Received: 11 May 2021 | Revised: 11 November 2021 | Accepted: 19 January 2022

DOI: 10.1002/acp.3921

RESEARCH ARTICLE

WILEY

Is working memory training efficient? Effects on IQ and school performance in Brazilian children

Drielle Barbosa-Pereira[1] | Pedro S. R. Martins[1] | Luiz Alves Ferreira-Junior[1] | Marina Freitas Alves da Costa[2] | Janaína Paula Chaves Paixão[3] | Renato Ramalho Costa[4] | Renata Saldanha-Silva[5] | Marcela Mansur-Alves[1]

[1]Department of Psychology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[2]Department of Psychology, Pontifícia Universidade Católica of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[3]Graduate Program in Systems Information and Knowledge Management, FUMEC University, Belo Horizonte, Minas Gerais, Brazil

[4]Graduate Program in Electrical Engineering, Federal University of Minas Gerais, Belo Horizonte, Brazil

[5]Department of Psychology, Faculdade de Ciências Médicas of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

Correspondence
Drielle Barbosa-Pereira, Universidade Federal de Minas Gerais, Departamento de Psicologia, Faculdade de Filosofia e Ciências Humanas, Avenida Antônio Carlos, 6627, Campus Pampulha, Belo Horizonte, MG 31270-901, Brasil.
Email: dribarbosa20@gmail.com

Funding information
Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Brasil (FAPEMIG), Grant/Award Number: APQ-00802-15

Abstract

Working memory training (WMT) for children has received great attention in recent scientific investigations. However, evidence on program efficacy is still sparse. This study aimed to investigate the effect of a WMT software program (*Programa de Ativação da Memória de Trabalho* [PRAMEMT]) in promoting gains on cognitive and school related abilities. To investigate its effect, 78 school children attending 1st to 5th grades (52.6% girls; mean age = 7.87 years; SD = 1.44) were randomly allocated to a control ($n = 37$) or an experimental ($n = 41$) group. The children participated in five intervention sessions. The control group answered questions related to the same animals and themes featured in the PRAMEMT, but unrelated to working memory. The results suggested that there were gains in the performance intelligence quotient ($p < .05$; d = −0.54) and phonological awareness ($p < .05$; d = −0.52) in the experimental group; however, there was no change in overall cognitive performance. Possible mechanisms for these findings are discussed.

KEYWORDS

children, intelligence, phonological awareness, working memory, working memory training

1 | INTRODUCTION

Working memory training (WMT) could be defined as stimulation of working memory (WM) during a predetermined period of time through the repetition of standardized tasks that aim to improve WM performance and related cognitive abilities. Currently, there is a broad ongoing discussion on the effect of WMT programs for children (Foster et al., 2017; Mansur-Alves & Saldanha-Silva, 2017; Melby-Lervåg et al., 2016; Redick, 2019; Sala & Gobet, 2017). Those programs are implemented with the assumption that they can increase cognitive functioning; therefore, there is considerable interest in achieving positive results (Pergher et al., 2020). WM is an important executive function (EF) which is a temporary system that holds and manipulates for a short period of time a limited amount of information. (Baddeley, 2003, 2012, 2017; Baddeley & Hitch, 1974). It is required to perform abilities such as reasoning, mathematics, and reading (Ilkowska & Engle, 2010; Sánchez-Pérez et al., 2018; Zhang et al., 2018).

Furthermore, recent evidence suggests that WMT can lead to gains in WM and related abilities, such non-verbal skills and school achievement (Au et al., 2015; Sánchez-Pérez et al., 2018; Zhang et al., 2018). The effects of WMT can be partially explained by neural plasticity, the potential modifiability of a person's cognitive abilities and brain activity. Neural plasticity is observed more easily in childhood (Au et al., 2015;

Karbach & Schubert, 2013; Sala et al., 2019; Vernucci et al., 2020). At the same time, it is unclear how many hours of training should the intervention program have to show improvement in the target cognitive domains (Melby-Lervåg et al., 2016; Sala & Gobet, 2017). On one hand, a meta-analysis conducted by Melby-Lervåg et al. (2016) suggested that training duration did not moderate training effects; on the other hand, a meta-analysis conducted by Sala and Gobet (2017) demonstrated a marginally significant effect of training duration on cognitive abilities. Additionally, the systematic review performed by Pergher et al. (2019) also suggested that many other factors may impact the training results. According to the authors, the training task features (such as stimulus modalities, adaptativity, feedback to the participant, and duration) are candidates for interfering in the observed training effects. Furthermore, the transfer task features (such as the WM task used in the assessment, and the general psychometric properties of the other tasks used to investigate the training effects) can also affect the intervention results.

Many studies have been conducted with typical development (TD) children because this population is larger than that of children with developmental disabilities; the distribution range of WM abilities is higher in TD children than in other populations, and the results of WMT studies with TD children are more generalizable (Sala & Gobet, 2020). In addition, improvement of cognitive abilities has presented great clinical potential, particularly in recent years (Redick, 2015). The impact of WMT on real-life skills of children, such as school achievement, has been the most researched effect within the field of cognitive training (Melby-Lervåg et al., 2016; Sala & Gobet, 2017; Vernucci et al., 2020), considering that WM is an important predictor of school achievement and is correlated with general cognitive abilities (Alloway et al., 2013; Gathercole & Alloway, 2004; Sánchez-Pérez et al., 2018). For example, Zhao et al. (2011) trained 33 children aged 9–11 years and found significant gains in fluid intelligence in favor of the experimental group ($d = 0.98$). However, the effects of WMT on improving fluid intelligence remain inconsistent (Mansur-Alves & Saldanha-Silva, 2017; Melby-Lervåg et al., 2016). According to the results of the meta-analyses conducted by Mansur-Alves and Saldanha-Silva (2017), there is no consistent evidence of training protocols that improve fluid intelligence ($\underline{g} = 0.09$). The meta-analysis performed by Melby-Lervåg et al. (2016) found a similar overall effect size ($\underline{g} = 0.05$). Accordingly, the study by Melby-Lervåg et al. (2016) also found no evidence of WMT effects on verbal and non-verbal skills, reading comprehension, and mathematics in studies with active control groups. Nonetheless, recent studies have found positive results regarding training effects. For example, Fälth et al. (2015) used the CogMed™ software in a sample of 32 schoolchildren and found gains of ~13 points for a single word-reading task. In their study, the sample size was small, the control group was passive, and they did not control for intelligence prior to the training sessions. Honoré and Noël (2017) also used CogMed™ in a sample of 46 children aged 5 and 6 years. Their results suggested transfer for Arabic number comparison ($\eta^2 = 0.07$). Additionally, for academic achievement, the meta-analysis by Sala and Gobet (2017) suggests gains only in mathematics ($\underline{g} = 0.20$).

According to Pergher et al. (2019), the absence of a general standard for the type of tasks used in WMT studies may also impact the heterogeneity in meta-analysis results. Generally, WMT studies do not consider the developmental aspects of WM, and use tasks that may not be adequate for the age group assessed. The tasks requiring the visuospatial sketchpad tend to be more appropriate for younger children, that is, preschoolers (Henry, 2012). The tasks involving the phonological loop require familiarity with alphanumeric symbols and tend to be too complex for younger children, such as *n*-back tasks, which involve functions within the central executive (the most complex WM component according to Baddeley's, 2012 model). Hence, training protocols with phonological loop and *n*-back tasks are more likely to provide more gains to older children (Gathercole et al., 2004; Henry, 2012; Uehara & Landeira-Fernandez, 2010). We emphasize that the WMT should consider training all the WM components while maintaining the focus on activities that would benefit each age group the most.

