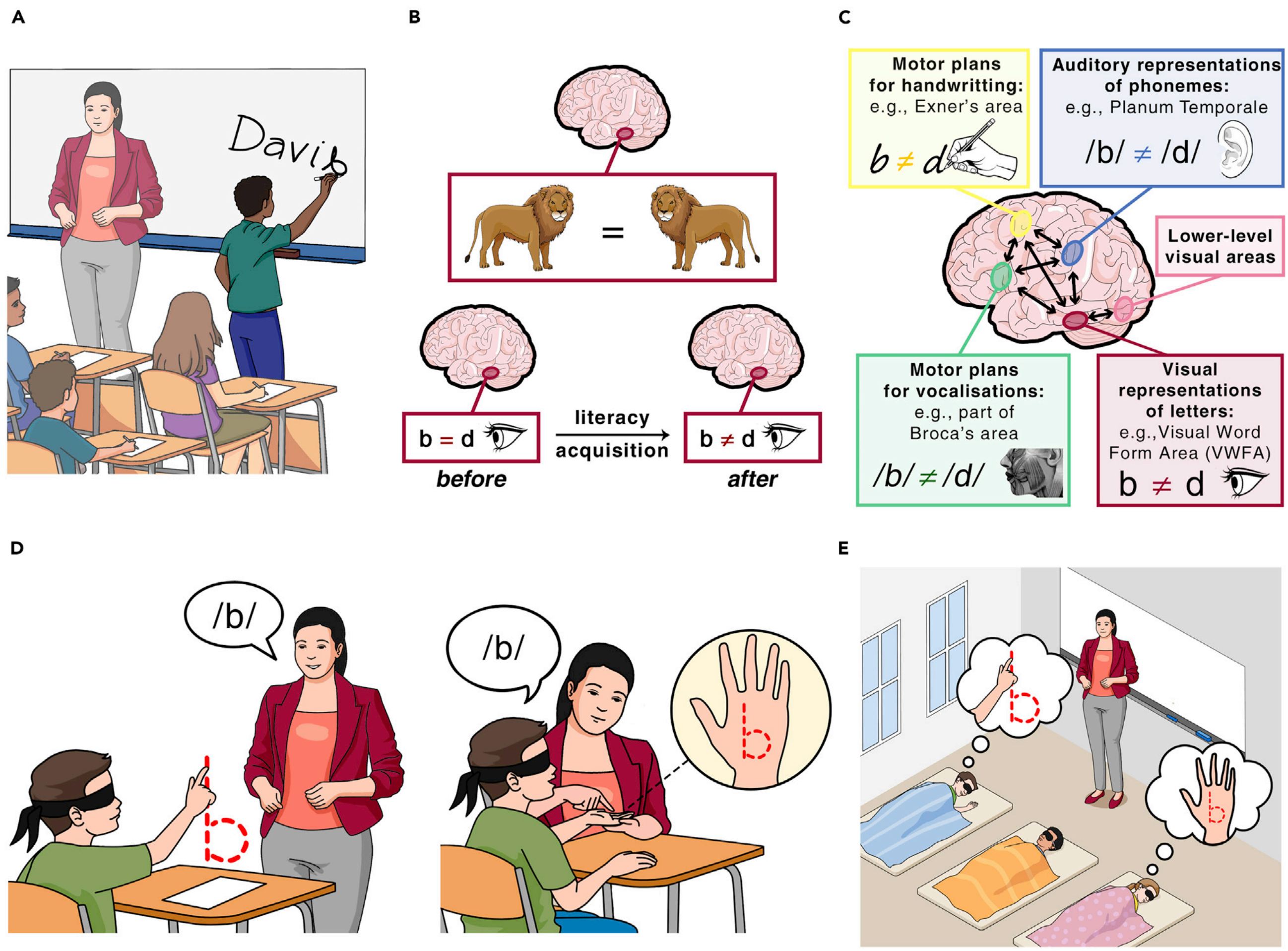


Figure 1. Overview of the study rationale

(A) Confusion with the correct orientation of letters is common among first graders, with writing errors being paralleled by visual perception errors.

(B) Mirror invariance in high-level visual cortex is suspected to be the main source of mirror confusion for letters (Dehaene et al., 2010; Pegado et al. 2011) and could hamper reading fluency acquisition.

(C) Our hypothetical learning model of how mirror invariance for letters is slowly overcome in school in a non-systematic way (via multisensory-motor mappings) [modified from Pegado et al. 2014a).

(D) Three randomized controlled trials (RCTs) were performed to causally probe the impact of a brief targeted training intervention (30 min/day for 3 weeks), based on our model (maximazing multisystem mappings), on mirror invariance for letters and on reading fluency. Multisensory-motor activities were used, such as "air-writing" (left) and tactile perception of letter traces (right), while listening to and/or producing letter sounds.

(E) To test whether sleep could improve learning consolidation of the targeted training (Stickgold et al, 2000; Marshall et al. 2006; Lemos, Weissheimer and Ribeiro, 2014; Cabral et. al, 2018; Cousins et al. 2019), one group of participants (T+S group) took post training naps (up to 2 h). Figure reprinted with permission from Torres et al., 2021.

⚠ CRITICAL: You must ensure that there is equivalent baseline performance across groups. It is important to ensure that the experimental and control groups have equivalent performances at *baseline*. This prevents potential confounds, which can make unclear the interpretation of the results. If baseline performance differs across groups, e.g., the experimental group is better already at baseline, this initial difference could explain why only this group improved after the intervention, since those with better initial skills can be expected to learn faster. Note however that baseline differences could be problematic even in the opposite situation: if the experimental group shows worse performance at baseline, post-intervention changes in performance only in the experimental group could be attributed to 'ceiling effects' in the control group, i.e. more "room for improvement" in the experimental group.

Thus, it is important to have equivalent performances across groups at baseline. To achieve this goal, it is recommended to assign participants to groups only *after* (not before) the analysis of baseline performances. We used a combined measure (average) of the two tasks (error rates at the visual and writing tasks) recorded at baseline to rank the participants. This allows the consideration of different performance *strata* (for instance high, medium, and low performers). Then, for each stratum, participants should be randomly assigned to the experimental and control groups, so that high performance participants were equally distributed across groups (and the same applied to medium and low performers). This procedure ensures that each group receives a balanced *mix* of the different performance strata, resulting in similar average performances at baseline across experimental and control groups. Finer-grained stratification can also be used. In any case, the number of children across groups should be equivalent.

After this initial randomization, the groups should be differentially treated as following: a multisensory-motor training group using *mirror-letter* pairs such as b and d (T = Training); a group receiving the same training with mirror-letters being followed by post-training naps (T+S = Training + Sleep); a classical no-training control group (C = Control), and an additional active control group (AC = Active Control) that receives the same multisensory-motor training as the previous one but playing only with *symmetrical letters*, thus with no mirror discrimination learning involved. Note that during the three weeks intervention period, children in the training groups (T and T+S) and the AC group will not participate in the regular activities of the school during the 30-min duration of the training, but instead will "play" the multisensory-motor games in restricted groups (including colleagues from different classes). Thus, while controlling for unspecific factors (socio-emotional), the training in the AC group does not involve mirror discrimination learning. The C group only takes part in the tests but does not participate in any activity during the intervention period. The other groups (T, T+S, and AC) are invited to participate *daily* (weekdays) according to their respective training/nap sessions. We conducted our experiment after the break of the first semester. We recommend applying the intervention during three consecutive weeks, with no gaps between them, but future investigation should quantify the importance of this continuity, the interest of spacing the training weeks, and, in fine, determine the best intervention *regime* possible.

In our case, the participants were children between 5 and 7 years old (1^{st} Replica (n = 32; 15 females), mean age = 6.0 $\pm$ 0.56 years; 2^{nd} Replica (n = 60; 33 females), mean age = 5.95 $\pm$ 0.62 years; 3^{rd} Replica (n = 48; 23 females), mean age = 6.02 $\pm$ 0.64 years). The participants did not have any differences in visual acuity, and we did not exclude left-handers or children with learning difficulties (in our case no formal diagnosis was known for any participant). We do not recommend the exclusion of children with learning disabilities because this could be viewed as discriminatory, but researchers should decide *a priori* what they will do with their data (simply exclude, treat as anecdotal cases, treat as a separate group etc.). Considerable variability of performance across participants can be expected at baseline, but this variability can be kept constant *across groups,* by using the randomization procedure described above to balance the groups for initial performance. Future studies should determine whether adult illiterates and children with dyslexia, who typically present prolonged persistency of mirror confusion, could also benefit from such sleep-consolidated targeted training intervention.

To be able to well compare the groups, the ideal is of course to have a large number of participants (> 30) in each group, but unless several schools are involved, there is a natural limitation of the number of participants that could be investigated within a single school. In the case of multiple schools, it is better to have a sample from each school for each of the groups, instead of having groups fully formed from a single school. In other words, it is important to consider the variability across schools.

In the case of replica 1 in our study (reported in Torres et al., 2021), we had a very small number of *remaining* participants (5 or 6 per group). This precluded for instance the ability to detect significant long-term group differences between T vs. T+S in both the visual (for Error Rates), writing and reading tasks, when using the stringent Bonferroni-correction. In contrast, replicas 2 and 3 with more participants did not present this issue. Further, replica 2 (the most powered one), shows significant long-term effects on the T group (relative to the other control groups) in all three tasks (visual, writing, and reading). These results suggest that only replica 2 was sufficiently powered to show sensitivity to all relevant effects, pointing to a threshold of a minimum number of remaining participants required: a minimum of 11 *remaining* participants per group (corresponding to the number of participants in the visual task of replica 2). Note that this is the number of *remaining* participants, so a higher number of *initial* participants is necessary to take into account the high prevalence of drop-outs in a school-based study with several data points. It would be much better to have at the end, the *double* of subjects than this minimum number of remaining participants per group. To deal with the small number of participants per group, we have adopted a replication approach, with the total number of subjects being satisfactory. Thanks to the sensitivity of this experimental paradigm with strong effect sizes, the results were clear and reproducible. If 'one-shoot' research (one replica) is aimed, then a sufficiently large number of initial participants should be included.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data plus code have been deposited in OSF	Torres et al., 2021	https://osf.io/643jh/
Experimental models: organisms/strains		
Human subjects	Children between 5 and 7 years old (of all genders) were assessed in the Colegio Nossa Senhora das Neves school in the city of Natal, Brazil.	https://osf.io/643jh/
Software and algorithms		
The R Project for Statistical Computing	https://www.r-project.org software	N/A
PsychoPy	https://www.psychopy.org software	N/A
Photoshop and Illustrator	Adobe, San Jose, CA	N/A

MATERIALS AND EQUIPMENT

Tasks

Three tasks are used here. One "laboratory-style" visual discrimination task and two "school style" tasks: a paper and pencil writing task and a reading on paper task. The baseline measures are performed before training begins. Our participants were beginner 1st graders who were not reading yet at baseline, so we performed only the visual and the writing tasks at the baseline. However, in other samples the reading task can be applied at baseline (e.g., more advanced 1st graders, 2nd graders etc.). Each task is performed individually by the participants, in the presence of the researcher, and participants do not receive feedback. Whenever possible, the researcher that collects data should remain 'blind' to the groups of participants, to avoid any potential subjective bias. It is also recommended that the activities are carried out in a separate room in the school building, with adequate lighting conditions.

Visual task

Pairs of images are presented on a computer screen (laptop). Each trial starts with a fixation cross for 1,000 ms, followed by the first stimulus for 200 ms, then a new fixation cross for 300 ms, then the second stimulus for 200 ms. The second stimulus can be identical to the first one ('same'), a mirror version of it ('mirror'), or an entirely different one ('different'). Participants are instructed to decide whether the second stimuli are the "same" or "different" by pressing on the keyboard the right arrow to indicate 'same' items and the left arrow to indicate 'different' items. Participants had up to 2.6 s to answer. In separate trials, single letters or visual icons are presented in a sequence similar to that used previously (Pegado et al., 2014b; Pegado et al., 2011; Pegado et al., 2014c). Each trial starts with a fixation cross for 1000 ms, followed by the first stimulus for 200 ms, then a new fixation cross for 300 ms, then the second stimulus for 200 ms (the same, a mirror version, or a different one), which is 25% larger than the first stimulus, to avoid physical repetition. Two categories (letters vs. symbols) and three types of repetition (same, mirror, different) are used intermixed in a single block, for a total of 60 trials per recording. All trials are automatically sorted, so as to generate a random sequence of stimulus presentation. We present a slightly reduced subset of stimuli previously used in Pegado et al. (2011) with asymmetric single letters (lower case): b, f, h, k, s, z, c, e, a, g and asymmetric visual icons (for more details see Figure 8). The output file (Excel format) structure is that of one row per trial, with columns designating the item, condition, participant number, etc.

For more instructions see Figure 9 - Where the experimenter can use to explain the task to the children.

Writing task

Participants are asked to copy the following lower-case letters (b, c, a, f, e, d, g, h, k, s, z, p, j, q, r, t, y). All letters are presented in Arial font, size 90. Despite a more artificial "laboratory-like" style, it is probably better to use a laptop to present the letters for a better control of the presentation time (even slide presentation software with preprogrammed timing can be employed). Each child receives a blank sheet divided into squares (one for each letter copy). The letters are shown one at a time, in the order above, for 3 s. Immediately, after that, the participant must write the letter while blindfolded. Then, the same procedure starts for the next letter: the participant removes their blindfold, looks at the letter, and puts the blindfold back on, and then draws the letter. The child has up to 1 min to answer, and the task is scored considering the percentage of copy errors. Note that this binary evaluation of correct/incorrect copied letters can involve some subjectivity and thus it is critical here to perform this evaluation using a judge that is 'blind' to the group of participants, as we did.

Instructions: "In this game, you are going to copy the letters that I show you, in each space in this sheet (show the sheet to the child). When I show you a letter, you copy it onto the empty square, like this (do the first one yourself). After that, I'll show you another letter and then you copy it onto the next empty square, like this (do the second one yourself). Do you understand? So, now it's your turn!" Continue the task but let the child copy this time. Then explain that you are going to show each letter for only 3 s and then you (the child) must close your eyes (using the eye-masks) and copy the letter with the eyes closed. Ask whether he/she understood, further explain the game if necessary and then replace the sheet of paper with a new one before starting the data collection. The letters used in the three familiarization examples are excluded from the testing set.

Reading speed task

Children are requested to read a list of words (.e.g., 'fogo' [fire], 'maluco' [crazy], 'rei' [king]…) and another list of pseudo-words ('agroz', 'monar', 'lamer'…) printed on separate sheets of paper; for a complete list, see https://osf.io/643jh/. Participants receive the 'go' signal, turn the sheet of paper and start to read the item as fast as possible for 1 min (each list). These lists are essentially the same used in our previous studies done with adults, but stimuli outside the typical vocabulary of young children were excluded. After one minute, the task is stopped, and the number of items read is recorded. The reading speed scores are calculated by averaging the number of words and pseudo-words read in one minute each (two separate lists). This is a very simple and fast reading test that proved to be robust to distinguish readers from different levels in our previous studies (Pegado et al., 2014b; Pegado et al., 2011; Dehaene et al, 2010).

Instructions: In this task, you will have one minute to read aloud a list of words (first sheet). You should read the words in the same column, from up to down, only then, the next column should be read. Show a fake list and first simply point word by word to make the reading direction clear, then read the first three words and ask the child to continue up to the 10th word. After that you say that the child should correctly read the words as fast as possible. Ask whether everything is clear.

⚠ CRITICAL: You must ensure that the task was well-understood before the start of the real test. This is important because you will not have a ''second shot'', i.e. if you need to stop the task in the middle to start again, some words will already be known by the participant, which could bias the results.