Nevertheless, some methodological issues may have been a confounding factor in WMT effects. Pergher et al. (2019) suggested that there is great variety in the tests and tasks used to assess program effect. The tests present different and often inadequate psychometric properties (Pergher et al., 2019). The second-order meta-analysis performed by Sala and Gobet (2020) with TD children also corroborated the issues pointed out by Pergher et al. (2019). Therefore, there is a need for more rigorous tests to assess the effect of WMT programs. The lack of studies on the validity of WMT has also contributed to divergent effect results. Training programs are rarely examined for their content validity (Barbosa-Pereira et al., 2019; Golino et al., 2017). If the programs do not present satisfactory validity indices or adequate psychometric parameters, the cognitive stimulation content may be considered controversial. Therefore, it is necessary that WMT studies increase their methodological rigor so that more consistent results can be achieved (Pergher et al., 2019).

Other features can also mask the gains of WMT, such as the characteristics of the intervention protocol (Pergher et al., 2019). According to Pergher et al. (2019), some programs have symbolic reinforcements (feedback) and some do not. Usually, the feedback is given to the children during progress in the session and are based on the participants' accuracy at previous levels or attempts (Pergher et al., 2019). A systematic review conducted by Pergher et al. (2019) showed that most studies conducted with adults did not report the presence or absence and the type of feedback used in the program. Furthermore, the studies that reported the use of feedback varied in providing the answer to the participant at the end of each block, session, or attempt (Pergher et al., 2019). Although that systematic review was conducted with samples of adults, no studies that show how feedback is present in WMT programs have been conducted with children so far. Nonetheless, rewards and material reinforcement, when present in training, can vary in the type of material, such as money and small objects (Katz et al., 2018). The presence of prizes, as well as of feedback, is important for maintaining the motivation and attention of the participant in conducting the tasks (Peijnenborgh et al., 2016). In addition, adaptive protocols can be an appropriate option, since they adjust the difficulty to the participant's ability (Holmes et al., 2009). In adaptive training, the participant advances or returns according to the accuracy rate obtained in previous attempts or levels (Landsberg et al., 2012).

Considering the divergent results on transfer effects, as well as on an attempt to refine methodological aspects of WMT studies (using active control group design, random allocation, and a WMT program that was content validated), the present study aimed to investigate the effect of a new WMT tool: the Working Memory Activation Program (PRAMEMT [*Programa de Ativação da Memória de Trabalho* in Portuguese], a software program described in detail in the Methods section). The PRAMEMT has six characters that engage with the children and stories of five main characters (animals: Owl, Toad, Bee, Jaguar, and River Dolphin). Participants have to help the animals achieve their respective goals. For example, in the Toad's task (visuospatial WM), the child is asked to help the toad to clean and cross a lake. The reinforcement scheme is composed of in-game sound feedback given after each attempt and prizes given according to the participant's performance, and all tasks are adaptive. Four out of five tasks stimulate visuospatial WM, and one task stimulates verbal WM, since the target age group are children still in the literacy process.

This study investigated the effect of the PRAMEMT on improving verbal intelligence quotient (VIQ), performance intelligence quotient (PIQ), performance on the full-scale IQ score (FSIQ), and school related abilities: phonological awareness, reading, writing, and mathematics in children aged 6–10 years. Considering the developmental patterns of WM components, the study sample was also stratified into two age groups: younger (6 and 7 years) and older (8–10 years) to analyze the specific gains in the dependent variables.

Considering the association between WM, verbal, nonverbal abilities, and school achievement (Ilkowska & Engle, 2010; Sánchez-Pérez et al., 2018), and recent evidences that show positive effects of WMT on these abilities (Au et al., 2015; Sánchez-Pérez et al., 2018; Zhang et al., 2018), we expected that the PRAMEMT would promote significant effects on verbal (VIQ and verbal WM), nonverbal (PIQ and visuospatial WM), and school related (phonological awareness, reading, writing, and mathematics) abilities. Training effects in abilities similar to the trained ones are found in some studies (Au et al., 2015; Sánchez-Pérez et al., 2018; Zhang et al., 2018), therefore we expected that PRAMEMT would show effects in the WM tasks. Finally, we also expected greater gains in older children, considering that the developmental aspects of WM become more consolidated as children grow older (Carneiro, 2008; Uehara & Landeira-Fernandez, 2010).

2 | MATERIALS AND METHODS

2.1 | Participants

The sample size was calculated based on the magnitude of the effect size for the difference between the control and experimental groups before and after the intervention in typically developing children, based on the meta-analysis by Sala and Gobet (2017). The average effect on verbal and visuospatial WM was calculated only for active control groups. We used the G*Power 3.1 software (Faul et al., 2009) and the mixed repeated measures analysis of variance (ANOVA) (for within and between subjects) was the statistical test selected. In order to find a small effect ($g = 0.46$; $f = 0.23$), assuming a 5% probability of type I error, 95% statistical power, considering two groups and two repeated measures (0.5 correlation between the assessments), a sample of at least 64 individuals would be considered satisfactory. The sample was selected in a non-probabilistic way and consisted of children with TD, aged 6–10 years, from a public school in the city of Belo Horizonte, state of Minas Gerais, Brazil. The children's participation was completely voluntary and prior consent was obtained from their legal guardians. All procedures were performed in accordance with the Declaration of Helsinki. This study was approved by the Research Ethics Committee of the aforementioned Institution under protocol no. 1.016.602. The inclusion criteria were as follows: (1) absence of behavioral indicators of conduct problems, hyperactivity, and relationship problems with peers (indicated on a parent-rated scale); (2) children with general intelligence (intelligence quotient - IQ) ≥70; (3) school performance within the average range (z-scores >-1.5) in mathematics, reading, and spelling.

The initial sample consisted of 122 schoolchildren attending 1st to 5th grades of a public elementary school (Figure 1). Twenty-four children were excluded based on the Strengths and Difficulties Questionnaire (SDQ) behavioral criteria. From the remaining sample, 19 children were excluded due to poor school achievement and low IQ. During the intervention, one child was transferred to another school and was excluded from further analysis. A total of 78 children completed all the cognitive assessment tasks and underwent all intervention sessions (approximately 61% of the initial sample). The mean age of the sample was 7.87 years (standard deviation [SD] = 1.44), 52.6% were girls, and 32.05% were 3rd grade elementary school students. Regarding sociodemographic characteristics, 53.3% of the sample belonged to the lower socioeconomic status. The children were randomly allocated to the experimental or active control group. There was no association between sex and group (control vs. experimental), $\chi^2(1) = 2.45$, $p = .117$; socioeconomic status and group, $\chi^2(4) = 3.39$, $p = .495$; or between school year and group, $\chi^2(4) = 0.38$, $p = .984$.