When you are sure the participant understood well the task and the words are been read by columns you can give them the sheet of words turned face down and explain that he/she should turn it and start reading after your signal. Ready? 3, 2, 1, Go! Using a chronometer let one minute pass then stop and count the number of words read (based on the position of the last word read). Then explain that we will do the same procedure, but this time with 'pseudowords that look like words because we can read them, but they do not exist in our language. Show the list of pseudowords and ask the child to read the first 10 examples. Then start the same procedure described above.

Post-training naps

Participants in the T+S group are allowed to sleep for up to 2 h daily after lunch (between 12:30–14:30) during the 3 weeks of the intervention. Participants should be conducted to a quiet room, where a mat and an eye mask will be handed to each participant.

Instructions: Now it's nap time. Please be quiet and sleep well.

⚠ CRITICAL: You must ensure that the sound, light, temperature, and humidity conditions are adequate to facilitate sleep, and that the participants do not engage in disruptive conversations during the nap time. The allocation of the sleep period to the post-lunch interval facilitates the intervention because of postprandial somnolence.

STEP-BY-STEP METHOD DETAILS

Part 1: Training

''Multisystem mappings'' training

Participants in the T, T+S, and AC groups receive daily (weekdays) training sessions of 30 min for 3 weeks. We suppose that the distribution of the training time across the three weeks (instead of compressing it into a few days of training) is important for its efficiency, but future studies should probe the efficiency of different training regimes. Perform training in the morning, to precede nap time (that is performed after lunch). If only one researcher performs the training for all groups, he/she should do it serially in sequential 30 min time slots. In our case it was practical for the school to have the training time at 09:00, 09:35, and 10:10 but future investigations should determine it would be physiologically better if the training is placed closer to the lunch time, i.e., late morning, to reduce the gap between the training and its consolidation. It is important to counterbalance the order of the three training groups for each time slot across days by using a Latin square procedure.

A typical training session comprises a sequence of short (a few minutes) multisensory-motor activities presented as games. Perform the training in reduced groups (between 8 to 15 children), for up to 2 h always with an equivalent (initial) number of participants across the T, T+S and AC groups per replica. Use only one pair of letters in each session. T and T+S groups play with opposed mirror-letters while the AC group plays with symmetrical letters. The training is built based on our ''multi-system mapping'' hypothetical mechanism of mirror discrimination learning for letters (Pegado et al., 2014c). We created activities aiming to *maximize* mappings between systems representing letters and used pairs of mirror-letters in the training groups in order to improve the discrimination between them. The results revealed that mirror discrimination learning for letters generalises to the other assymetrical letters (non mirror-letters) : see Figure 4 of the original work). Beyond the three

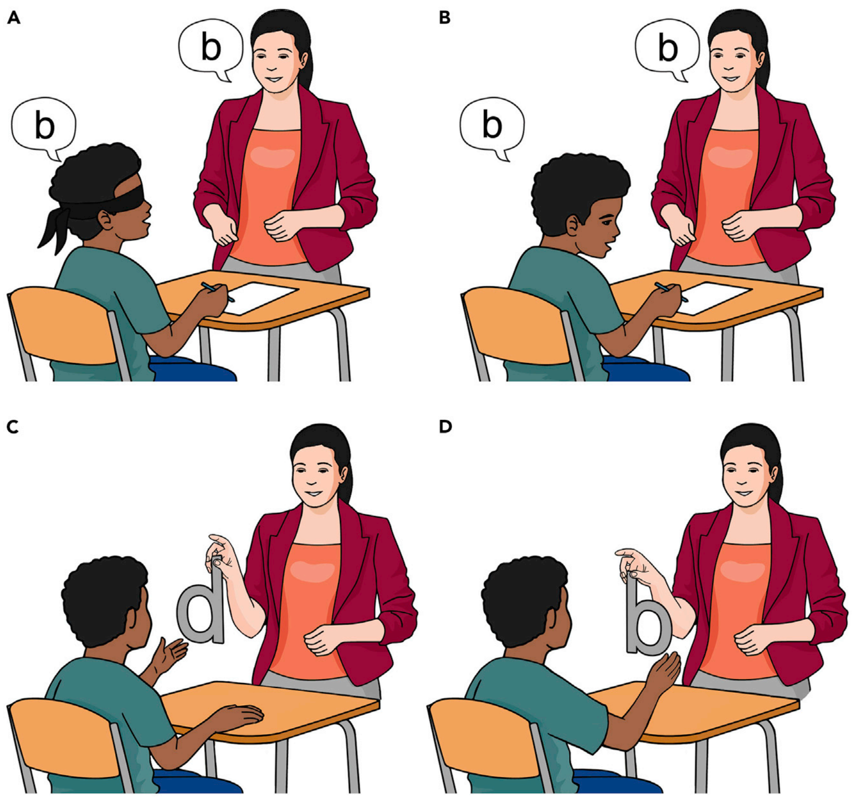


Figure 2. Additional examples of "multisystem mapping" training activities

After performing several multisensory-motor activities (e.g., writing letters on a paper) while blindfolded (A), which represents about 70% of the training time, participants remove the eye-masks and perform the same activities (for a shorter period of time) but now with eyes open (B). In this second phase of the training, a 'perception for action' activity is included, by asking participants to indicate, as fast as possible, parts of letters (e.g., the round part of letters 'd' or 'b') by using the left (C) and right (D) hands (simulating actions upon the letters, such as grasping), in the T and T+S groups. AC groups are requested to indicate parts of symmetrical letters (e.g., central part of letter A or letter X) with either hand. Figure reprinted with permission from Torres et al., 2021.

systems traditionally involved in letter representations at school, i.e., auditory, visual and writing gestures systems, we also included tactile perception and 'perception for action' activities to expand the possibilities of mappings towards the somatosensory and the dorsal system respectively.

To reduce the influence of vision and to increase attention to other sensory and motor representations, for around 70% of the session time, children have their eyes closed (except for 'perception for action' activities, as explained below). The training sessions should typically start by blocking visual inputs with eye-masks. Three types of activities (described below) are proposed with eyes closed: 1) "air-writing"; 2) writing on paper; and 3) "perceiving letters on hands". Closer to the end of the session (~ 30% of the remaining time), the children open their eyes and perform the same activities again (for a shorter duration). The aim of this phase is to map the auditory, tactile, vocal and writing representations with visual representations of letters. Further, at this phase, introduce one additional activity with eyes opened: "perception for action" (Figure 2).

Experimental training: "Multisystem mapping" for mirror letters

Timing: 20/30 min (20 min of effective training + 10 min for bringing, installing and instructing the participants)

Training Group: the two experimental groups, i.e., Training (T) and Training + Sleep Group (T+S)

Note: Use opposed mirror-symmetric letter pairs in this training.

Letters included:

Monday and Wednesday: Distinction between letters 'b' and 'd'

Tuesday and Thursday: Distinction between letters 'p' and 'q'

Friday: Distinction between letters 's' and 'z'

1. Blindfold activities. Sitting on their chairs, children receive a fold, which they place on their eyes at the beginning of the session. The aim is to reduce the dominance of the visual system over other systems (Rock & Victor, 1964), eventually confusing them with visual mirror confusion) and, at the same time, to favor attention focused on the other sensory-motor systems. They also have a paper and a pencil in front of them.

   Note: Propose the activities for one letter first, then to the other letter. Finally, invite the participants to find out which of the two is being written.

   a. The experimenter announces the beginning of each activity to the group.
   b. **Air writing.** Guide each child to make movements in the air with their index finger of preferred hand. Initially, these movements mimic some elementary shapes (circles, semicircles, vertical and horizontal lines), to 'warm up' (familiarisation). Then, guide the participant to make movements following the written gestures of the proposed letters (mirror-letters in this case).
      i. First, use a hand to guide the participant to passively perform the movement. Then, ask the participant to actively perform it (without the guidance of the experimenter's hand, only by verbal instruction). After that, ask the participant to 'guess' a letter being traced (passively). Different sizes of letters are proposed, with the aim to create a more abstract *size invariant* representation of letters.