To investigate the equivalence in the outcome variables between groups in the pretest, a series of t-tests were conducted. There were

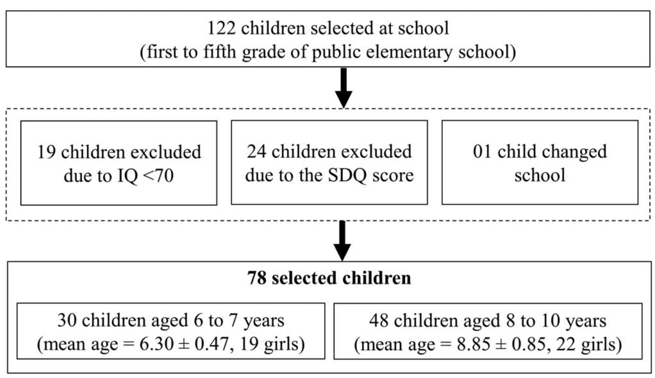

FIGURE 1 Sample selection process

no significant differences between groups in the pretest (all *p*-values > .05). In addition, all the effect sizes (Cohen's *d*) could be considered small using the Cohen's (1988) criterion, and varied between −0.07 and 0.40. The groups did not differ regarding age: $t(76) = 0.12$, $p = 0.908$, $d = 0.03$. For all comparisons of means and test statistics in the pretest, please refer to Table S1 in the Data S1. Aiming to examine the effect of the PRAMEMT by age, the sample was divided into two groups: the first consisted of children aged 6 and 7 years, whereas the second comprised children aged 8–10 years.

2.2 | Younger sample (children aged 6 and 7 years)

This group was composed of 30 schoolchildren (mean age = 6.30 years, SD = 0.47; 63.3% girls). Data analysis showed no association between gender and group (control vs. experimental), $\chi^2(1) = 2.01$, $p = .156$; socioeconomic status and group, $\chi^2(3) = 0.72$, $p = .868$; or between school year and group, $\chi^2(2) = 0.27$, $p = .875$.

A series of *t*-tests was performed to assess equivalence in pretest scores in the younger group. The groups did not differ according to age, $t(28) = -0.94$, $p = .355$, $d = -0.31$. In general, the results suggest that the groups did not differ in any of the investigated variables. However, moderate effect sizes were found in the PIQ, $t(25.18) = 1.94$, $p = 0.064$, $d = 0.67$ (positive values suggest higher means in the control group). In addition, the other effect size values can be considered small, varying between 0.02 and 0.49. For complete comparisons between means and test statistics in the pretest, please refer to Table S1 in the Data S1.

2.3 | Older sample (children aged 8–10 years)

This group comprised 48 schoolchildren (mean age = 8.85 years; SD = 0.85 years; 45.8% female). There was no association between sex and type of group (control vs. experimental), $\chi^2(1) = 0.80$, $p = .371$; socioeconomic status and type of group, $\chi^2(4) = 5.53$, $p = .237$; or between school year and type of group, $\chi^2(2) = 0.257$, $p = .880$.

A series of *t*-tests was carried out to investigate differences between groups in the pretest. The results are presented in Table S1. Results suggest that groups showed differences only in phonological awareness in favor of the control group, $t(46) = 2.01$, $p = .051$, $d = 0.57$. For all other variables, the *d* values could be considered small, ranging between ~0 and 0.34. The groups did not differ regarding age, $t(46) = 0.46$, $p = .650$, $d = 0.13$. For complete comparisons between means and test statistics in the pretest, please refer to Table S1 in the Data S1.

2.4 | Instruments

Strengths and Difficulties Questionnaire: The SDQ (Muris et al., 2003; Saur & Loureiro, 2012) is a screening test of mental disorders that assesses behaviors of teenagers and children aged 4–16 years. The questionnaire contains 25 items assessing Emotional Symptoms, conduct problems, Hyperactivity/Inattention, Peer Problems, and Prosocial Behavior. For the present study, only the parent's version was used. Responses are based on a 3-point Likert scale (0 = "not true"; 1 = "somewhat true"; 2 = "certainly true"). For the five subscales, the summed scores can vary between 0 and 10. The total problems score is found by summing up all sub-scores, except for prosocial behavior. The cutoff score for total difficulties was 17. For the parent's version, the internal consistency of the total score was 0.80 (Saur & Loureiro, 2012).

The Wechsler Abbreviated Scale of Intelligence (WASI): The WASI (Irby & Floyd, 2013; Trentini et al., 2014) is used to assess the intelligence of people aged 6–89 years. The WASI is composed of four subtests: Vocabulary and Similarities, which evaluate verbal or crystalized abilities (VIQ), and Block Design and Matrix Reasoning, which evaluate nonverbal or fluid abilities (PIQ). Scores on the four subtests are summed for the FSIQ. The average internal consistency values for the age group of 6–16 years presented in the manual were .89 (VIQ) .89 (PIQ), and .92 (FSIQ), suggesting high reliability.

Brazil Socioeconomic Questionnaire: The Brazil Socioeconomic Questionnaire (Questionário Socioeconômico – Critério Brasil; Brasil, 2018) was used to measure socioeconomic status. It is a standardized questionnaire that measures socioeconomic characteristics such as marital status, number of children, educational level, and family income. The questionnaire was answered by the child's parent and/or legal guardian.

Brief Neuropsychological Assessment Battery: The Brief Neuropsychological Assessment Battery (Neupsilin-Inf) (Salles et al., 2015) provides a neuropsychological profile for children aged 6–12 years. The battery assesses eight neuropsychological functions: time and space orientation, attention, visual perception, memory, mathematical abilities, language, visuoperceptual abilities, and EFs. However, in this study, we used only the subscales of WM (verbal and visuospatial), language (phonological awareness, reading, and spelling) and mathematical abilities. Almost all Neupsilin-Inf tasks present internal consistency ≥0.60 (73.9% of tasks) and temporal stability ranges from 0.61 to 0.94. The raw scores were transformed to age standardized z-scores according to reference values of the Neupsilin-Inf. The following tasks were used in the present study:

Phonological awareness: In the first task, the examiner says three words and asks the child to indicate the two that rhyme. In the second task, the examiner says a word and excludes the first phoneme. Then the child should say what the word without the phoneme is. These tasks measure phonological awareness. The total score of phonological awareness has a temporal stability coefficient (*r*) of 0.74.

Reading: In this task, the examiner presents words to the child and then the child reads them. The total score is the sum of words read correctly. This task has $r = 0.91$.

Spelling: In this task, the examiner reads several words aloud. Then the child is asked to write the words down. The total score is the sum of the correctly written words. This task has $r = 0.94$.

Mathematical abilities: It is composed of two parts. First, the child counts how many matchsticks are in the image presented by the examiner. The examiner then dictates an arithmetic expression that should be written and solved by the child. The total score is the sum of the correct arithmetic expressions and the correct count. This task has $r = 1.00$.

Digit span: In this task, the examiner asks the child to repeat numbers that were dictated out loud. This task measures phonological short-term in the forward order (i.e., the child repeats the numbers in the same order they were presented) and verbal WM in the backward order (i.e., the child repeats the sequence in the inverse order). The span ranges between 2 and 5, with eight items for each order. The total score is the sum of the items repeated correctly (maximum = 28). This task has $r = 0.57$.

Pseudoword repetition: In this task, the child is asked to repeat pseudowords in the same order they had been presented by the examiner. The task contains eight items and the span varies from 1 to 4. The score is calculated by summing the pseudowords repeated correctly (maximum total of 20). This task has $r = 0.68$. The sum of the backward digit span and pseudoword repetition scores is the VWM total score.

Visuospatial working memory (VSWM): In this task, the child taps squares on a sheet of paper in backward order of the examiner. This task contains eight items and span varies between 2 and 5. The total score is calculated by summing up all the correct repetitions (maximum total of 28). This task has $r = 0.66$. The sum of the VWM and VSWM scores is the WM score.

2.4.1 | Cognitive training software

Working Memory Activation Program (Figure 2): This WMT program has an adaptive system that stimulates three WM components: simple span, complex span, and updating. Simple span tasks require only information storage and recall in the same serial order. Complex span tasks demand storage and processing. Finally, updating tasks require previously acquired information to update with the new information in order to manipulate the first. In the *n*-back tasks, for example, a set of images is presented to the child one at a time and they are asked to press the target stimulus when the object shown is equal to the one shown *n* times before. The target *n* increases steadily.

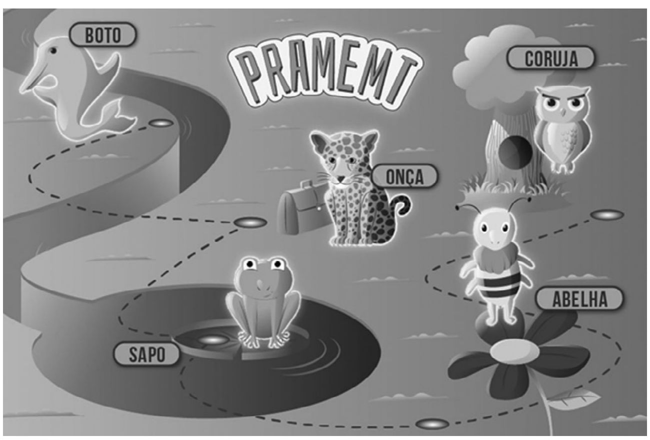

FIGURE 2 Working Memory Activation Program's (PRAMEMT) tasks. "boto" = river dolphin, "sapo" = toad, "onça" = jaguar, "abelha" = bee, "coruja" = owl. Source: Barbosa-Pereira et al. (2019)

3 | MEMORY ACTIVATION PROGRAM

This software is composed of five training tasks featuring animals from the Brazilian fauna: owl, bee, toad, jaguar, and river dolphin. The content validity of the PRAMEMT was investigated in a previous study. In general, the software presented acceptable indices of content validity and was readily understood by children aged between 5 and 8 years (for full descriptions, see Barbosa-Pereira et al., 2019). The program's adaptive system increases or decreases the level of difficulty according to the participant's performance, optimizing the learning process (Foster et al., 2017; Landsberg et al., 2012). On the owl, toad, jaguar, and river dolphin tasks, the child starts at an easier level and progresses according to their correct score. If the child answers at least 85% of the 5 initial trials correctly, they move on to the next level. Otherwise, the child continues to do the second part with more 5 trials. When the child reaches 85% or more correct answers in this part, they move on to the next level. A similar procedure is adopted in the other levels. However, if the child scores less than 70%, they move back to the previous level. The game is finished if the child repeats this situation twice. When repetition occurs on the starting level, the task is finished. In this sense, some children may stop training on some tasks before others.

The bee task has a different adaptive system. In the first part, which contains 15 trials, the child moves on to the next level when making fewer than three errors. If the child answers three or more trials wrongly, they should do 15 more trials of part 2 at the same level. In this situation, if the child scores 5 or higher and has more than three errors, they move back to the previous level. When this occurs at the first level, the game is finished. During the intervention, participants received two sources of reinforcement feedback. The program features sounds indicating whether the participants' responses are correct. In addition, the examiner awards the children with small stickers according to their performance to complete a sticker album. The stickers earned by the participants are similar to the digital rewards shown in the PRAMEMT main screen. The tasks featured in the PRAMEMT are described below:

Owl task (visuospatial, simple span): In this task, an owl appears randomly in a set of trees (positions) and the child must memorize the positions the animal appeared in the forward order. After that, the child must tap the trees. This task has six levels of difficulty that increase as the task progresses.

Bee task (visuospatial, n-back): In this task, the child must indicate whether or not the flower on the screen was the same color as the flower shown *n* times before. The task has four levels.

Toad task (visuospatial, complex span): In this task, a toad randomly appears on lily pads (positions) in a lagoon and the child must memorize the positions the animal appeared. After that, the child must indicate the positions of the toad in backward order. This task has four levels of difficulty that increase as the task progresses.

Jaguar task (verbal, complex span): In this task, the child must indicate the different color stimuli that appear on the screen. After that, a screen with eight colors is presented to the child and they must indicate the previously identified stimuli in the order they appeared. At the first level, the participant is only required to memorize two colors. At the fourth (and final) level, five stimuli must be identified and memorized.

River Dolphin task (verbal, complex span): In this task, the child is asked to count the number of white boats that appear on the screen. Afterwards, the child is asked to indicate the exact number of white boats that appeared on the screen. The answer is given in a matrix ranging from 1 to 9. This task has four levels.

3.1 | Active control group task

The control group participated in a quiz that featured the same animals and themes as those of the PRAMEMT. The task had five modules with 15 questions each. The questions were presented using the PowerPoint® software. If the child could not read, the examiner would read the questions out loud. Each question had two options and one correct answer. After the child had answered five questions in a module, they received a reinforcement similar to that of the PRAMEMT rewards (stickers). The questions were about the habitat of the animals, what they eat, where they are usually found in Brazil, and their physical characteristics and common habits. The questions were answered in groups or pairs of children. Unlike in the experimental group, no child performed the task of the active control group individually. In addition, every answer was also followed by a sound that indicated whether it was correct or incorrect.

3.2 | Data collection

Data collection occurred between August 2018 and June 2019. The sample was randomly assigned to either the control or the experimental group. All children were assessed in the pretest for intelligence, cognitive, and school-related abilities. The examiners were undergraduate and graduate psychology students who had been previously trained. The cognitive training and control group activities occurred in five weekly sessions, of about 20 minutes. After the fifth session, a posttest assessment was conducted. Pretest and posttest were carried out in two sessions for each child. Each assessment session lasted approximately 40 min.

The experimental group tasks were performed individually using a tablet. A member of the research team was always together with children for answering their doubts and questions. The children listened to the instructions on earphones. An animal was chosen for each training session. In contrast, the activities were conducted in group in the active control group. The questions were shown to the participants using a tablet and the guidelines were presented by the research team. Unlike in the cognitive training, no instructions were previously recorded. Each meeting had an animal as the theme.

3.3 | Data analysis

The analyses were conducted using the SPSS 20.0 and JASP 0.14.1.0 software programs. Prior to all analyses, we screened our data for normality using the Shapiro–Wilk test. In addition, we searched for differences in the means between the experimental and control groups in the pretest assessment using a series of t-tests. The PRAMEMT software effect on promoting cognitive gains was investigated using 2×2 mixed effects repeated measures ANOVA. As there is a variety of effect sizes used for evaluating ANOVA, we relied on the generalized eta-square ($\eta^2 g$). One could argue that $\eta^2 g$ can be both a reliable effect size indicator and a comparable measure between different ANOVA designs (Olejnik & Algina, 2003). Moreover, aiming to rely on an effect size metric that is common in the literature, we also calculated the Cohen's d. The calculation considered differences between the scores in both assessment times (posttest - pretest). All d values were corrected for comparisons using small samples (Durlak, 2009 - Equation (2)). Whenever a difference was found in the pretest, ANCOVA was conducted using gain (posttest - pretest) as the dependent variable and the scores on the pretest as the covariate. For analyses using covariates, Cohen's d was calculated using the estimated marginal means from the ANCOVA measures. The same procedures were applied to the overall sample and by age group. In order to avoid type-I errors during hypothesis testing with multiple dependent variables, we also relied on the Holm-Bonferroni method to correct our p-values to an alpha level of .05 (Holm, 1979).