   Note: During the activity, pronounce the sound of the letter (letter name) continuously, either the researcher or the participant (alternating). The aim is that auditory and vocal representation of letters could map to writing gestures of letters.

   c. **Writing on paper.** This activity is similar to the previous one, but this time using the pencil and the sheet of paper placed in front of the participants. The difference here is that participants feel the resistance of the pencil on the paper when tracing the letters and the writing gestures are smaller. Apply the same principles here: an initial warming-up with elementary shapes and then starting letter gestures with 1) passive movements first, then active movements and finally guessing the letter proposed by the experimenter; 2) the use of different letter's sizes; and 3) concomittant pronunciation of the letter name (by the researcher or by the participant, alternatively) during the activity, showing the child that when we hear or pronounce the letter "b" for instance it is different from when we do the same for the letter "d".
   d. **"Perceiving letters on hands".** This activity consists in tracing a letter in the hand of a participant using the researcher's finger and asking the participant to say out loud the letter name (Figure 1), using the same principles as described above. This activity aims to map tactile representations with articulatory and auditory representations of letters.

   Note: Conduct each of these three training activities described in steps 1b–1d with each child in a group by taking turns, moving through the group, one child at a time, while the other children wait.

2. Eyes-open activities
   a. **Repetition of previous activities but now performed with eyes-open.** After the activities with eyes closed, repeat the same activities with eyes open (but this time for a shorter period). The

aim is to map the *visual* representations of letters with all the letter representations on the other systems.

b. **"Perception for action".** Here we introduce a new activity: "perception for action". The game consists to indicate, as fast as possible, parts of letters (Figures 2C and 2D).
   i. Ask participants to indicate, as fast as possible, parts of letters (e.g., the round part of letters 'd' or 'b') by using the left and right hands ((Figures 2C and 2D); simulating action upon the letters such as grasping.

Note: The aim is to map the dorsal visual system ("perception for action") with the ventral visual system ("perception for recognition"), as even illiterates, that did not adquire mirror discrimination (for recognition), are able to perform visual tasks where 'action upon the images' can be simulated (e.g., to handle a cup of tea with the correct hand) (Kolinsky and Verhaeghe, 2017).

Control training: "Multisystem mapping" but for symmetrical letters

Timing: 20/30 min (20 min of effective training + 10 min for bringing, installing and instructing the participants)

Control Group: Active Control (AC)

Note: Symmetrical letters are used in this training.

Letters included:

Monday and Wednesday: Distinction between letters 'A' and 'X'

Tuesday and Thursday: Distinction between letters 'M' and 'U'

Friday: Distinction between letters 'T' and 'H'

3. **Blindfold and eyes-open activities.** Apply the exact same activities performed by the experimental group (steps 1 and 2) with the active control (AC) group, but with the critical difference to use letters that are all *symmetrical* letters (across the vertical axis) to avoid any mirror discrimination learning.

Part 2 : Sleep

Post-training sleep consolidation

Timing: Up to 2 h duration, starting closely after the morning training (but after lunch)

4. Invite participants in the T+S group and sleep control groups to have post-training nap sessions, immediately after lunch, during the 3 weeks of the intervention, within the school setting. Participants should nap for up to 2 h a day in a quiet room of the school on mats and wearing eye masks.

Note: while it is interesting to measure neural activity and other physiological correlates of sleep, our naturalistic experimental design does not include devices to gauge activity. Instead, the experimenter records the amount of time that each participant *remains still* during the nap opportunity, as a proxy of the amount of post-training sleep, for each day of intervention.

Note: A previous school-based study (Cabral et. al, 2018) has shown the importance of nap duration, with children receiving 30 minutes or more of sleep presenting a clear advantage

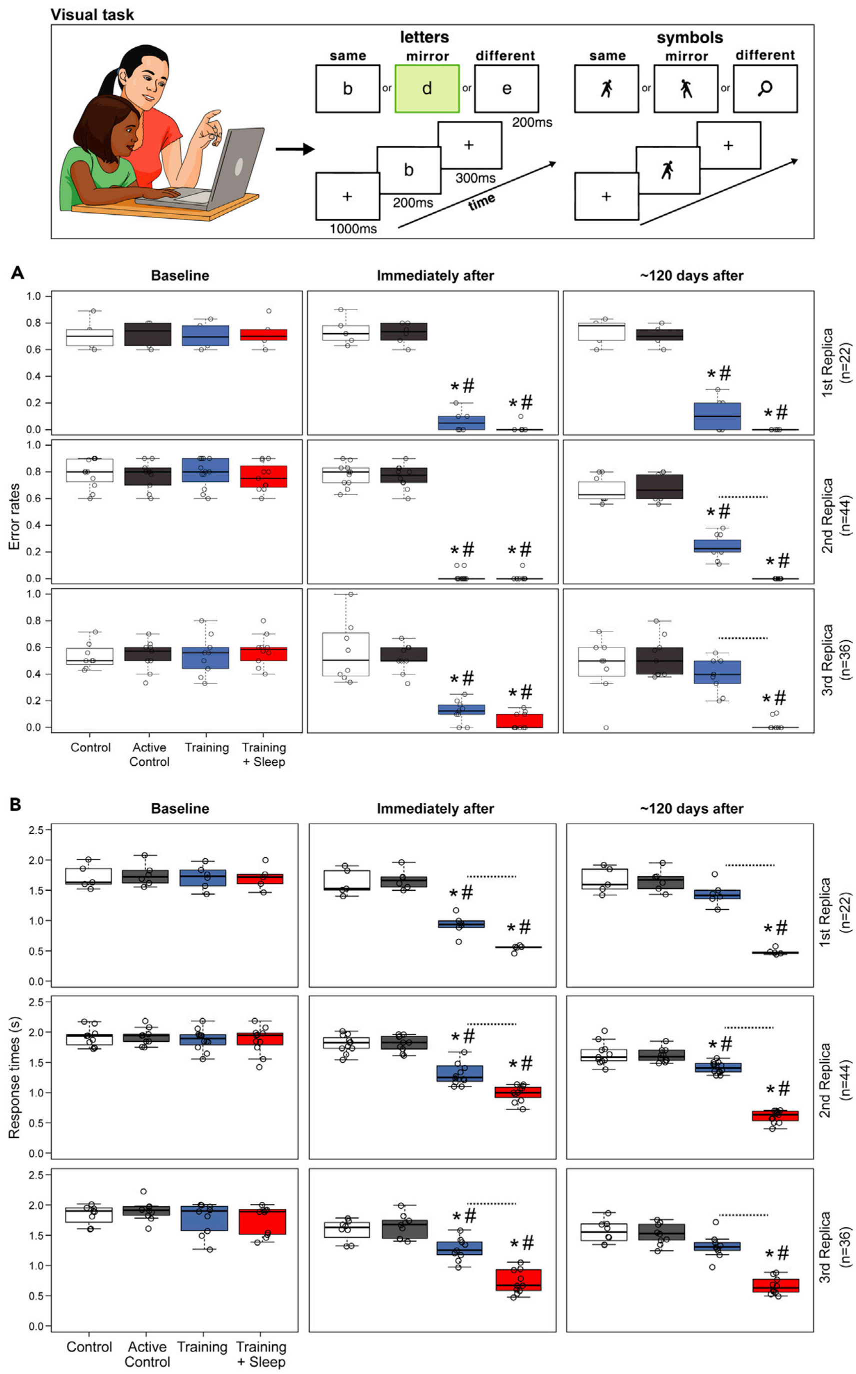


Figure 3. Impact of the intervention on visual mirror discrimination for letters

Top (task): In this visual task, participants decided whether two stimuli sequentially presented on a laptop were the "same" or "different", pressing corresponding buttons. Two categories of stimuli (letters or iconic symbols) and three trial types (same, different, or mirror) were present. Note that mirror trials should be assigned as "different". Stimulus duration is indicated below each frame. Bottom (results): Results concern the "mirror trials" for letters category (indicated here by a green frame in the top image): (A) Error rates (ER) and (B) Response times (RTs) with only correctly responded trials considered. Note that these results concern all 10 letters of the experiment and not only mirror-letters. Columns: Pre-intervention baseline (left), immediately after the 3-week intervention (middle), and ~120 days after the end of the intervention (right). Each dot represents one participant. Symbols indicate statistically significant differences in pairwise group comparisons (Wilcoxon rank-sum test; $p < 0.05$ Bonferroni-corrected) from the C group (*), from the AC group (#), and between the T and T+S (—) groups. The AC group underwent the same activities as the

Figure 3. ***Continued***

training group but played only with symmetrical letters (e.g., A and X), thus with no mirror discrimination learning. Boxplots: Central horizontal line, median; box, 25th and 75th percentiles; whiskers, minimum and maximum; outlier = box limits +/- 1.5 interquartile range. Figure reprinted with permission from Torres et al., 2021.

relative to children receiving less than 30 minutes of sleep. We recommend long nap time (2 hours window) but future studies should determine the minimum amount for *optimal* effects.