4 | RESULTS

4.1 | Effect on the overall sample

Results of the Shapiro–Wilk test suggest that 70% of all the conditions (control vs. experimental in pretest vs. posttest) could be treated essentially as normally distributed variables. The results regarding the gains from the overall sample are presented in Table 1. Taking a significance level of 5% ($p < .05$) as reference, there were only two significant interactions. For PIQ, the experimental group showed greater gains compared with those of the control group, [$F(1, 76) = 5.87$, $p = .018$, $\eta^2 g = .01$, $d = -0.54$]. Similar results were found for phonological awareness, [$F(1, 76) = 5.59$, $p = .021$, $\eta^2 g = .02$, $d = -0.52$]. The eta effect sizes were small in both cases (Cohen, 1988). With the p-values corrected for multiple comparisons, there was no significant effect of time vs. group.[1] All effect sizes are displayed in Figure 3. Descriptive statistics for all conditions are shown in the Data S1.

4.2 | Effect on the younger sample (children aged 6–7 years)

All intervention effect coefficients are presented in Table 1. Considering that a moderate effect size was found in the pretest for PIQ, this variable was entered as covariate in the ANOVA models. Specifically,

TABLE 1 Test statistics for mean gains in the overall sample and by age group

	Overall sample						Younger sample						Older sample					
	CG (n = 37) M (SD)	EC (n = 41) M (SD)	t	df	p	d	CG (n = 14) M (SD)	EC (n = 16) M (SD)	t	df	p	d	CG (n = 23) M (SD)	EC (n = 25) M (SD)	t	df	p	d
Age (years)	7.89 (1.51)	7.85 (1.41)	0.116	76	0.908	0.03	6.21 (0.43)	6.38 (0.50)	−0.94	28	0.355	−0.31	8.91 (0.85)	8.80 (0.87)	0.456	46	0.65	0.13
Measure	Effect	df	MSE	F	p	η^2_G	Effect	df	MSE	F	p	η^2_G	Effect	df	MSE	F	p	η^2_G
Digit span backward	Time	1, 76	2.6	6.31	0.014	0.02	Time	1, 27	0.90	2.17	0.152	0.02	Time	1, 45	0.37	0.91	0.345	0.00
	Time * group	1, 76	0.12	0.3	0.587	0.00	Time * group	1, 27	0.25	0.60	0.444	0.01	Time * group	1, 45	0.13	0.33	0.571	0.00
	group	1, 76	2.48	1.61	0.208	0.02	group	1, 27	0.42	0.29	0.594	0.01	group	1, 45	0.28	0.19	0.666	0.00
							Performance IQ (covariate)	1, 27	8.87	6.10	0.020	0.15	Phonological awareness (covariate)	1, 45	0.05	0.03	0.861	0.00
Pseudoword repetition	Time	1, 76	2.33	3.77	0.056	0.01	Time	1, 27	0.10	0.20	0.662	0.00	Time	1, 45	1.6	2.19	0.146	0.01
	Time * group	1, 76	2.02	3.27	0.075	0.01	Time * group	1, 27	1.75	3.61	0.068	0.02	Time * group	1, 45	0.32	0.43	0.514	0.00
	group	1, 76	7.69	3.93	0.051	0.04	group	1, 27	2.84	1.28	0.269	0.04	group	1, 45	5.62	3.05	0.088	0.05
							Performance IQ (covariate)	1, 27	0.19	0.09	0.772	0.00	Phonological awareness (covariate)	1, 45	0.42	0.23	0.635	0.00
Verbal working memory	Time	1, 76	3.99	10.97	0.001	0.03	Time	1, 27	0.34	1.01	0.324	0.01	Time	1, 45	1.46	3.7	0.061	0.02
	Time * group	1, 76	0.83	2.27	0.136	0.01	Time * group	1, 27	0.96	2.83	0.104	0.02	Time * group	1, 45	0.3	0.75	0.391	0.00
	group	1, 76	6.84	4.23	0.043	0.04	group	1, 27	1.35	0.85	0.365	0.03	group	1, 45	3.08	1.85	0.18	0.03
							Performance IQ (covariate)	1, 27	4.76	3.00	0.095	0.08	Phonological awareness (covariate)	1, 45	0.06	0.04	0.851	0.00
Visuospatial working memory	Time	1, 76	13.78	31.95	< .001	0.08	Time	1, 27	0.33	0.73	0.400	0.01	Time	1, 45	7.09	16.34	< .001	0.08
	Time * group	1, 76	0.09	0.2	0.654	0.00	Time * group	1, 27	0.15	0.33	0.570	0.00	Time * group	1, 45	0.35	0.8	0.375	0.00
	group	1, 76	0.05	0.03	0.872	0.00	group	1, 27	0.01	0.00	0.950	0.00	group	1, 45	2.75	2.01	0.163	0.03
							Performance IQ (covariate)	1, 27	13.95	7.75	0.010	0.19	Phonological awareness (covariate)	1, 45	2.82	2.06	0.159	0.03
Total working memory	Time	1, 76	13.59	38.36	< .001	0.07	Time	1, 27	0.02	0.05	0.822	0.00	Time	1, 45	6.57	17.87	< .001	0.07
	Time * group	1, 76	0.57	1.62	0.207	0.00	Time * group	1, 27	0.13	0.36	0.553	0.00	Time * group	1, 45	0.59	1.6	0.213	0.01
	group	1, 76	2.1	1.05	0.309	0.01	group	1, 27	0.12	0.06	0.815	0.00	group	1, 45	0.09	0.05	0.823	0.00
							Performance IQ (covariate)	1, 27	12.97	5.95	0.022	0.16	Phonological awareness (covariate)	1, 45	0.73	0.42	0.521	0.01

TABLE 1 (Continued)

	Overall sample						Younger sample						Older sample					
	CG (n = 37) M (SD)	EC (n = 41) M (SD)	t	df	p	d	CG (n = 14) M (SD)	EC (n = 16) M (SD)	t	df	p	d	CG (n = 23) M (SD)	EC (n = 25) M (SD)	t	df	p	d
Verbal IQ	Time	1, 76	75	1.75	0.19	0.00	Time	1, 27	94.16	1.72	0.201	0.01	Time	1, 45	43.89	1.18	0.284	0.00
	Time * group	1, 76	26.64	0.62	0.433	0.00	Time * group	1, 27	26.30	0.48	0.494	0.00	Time * group	1, 45	13.59	0.36	0.549	0.00
	group	1, 76	211.52	0.51	0.478	0.01	group	1, 27	166.77	0.45	0.510	0.01	group	1, 45	0.36	0	0.976	0.00
							Performance IQ (covariate)	1, 27	40.56	0.11	0.745	0.00	Phonological awareness (covariate)	1, 45	373.84	0.93	0.34	0.02
Performance IQ	Time	1, 76	546.37	29.7	< .001	0.03	group	1, 27	54.36	1.84	0.186	0.06	Time	1, 45	513.1	27.44	< .001	0.05
	Time * group	1, 76	108.06	5.87	0.018	0.01	Performance IQ (covariate)	1, 27	129.68	4.39	0.046	0.13	Time * group	1, 45	71.06	3.8	0.058	0.01
	group	1, 76	260.99	1.3	0.257	0.02							group	1, 45	0.01	0	0.994	0.00
													Phonological awareness (covariate)	1, 45	528.81	2.46	0.124	0.05
Full scale IQ	Time	1, 76	369.41	14.09	< .001	0.02	Time	1, 27	86.61	3.01	0.094	0.02	Time	1, 45	336.3	13.96	< .001	0.02
	Time * group	1, 76	6.41	0.25	0.622	0.00	Time * group	1, 27	0.16	0.01	0.941	0.00	Time * group	1, 45	4.52	0.19	0.667	0.00
	group	1, 76	313.84	1.07	0.305	0.01	group	1, 27	6.54	0.04	0.839	0.00	group	1, 45	0.19	0	0.981	0.00
							Performance IQ (covariate)	1, 27	1796.90	11.57	0.002	0.27	Phonological awareness (covariate)	1, 45	635.53	2.07	0.158	0.04
Phonological awareness	Time	1, 76	0.44	1.77	0.188	0.01	Time	1, 27	0.13	1.83	0.188	0.00	group	1, 45	1.09	2.53	0.119	0.03
	Time * group	1, 76	1.4	5.59	0.021	0.02	Time * group	1, 27	0.01	0.19	0.664	0.00	Phonological awareness (covariate)	1, 45	12.85	29.96	< .001	0.39
	group	1, 76	0.23	0.23	0.631	0.00	group	1, 27	0.09	0.06	0.813	0.00						
							Performance IQ (covariate)	1, 27	2.05	1.27	0.270	0.04						
Mathematics	Time	1, 76	7.36	30.19	< .001	0.07	Time	1, 27	0.31	1.17	0.290	0.01	Time	1, 45	3.8	15.78	< .001	0.09
	Time * group	1, 76	0.85	3.48	0.066	0.01	Time * group	1, 27	0.62	2.30	0.141	0.01	Time * group	1, 45	0.51	2.13	0.152	0.01
	group	1, 76	0.18	0.16	0.686	0.00	group	1, 27	0.09	0.06	0.808	0.00	group	1, 45	0	0	0.972	0.00
							Performance IQ (covariate)	1, 27	2.35	1.65	0.210	0.05	Phonological awareness (covariate)	1, 45	1.29	1.93	0.171	0.03
Reading	Time	1, 76	0.02	0.05	0.831	0.00	Time	1, 27	0.07	0.58	0.455	0.00	Time	1, 45	0.25	0.34	0.561	0.00
	Time * group	1, 76	0.05	0.11	0.746	0.00	Time * group	1, 27	0.24	1.88	0.181	0.01	Time * group	1, 45	0.19	0.26	0.613	0.00

(Continues)

TABLE 1 (Continued)

	Overall sample						Younger sample						Older sample					
	CG (n = 37) M (SD)	EC (n = 41) M (SD)	t	df	p	d	CG (n = 14) M (SD)	EC (n = 16) M (SD)	t	df	p	d	CG (n = 23) M (SD)	EC (n = 25) M (SD)	t	df	p	d
	group	1, 76	2.34	2.2	0.143	0.02	group	1, 27	0.40	0.23	0.637	0.01	group	1, 45	1.26	1.88	0.177	
							Performance IQ (covariate)	1, 27	2.19	1.25	0.273	0.04	Phonological awareness (covariate)	1, 45	0.26	0.39	0.538	0.00
Spelling	Time	1, 76	5.75	18.78	< .001	0.04	Time	1, 27	0.20	1.11	0.302	0.00	Time	1, 45	2.49	6.5	0.014	0.03
	Time * group	1, 76	0.03	0.1	0.756	0.00	Time * group	1, 27	0.24	1.36	0.254	0.00	Time * group	1, 45	0.09	0.25	0.623	0.00
	group	1, 76	0.22	0.15	0.696	0.00	group	1, 27	0.53	0.29	0.596	0.01	group	1, 45	0.39	0.32	0.574	0.01
							Performance IQ (covariate)	1, 27	0.58	0.31	0.582	0.01	Phonological awareness (covariate)	1, 45	1.98	1.64	0.207	0.03

Abbreviations: df, degrees of freedom; M, mean; MSE, mean-squared error; η^2G, generalized eta square.

the analysis was performed using the gain variable (performance in the posttest minus the score in the pretest) in an univariate covariance model. None of the interactions were significant. Moreover, there was also a moderate gain (Cohen's *d*) in favor of the control group for verbal WM [$F(1, 27) = 2.83$, $p = .104$, $\eta^2 g = .02$, $d = 0.62$], pseudoword repetition [$F(1, 27) = .61$, $p = .068$, $\eta^2 g = .02$, $d = 0.70$], reading [$F(1, 27) = 1.88$, $p = .181$, $\eta^2 g = .01$, $d = 0.51$], and mathematics [$F(1, 27) = 2.30$, $p = .141$, $\eta^2 g = .01$, $d = 0.56$].

4.3 | Effect on the older sample (children aged 8–10 years)

Considering that there were significant differences in phonological awareness in the pretest, this variable was entered as a covariate in all models. All coefficients and descriptive statistics of the mean gains are presented in Table 1. None of the models suggested significant interactions. However, there was a moderate effect size, suggesting that the experimental group had a different pattern of gains in PIQ compared with the control group [$F(1, 45) = 3.80$, $p = .058$, $\eta^2 g = .01$, $d = -0.56$]. For the remaining variables, the effect sizes were small.

5 | DISCUSSION

This study investigated the effect of a WMT, the PRAMEMT (Barbosa-Pereira et al., 2019), in improving cognitive and school performance in a sample of children aged 6–10 years. In the present study, 78 schoolchildren were assessed and participated in the intervention program. Moreover, in order to deeply investigate developmental effects in gains, we stratified our sample into two subgroups: younger (aged 6–7 years) and older (aged 8–10 years) children. The results were similar in the overall sample and in each age group. We found evidence of gains in PIQ and phonological awareness, especially when the effect sizes were considered.

The Block Design and Matrix Reasoning WASI subtests (Trentini et al., 2014) were the assessment tools most similar to the training content. In that sense, considering that the PRAMEMT focuses on visuospatial WM tasks, the mean gain of ~5 points in PIQ could be partially explained by this similarity. Nonetheless, there is consistent evidence that fluid intelligence and WM are closely related (Little et al., 2014) and even share common neural networks (Jung & Haier, 2007). However, in the younger age group, the mean gain in PIQ was not significant, even though the effect size was almost moderate ($d = -0.47$). Moreover, for the older subsample, there was a closeness to significance and a moderate effect size ($p = .058$, $d = -0.56$). The sample size may have biased the estimation process, underestimating the pattern observed in the whole sample. The results for fluid intelligence were not the same as those found in the review conducted by Mansur-Alves and Saldanha-Silva (2017). Those authors observed small, limited gains for fluid intelligence ($g = 0.09$). However, it should be noted that the gains in PIQ were not found in