Note: even though we could hypothesize that the best timing to apply sleep would be immediately after the training, we realize during a previous piloting that this is physiologically unfeasible and it is much more natural to apply naps after lunch (see troubleshooting).

EXPECTED OUTCOMES

The results from the application of the protocol described here were consistently replicated in three randomized controlled trials (RCT) and are displayed in Figures 3, 4, 5, 6, and 7 below. All three RCT replicas consistently show short-term improvement in mirror discrimination for letters in the two training groups but not in the two control groups. For a more comprehensive discussion on the results, please refer to Torres et al. (2021).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses and plots can be performed using standard software, such as R software (we use version 1.2.1335). First, we perform a normality test for the results of each task. As usually not all data turns out to be normally distributed in this type of naturalistic design, we generally use non-parametric statistical tests to compare the group scores. Kruskal-Wallis tests for group effects with significant p-values are followed by Wilcoxon ranksum tests, with Bonferroni correction for the six possible pairwise comparisons across the four groups (C, AC, T, T+S). Two exclusion criteria are used to guarantee the quality of the data for the visual task. First, we exclude participants who fail, in any of the sessions, to press each of the two computer keys used in the test at least once, a situation that could reflect, for instance, lack of task understanding or a potential technical problem. Second, trials with no response are excluded. Note that in our case about 20% of the data were excluded using these criteria. Blind correction (for group identity) is important for the writing test to avoid potential subjective bias. Ideally, all tasks should be conducted by a second experimenter, blind to the group of the participant.

LIMITATIONS

Performing a protocol like the one described here in a school environment poses several challenges, such as a high number of dropouts, a small number of participants per group, and restrictions to short and few tests in order to minimize the disturbance on the school dynamics. Another possible limitation is teacher bias, so we strongly recommend that teachers remain blind to the experimental design and to the group assignment. In addition, one could argue that child motivation might also be a confounding variable, since some children go out of regular classes for 30 min every weekday to play new "nice games". Therefore, we recommend an active control (AC) group, which is included in this daily class dismissal and participate to the same "nice games", to control for socio-emotional and motivation effects.

Our protocol cannot fully disentangle between a strict mirror discrimination effect and a more global improvement in visual letter discriminability. For instance , a potential lack of sensibility in the 'same' and 'different' conditions (no room for improvement), could omit a potential global improvement in letter processing. Additional tasks could be added, if time permits, in order to test potential changes in other aspects of letter processing.

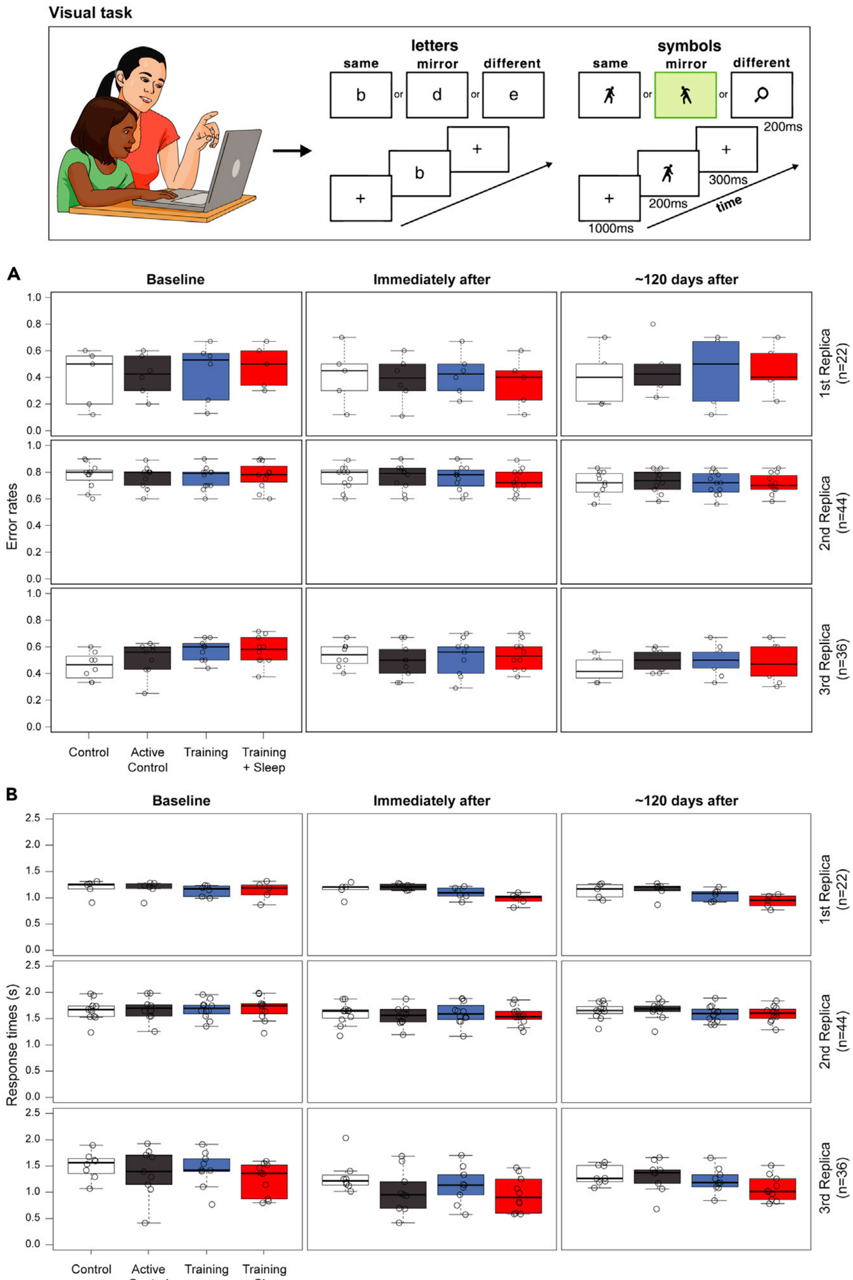


Figure 4. The intervention had no impact on the control condition (mirror trials for symbols)

Note that training does not change performance for non-alphabetic stimuli. (A) Error rates and (B) response times (RTs) for the control condition mirror trials for symbols (this condition is indicated by a green frame in the top image). In all measurements, replicas, and time points there is no significant group effect (except for a small intervention effect on RTs in replica 1 immediately after the intervention). Columns: Measures at the preintervention baseline (left), at the end of the 3-week intervention (middle), and at ~120 days after the end of the intervention (right). Each dot represents one participant. Boxplots: Central horizontal line, median; box, 25th and 75th percentiles; whiskers, minimum and maximum; outlier = box limits +/- 1.5 interquartile range. Figure reprinted with permission from Torres et al., 2021.

TROUBLESHOOTING

Problem 1

Problems during sleep time (step 4)

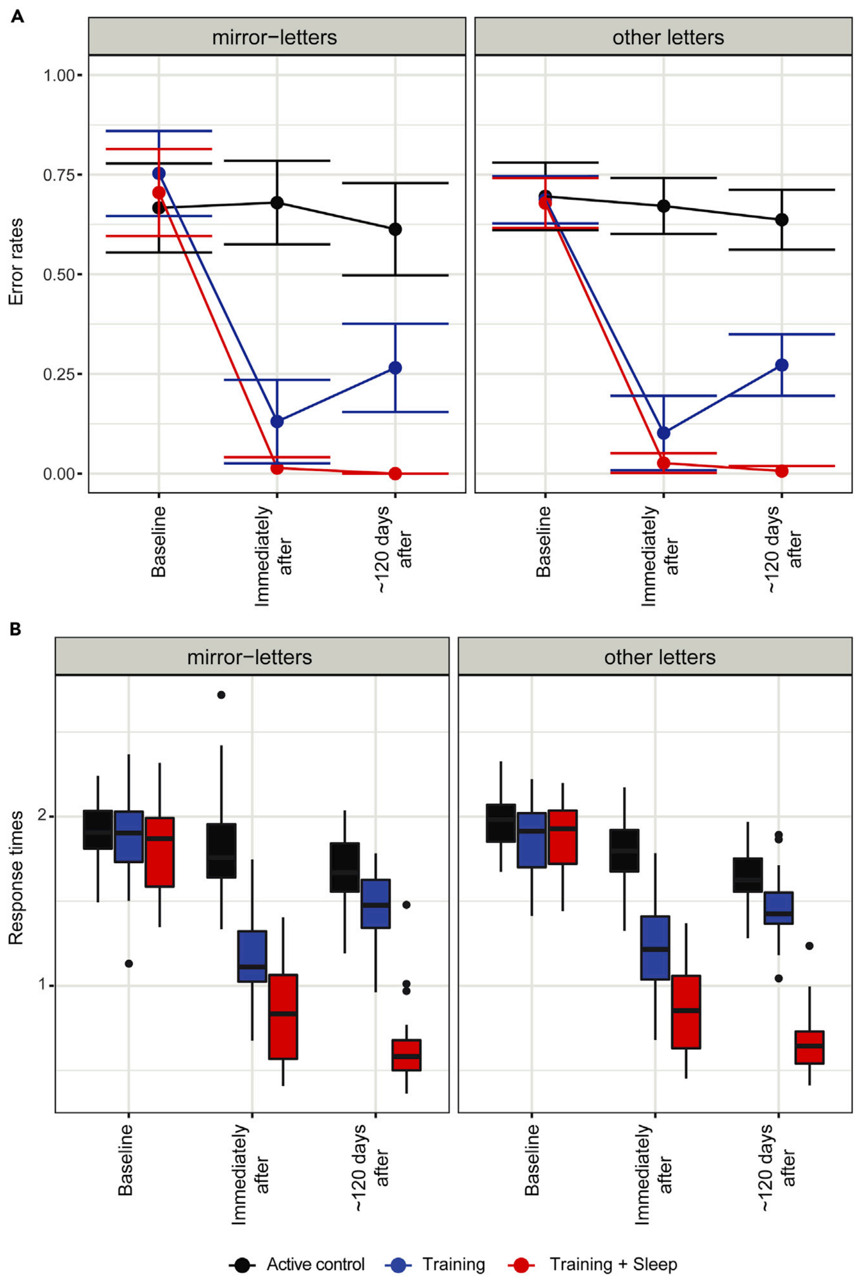


Figure 5. Mirror discrimination learning generalizes to other letters (collapsed data across the three replicas)

(A) Error rates for 'mirror-letters' vs 'other letters', restricted to the 'mirror' condition.

(B) Response times for 'mirror-letters' vs 'other letters', restricted to the 19 'mirror' condition. This exploratory analysis shows a 'generalization' of learning mirror discrimination for other letters (collapsed data). Vertical bars illustrate pointwise 95% confidence intervals. Figure reprinted with permission from Torres et al., 2021.

Potential solution

In our first piloting of the protocol, we tried to have the children sleep in the period immediately after training to maximally shorten the post-training interval preceding sleep consolidation. However, when we tried to implement this schedule by the end of the morning, we were faced with a complete failure, because the children become hungry and did not fall asleep easily. The expectation of lunch made them quite restless, and the experiment ended up disturbing the overall school dynamics

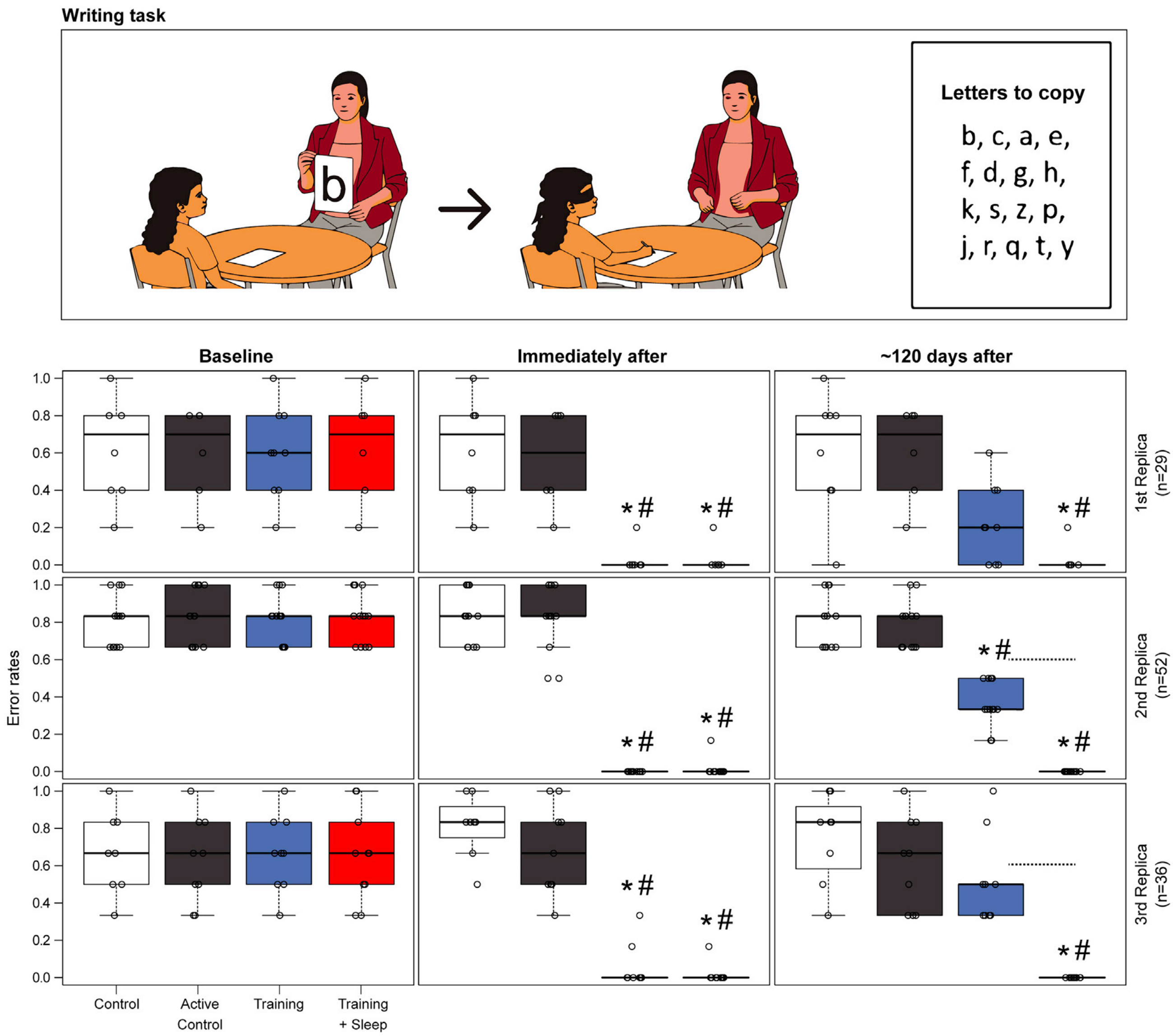


Figure 6. The impact of the intervention on writing errors

Top (task): Participants observed single letters for 3 s, then with eye-masks wrote the letters with pencil on paper. Bottom (results): Error rates for writing letters (including mirror errors). Columns: Pre-intervention baseline (left), immediately after the 3-week intervention (middle), and ~120 days after the end of the intervention (right). Each dot represents one participant. Symbols indicate statistically significant differences in pairwise group comparisons (Wilcoxon rank-sum test; $p < 0.05$ Bonferroni-corrected) from the C group (*), from the AC group (#), and between the T and T+S (—) groups (see Data S1 (A and B)). Boxplots: Central horizontal line, median; box, 25th and 75th percentiles; whiskers, minimum and maximum; outlier = box limits +/- 1.5 interquartile range. In replica 3, the number of letters to copy was extended, including five additional letters (j, q, r, t, y). Figure reprinted with permission from Torres et al., 2021.