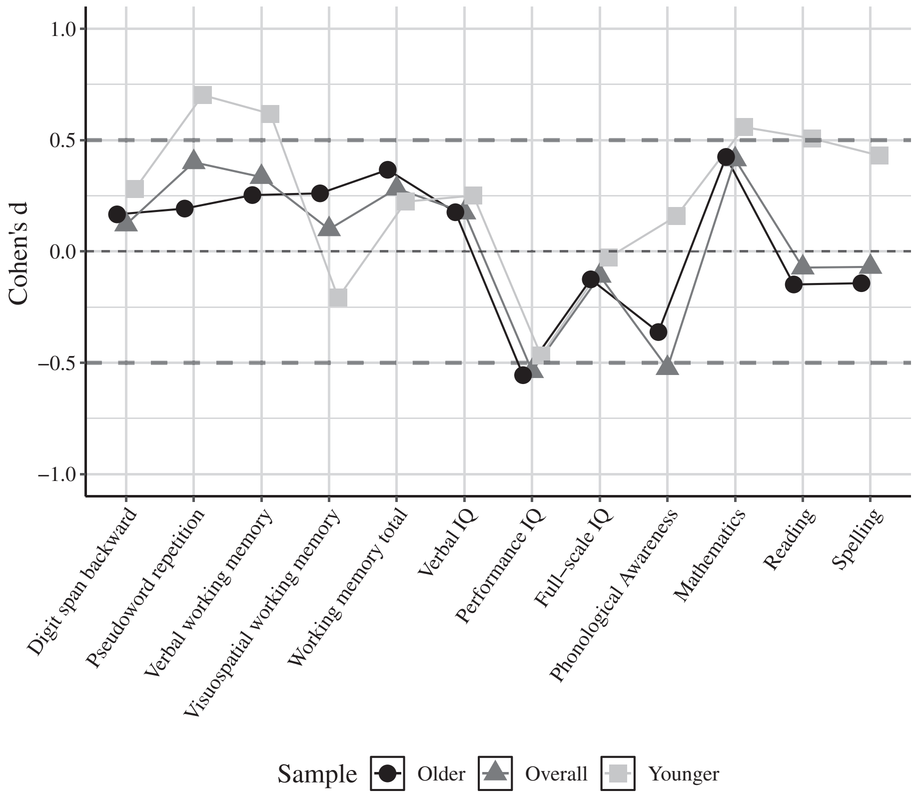

FIGURE 3 Effect sizes for mean gains in the overall sample and by age group

the FSIQ score. This finding suggests that the intervention was not fully able to improve the children's general cognitive abilities. Changes in general intelligence are hardly ever achieved in cognitive training studies (; Sala et al., 2019).

With regard to phonological awareness, there was a mean gain of ~0.30 standardized unit measures and a $d = -0.52$ in favor of the experimental group. The abilities required in phonological awareness tasks (rhyme detection and phonemic elision) are closely related to verbal WM capacity (Knoop-van Campen et al., 2018). The phonemic elision tasks are even more similar to the verbal WM tasks, requiring verbal manipulation of information (Lopes-Silva et al., 2014). Currently, the most accepted model for phonological processing includes lexical retrieval, phonological awareness, and verbal WM as similar and correlated components (Wagner & Torgesen, 1987). The complexity of the relationship between verbal WM and phonological awareness has been explored in different ways, such as in mediational models for the prediction of reading (Knoop-van Campen et al., 2018) and number writing (Lopes-Silva et al., 2014). For the younger and older samples, however, there was no significant interaction for gains in phonological awareness, ($p = .664$, $d = 0.16$) and ($p = .119$, $d = -0.36$), respectively. The larger effect size in the older sample can be justified by the development of phonological awareness (Snowling & Hulme, 1994); therefore, the training effects may be more evident in this sample.

No significant gains were observed for the remaining variables. In the overall sample, the effect sizes of school-related abilities were small (mathematics: $d = 0.41$; reading: $d = -0.07$; spelling: $d = -0.07$). According to Melby-Lervåg et al. (2016), Sala and Gobet (2017), and Mansur-Alves and Saldanha-Silva (2017), evidence of WMT on cognitive and school-related abilities are rare and tends to occur only in experimental studies with small sample sizes and passive control groups. On one hand, there was no significant interaction in the older sample and all effect sizes were small; on the other hand, the effect sizes favored the control group for school-related abilities (mathematics: $d = 0.56$; reading: $d = 0.51$; spelling: $d = 0.43$) in the younger sample. In addition, in the younger sample, there were moderate effects in pseudoword repetition ($d = 0.70$) and verbal WM ($d = 0.62$) in favor of the control group. These effects favoring the control group could be explained by the nature of our study design. The control group “intervention” was conducted in small groups that were given a simulated quiz on the general characteristics of the animals featured in the PRAMEMT. Consequently, this group had more social interaction with their peers and examiners. For example, for the younger children who did not have fully developed reading skills, the examiner had to read the instructions out loud to them. Therefore, one could assume that the social motivation factor regarding the assessment process was higher in the control group. This could have made the qualities of this group more evident. However, we intended to control the participants' motivation using two reinforcement schemes for both groups (fully described in the Methods section). Motivation is an important factor in training and learning in general (Katz et al., 2014, 2018). Hence, higher motivation and engagement in the assessment and intervention process will likely lead to improvement in behavioral measures. In addition, although group allocation was random, in the younger age group, children from the experimental group scored ~8 IQ points less than those in the older experimental group. Therefore, it is possible to argue that this difference in intelligence could have impacted the general understanding of the intervention.

In the overall sample, no effects were observed for gains in WM. Nonetheless, the cognitive tasks employed to investigate WM gains presented low variability in our sample. Some participants reached the maximum score for their age group, even though it was not the ceiling effect of the task (the standardized scores in the Neupsilin-Inf manual were calculated by age). At the time of data collection, the Neupsilin-Inf was the only standardized instrument available in Brazil to assess WM in children aged 6–10 years. In addition, the scoring criteria applied in the task manual may not have been fully adequate to discriminate changes in the latent trait. The WM scores are calculated by simply summing up all the stimuli retrieved by the child, even if the full sequence is not fully correct. For instance, if a child had to remember a five-number sequence (i.e., 2, 7, 4, 5, and 8) and their response was only for the first three stimuli, they would receive three points instead of zero (similar to the scoring criteria used for the WASI). Considering this, two individuals could receive similar scores even though their span capacities were different. The Neupsilin-Inf is currently used as a brief screening tool for developmental disabilities; therefore, not all cognitive domains were deeply covered. As outlined by Pergher et al. (2019), tasks that do not present rigorous psychometric properties may not reflect true training effects. In addition, in the present study, we did not assess the updating component of WM, even though this cognitive domain was stimulated by the PRAMEMT tasks (specifically, the Bee task). Consequently, one could argue that the full effects on WM investigation was not complete, mainly because there were no standardized measures available to assess the updating component of WM. The same pattern observed in the overall sample was also applied to the two age groups.

Duration of training and its relation to effects on cognitive and school-related abilities are still open to debate. For instance, there is mixed evidence regarding the number of training sessions and improvement in cognitive performance. According to Yang et al. (2017), dyslexic children who underwent 15 WMT sessions showed significant gains in orthographic knowledge ($\eta^2 p = 0.20$). The number of hours of intervention in the present study was not optimal according to Mansur-Alves and Saldanha-Silva (2017). Cognitive training studies should last at least 8 h to be effective (Mansur-Alves & Saldanha-Silva, 2017, but see Sala & Gobet, 2017 for results that point to a non-significant moderation effect on training duration). However, we were only able to perform ~1 h and 15 min of training for each child.