without efficient sleep time. Given this problem we opted for a more physiological approach, namely to have the children sleep immediately after lunch. This period is more suitable to fall asleep, given the postprandial state.

Another point is the children's engagement in sleeping. In this way, there must be a certain distance between the mattresses used by the children so that they are not too crowded and start to interact with each other losing the focus on sleep. In addition to the use of eye masks and exercise mats, in cold countries it may be necessary to add blankets.

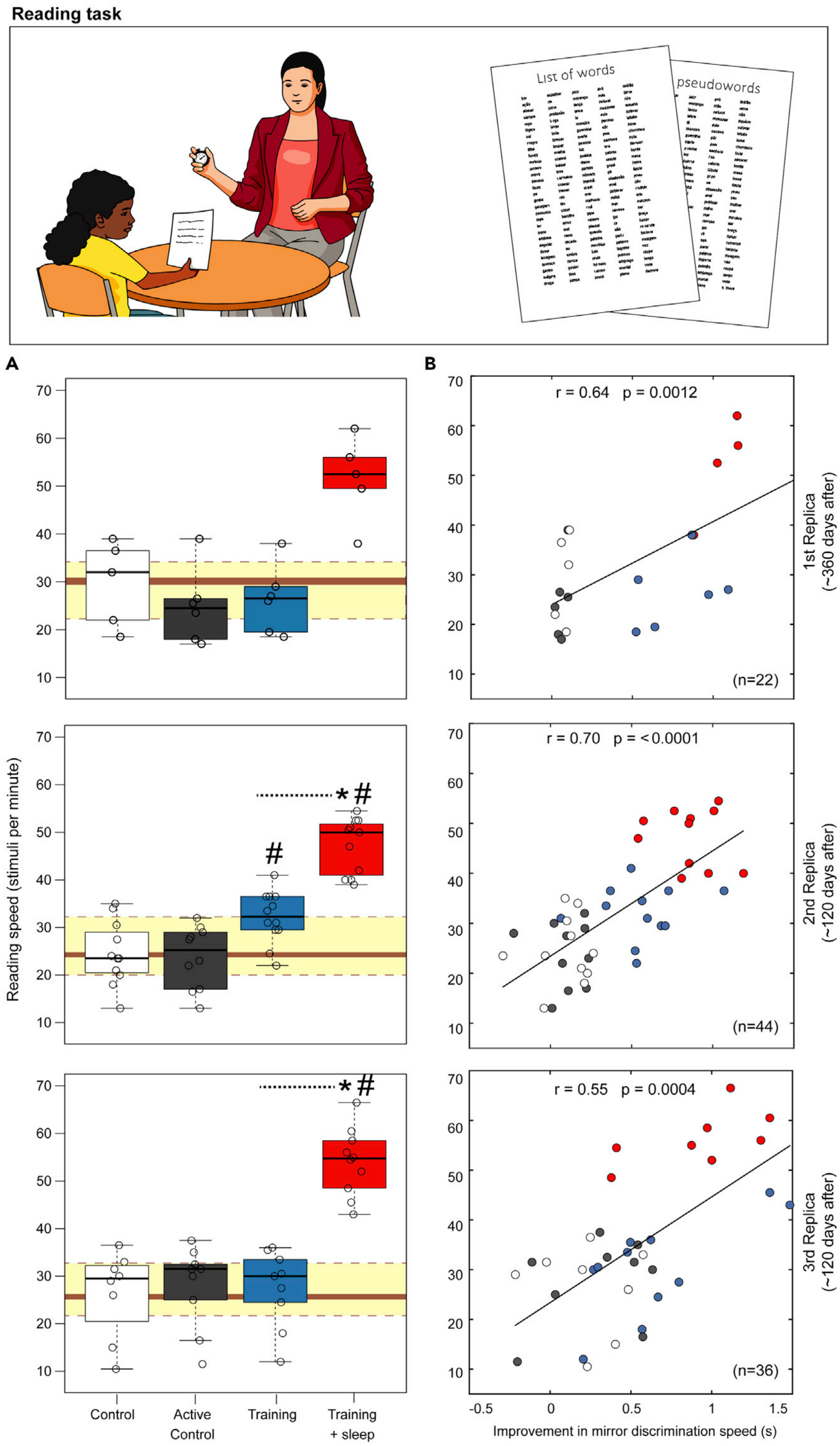


Figure 7. The impact of the intervention on reading speed

(A) Reading speed (average of words and pseudo-words read in one minute) for each group measured several months after the end of the intervention (indicated on the vertical axis in (B)). The lower and upper boundaries of the yellow stripe correspond to the 1st and 3rd interquartiles (and the brown line to the median) of extra control groups, controlling for the impact of afternoon naps with no training . (B) Reading speed as a function of improvement (relative to baseline) in the speed of mirror letter discrimination in the visual task. Each dot represents one participant. Symbols indicate statistically significant differences in pairwise group comparisons (Wilcoxon rank-sum test; $p < 0.05$

Figure 7. *Continued*
Bonferroni-corrected) from the C group (*), from the AC group (#), and between the T and T+S (—) groups (—) (see Data S1 (A and B)). Boxplots: Central horizontal line, median; box, 25th and 75th percentiles; whiskers, minimum and maximum; outlier = box limits +/- 1.5 interquartile range. Note that the low number of remaining participants in replica 1 precluded significantly different results for T+S group under this stringent correction. Figure reprinted with permission from Torres et al., 2021.

Problem 2

Risk of lacking engagement during tests and training (steps 1–3)

Potential solution

Given that the same activities during the *training* are done every day over three weeks, there is a risk that they become a little repetitive for some children. We thus reduced the initial planned training time of 2 h to just 30 min and alternated between three pairs of letters ('b' and 'd'; 'p' and 'q'; and 's' and 'z',) over the days. In addition, to keep motivation high, we also provide feedback and rewards in the form of verbal compliments for children's achievement during the activities. Note that feedback plays a critical role in learning (e.g., Frank et al., 2020).

Problem 3

Dropouts

Potential solution

Since this type of study is not performed in a laboratory but in a school setting, and especially because long-term retesting is required, a high rate of dropouts can be expected. In addition, several participants might leave the school in the following academic year, hindering long-term re-testing. It is thus recommended to apply long-term retesting by the end of the same academic year, to reduce these dropouts. A constant attention to obtain high quality data (e.g., ensuring confort, friendly environment and good comphrehension of the task by the children) can also prevent loss of data (exclusion).

Problem 4

Disturbing school dynamics

Potential solution

Schools generally have a dynamic routine and are full of activities. To avoid disturbing the school dynamics, researchers should discuss with the teachers to find out the best timing for evaluations and training, taking into account physiological and school schedules, for instance the snack time, recreational time, physical education and lunch time. It is important for long-term collaboration to establish a very good communication between researchers and teachers. It is also important to well explain the project to the parents and coordinators to be able to obtain their genuine engagement.

Problem 5

Limited number of participants in small schools

Potential solution

In case this protocol is implemented in a small school, where the total number of children in first grade is small, alternative designs could be implemented. First, to avoid extra control groups (for

Figure 8. Symbols used on the visual task
Figure reprinted with permission from Torres et al., 2021.

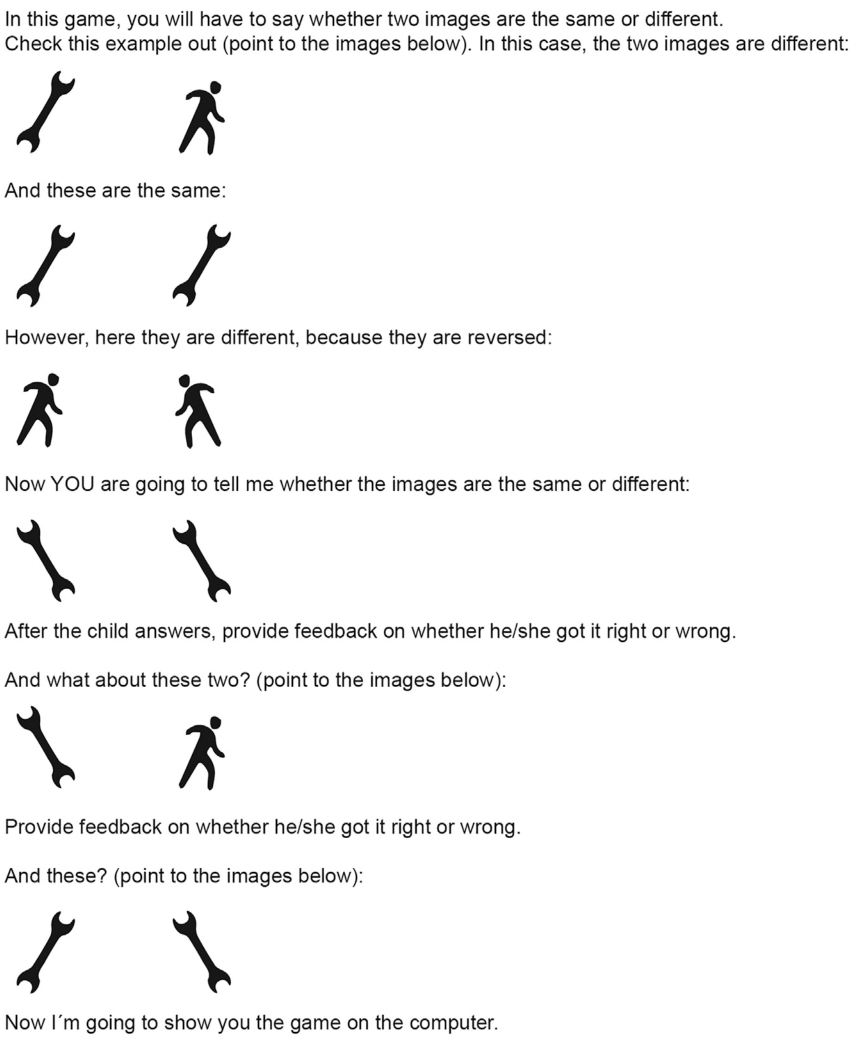


Figure 9. Explanation with illustrative images of the visual task where the experimenter can use to explain the task to the children

Figure reprinted with permission from Torres et al., 2021.

sleep), an alternative consists of having an experimental group (T+S), a control sleep group (CS), which trains as the T+S group but does not sleep, and a control training group (CT), which does not undergo any training, but sleeps. Another alternative, for very small schools (one class per level) is to only have one experimental group (T+S) and one control group receiving active control training plus sleep (AC+S). The individual contribution of the training and sleep factors could not be controlled in this case, but it would allow the probing of the main impact of the intervention relative to a control group. However, there is a risk of lack of power given the reduced number of inicial participants. A good strategy is to associate a second school receiving the same protocol and pull the data together.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contacts, Sidarta Ribeiro (sidartaribeiro@neuro.ufrn.br) or Felipe Pegado (felipe.pegado@univ-amu.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Original data and code have been deposited in OSF and can be found on https://osf.io/643jh/.

ACKNOWLEDGMENTS

We thank the Neves School for facilitating the experiments; the parents and children for permission to conduct them; and the Latin American School for Education, Cognitive and Neural Sciences (LA School) for providing training and inspiration to the authors. We thank the Google Latin America Research Award (2017) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for A.R.T., FWO (Fonds Wetenschappelijk Onderzoek) postdoc fellowship (12Q4615N) and ERC grant (742141) for F.P., grant #306659/2019-0 from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for J.W., grants #308775/2015-5 and #408145/2016-1 from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and grant #2013/07699-0 from the São Paulo Research Foundation (FAPESP) Center for Neuromathematics for S.R.

AUTHOR CONTRIBUTIONS

F.P. and S.R. conceived and supervised the project. A.R.T. conducted the experiments. F.P., A.R.T., and S.R. analyzed data. F.P., A.R.T., J.W., and S.R. wrote the manuscript. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Cabral, T., Mota, N.B., Fraga, L., Copelli, M., McDaniel, M.A., and Ribeiro, S. (2018). Post-class naps boost declarative learning in a naturalistic school setting. NPJ Sci. Learn. *3*, 14.

Cousins, J.N., van Rijn, E., Ong, J.L., Wong, K.F., and Chee, M.W.L. (2019). Does splitting sleep improve long-term memory in chronically sleep deprived adolescents? NPJ Sci. Learn. *4*, 8.

Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., and Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. Science *330*, 1359–1364.

Frank, S.M., Qi, A., Ravasio, D., Sasaki, Y., Rosen, E.L., and Watanabe, T. (2020). Supervised learning occurs in visual perceptual learning of complex natural images. Curr. Biol. *30*, 2995–3000.e3. https://doi.org/10.1016/j.cub.2020.05.050.

Kolinsky, R., and Verhaeghe, A. (2017). Lace your mind: the impact of an extra-curricular activity on enantiomorphy. J. Cult. Cogn. Sci. *1*, 57–64.

Lemos, N., Weissheimer, J., and Ribeiro, S. (2014). Naps in school can enhance the duration of declarative memories learned by adolescents. Front. Syst. Neurosci. *8*, 103.

Marshall, L., Helgadottir, H., Molle, M., and Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. Nature *444*, 610–613.

Pegado, F., Nakamura, K., Cohen, L., and Dehaene, S. (2011). Breaking the symmetry: mirror discrimination for single letters but not for pictures in the Visual Word Form Area. Neuroimage *55*, 742–749.

Pegado, F., Nakamura, K., and Hannagan, T. (2014a). How does literacy break mirror invariance in the visual system? Front. Psychol. *5*, 703.

Pegado, F., Comerlato, E., Ventura, F., Jobert, A., Nakamura, K., Buiatti, M., Ventura, P., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., et al. (2014b). Timing the impact of literacy on visual processing. Proc. Natl. Acad. Sci. U S A *111*, E5233–5242.

Pegado, F., Nakamura, K., Braga, L.W., Ventura, P., Nunes Filho, G., Pallier, C., Jobert, A., Morais, J., Cohen, L., Kolinsky, R., et al. (2014c). Literacy breaks mirror invariance for visual stimuli: a behavioral study with adult illiterates. J. Exp. Psychol. Gen. *143*, 887–894.

Rock, I., and Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. Science *143*, 594–596.

Stickgold, R., James, L., and Hobson, J.A. (2000). Visual discrimination learning requires sleep after training. Nat. Neurosci. *3*, 1237–1238.

Torres, A.R., Mota, N.B., Adamy, N., Naschold, A., Lima, T.Z., Copelli, M., Weissheimer, J., Pegado, F., and Ribeiro, S. (2021). Selective inhibition of mirror invariance for letters consolidated by sleep doubles reading fluency. Curr. Biol. *31*, 1–12. https://doi.org/10.1016/j.cub.2020.11.031.