We were only able to perform the single-blind procedure due to financial and human resource limitations. Only study participants did not know which condition (control vs. experimental) they were allocated. Another possible limitation is the active control group design of the study. As previously described, interaction between the examiners and the children in the control group seems to have been highly motivating for the children, even more than the application of the PRAMEMT. Thus, interest in the intervention/assessment process may have impacted the posttest results in favor of the control group. Furthermore, we were unable to control some variables that could influence the transfer effects, such as personality and attachment type. Nevertheless, this study also has several strengths, such as rigorous methods (random allocation and active control group) and a WMT fully developed for Brazilian children with reinforcement schemes. In addition, the WMT used was content validated (Barbosa-Pereira et al., 2019), which is uncommon for training instruments (Golino et al., 2017). Even though we have tried to overcome some challenges, such as those pointed out by Pergher et al. (2019), in the field of WM training, further studies are needed to improve methodological rigor, as we discussed previously.

5.1 | Final remarks

This study investigated the effects on cognitive and school-related abilities of the PRAMEMT, a WMT program, in children aged 6–10. We highlight that the number of children trained in the present study was higher than that in the majority of published studies (Sala & Gobet, 2017). The results are highly relevant, because only a few WMT effect studies have been conducted with Brazilian children, and even fewer WMT protocols have been constructed to address national characteristics (except for the studies by Mansur-Alves et al., 2013; Ramos & Melo, 2016; Alves & Bonfim, 2016). The program seems to promote specific gains in phonological awareness and PIQ; however, these improvements were not translated to general cognitive abilities. Future studies should address unanswered questions, such as how the training protocol compares with a paper-and-pencil version. Furthermore, the training duration could be modified to achieve a longer protocol. Moreover, even though we were able to find gains in some cognitive domains, the durability of such gains should be further investigated using follow-up designs. Finally, to the best of our knowledge, the present study was the first to investigate the effect of a WMT method developed for Brazilian children with two reinforcement schemes (in-game sounds and feedback). This study used a robust design, with random allocation and an active control group, and the results can potentially advance the discussion on the effect of WMT in children with TD.

ACKNOWLEDGMENTS

The study was funded by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Brasil (FAPEMIG), Grant number APQ-00802-15.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

ETHICS STATEMENT

All procedures were performed in accordance with the Declaration of Helsinki. The study was approved by our local ethical committee (approbation number: 1.016.602).

INFORMED CONSENT

The children's participation was completely voluntary and was conditioned to the consent of their legal guardian. All authors have read the final version of the manuscript and authorize the publication.

ENDNOTE

[1] We tested for outliers using the extreme value criteria (Tabachnick & Fidell, 2018). For this, a z-score higher than the absolute value of 3.29 is considered a potential outlier. Screening our data, we found three cases that could be potential outliers. In the pseudoword repetition: a z-score of −3.18 in the pretest (experimental group) and another case with a z-score of 3.82 in the posttest (control group). In addition, in the reading task we found another extreme value of 5.48 in the pretest (control group). Additional analyses using the Grubbs's outlier test revealed that only the scores of 5.48 and 3.82 were significant outliers. Therefore, we repeated the analyses without the outliers, but the results were largely the same. The results are, then, reported using the full data.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Drielle Barbosa-Pereira https://orcid.org/0000-0001-5027-9976
Pedro S. R. Martins https://orcid.org/0000-0002-6555-7649
Luiz Alves Ferreira-Junior https://orcid.org/0000-0003-2209-9323
Renata Saldanha-Silva https://orcid.org/0000-0003-2513-1106
Marcela Mansur-Alves https://orcid.org/0000-0002-3961-3475

REFERENCES

Alloway, T. P., Bibile, V., & Lau, G. (2013). Computerized working memory training: Can it lead to gains in cognitive skills in students? *Computers in Human Behavior*, *29*(3), 632–638. https://doi.org/10.1016/j.chb.2012.10.023

Alves, L., & Bonfim, C. (2016). Gamebook e a estimulação de funções executivas em crianças com indicação de diagnóstico de TDAH: processo de pré produção, produção e avaliação do software. *Educação e Contemporaneidade*, *25*(46), 141–157 https://www.revistas.uneb.br/index.php/faeeba/article/view/2723

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin and Review*, *22*(2), 366–377. https://doi.org/10.3758/s13423-014-0699-x

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D. (2017). *The concept of working memory: A view of its current state and probable future development. In exploring working memory* (pp. 99–106). Routledge.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation*. Academic Press.

Barbosa-Pereira, D., Luiz, A., Ferreira, J., Rodrigues De Souza, E., De Almeida Galvão, H., Meirelles Mendonça, C., Saldanha-Silva, R., & Mansur-Alves, M. (2019). Desenvolvimento e Validade de Conteúdo de um Programa Computadorizado de Treino Cognitivo para Crianças. *Revista Neuropsicologia Latinoamericana*, *11*(3), 52–65. https://doi.org/10.5579/rnl.2016.0450

Brasil, C. D. C. (2018). *Critério de classificação econômica Brasil. Associação Brasileira de Empresas de Pesquisa (ABEP)*. Associação Brasileira de Empresas de Pesquisa. http://www.abep.org/criterio-brasil

Carneiro, M. P. (2008). Memory development in children: What changes with age? *Psicologia: Reflexao e Critica*, *21*(1), 51–59. https://doi.org/10.1590/s0102-79722008000100007

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004

Fälth, L., Jaensson, L., & Johansson, K. (2015). Working memory training-A Cogmed intervention. *International Journal of Learning, Teaching and Educational Research*, *14*(2), 28–35. https://www.ijlter.org/index.php/ijlter/article/view/525

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Foster, J. L., Harrison, T. L., Hicks, K. L., Draheim, C., Redick, T. S., & Engle, R. W. (2017). Do the effects of working memory training depend on baseline ability level? *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(11), 1677–1689. https://doi.org/10.1037/xlm0000426

Gathercole, S. E., & Alloway, T. P. (2004). Working memory in the classroom. *Dyslexia Review*, *15*, 4–9 https://www.researchgate.net/profile/Tracy-Alloway/publication/254392644_Working_memory_and_classroom_learning/links/0deec539f66116d896000000/Working-memory-and-classroom-learning.pdf

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Golino, M. T. S., Schelini, P. W., & Golino, H. F. (2017). Investigating evidence of content and structural validity in cognitive training tasks for the elderly. *Avaliacao Psicologica*, *16*(3), 278–292. https://doi.org/10.15689/ap.2017.1603.12431

Henry, L. A. (2012). *The development of working memory in children*. SAGE PublicationsSage UK.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70. https://doi.org/10.2307/4615733

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9–F15. https://doi.org/10.1111/j.1467-7687.2009.00848.x

Honoré, N., & Noël, M.-P. (2017). Can working memory training improve preschoolers' numerical abilities? *Journal of Numerical Cognition*, *3*(2), 516–539. https://doi.org/10.5964/jnc.v3i2.54

Ilkowska, M., & Engle, R. W. (2010). Trait and state differences in working memory capacity. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of individual differences in cognition. The springer series on human exceptionality* (pp. 295–320). Springer. https://doi.org/10.1007/978-1-4419-1210-7_18

Irby, S. M., & Floyd, R. G. (2013). Test review: Wechsler abbreviated scale of intelligence, second edition. *Canadian Journal of School Psychology*, *28*(3), 295–299. https://doi.org/10.1177/0829573513493982

Jung, R. E., & Haier, R. J. (2007). The Parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135–154. https://doi.org/10.1017/S0140525X07001185

Karbach, J., & Schubert, T. (2013). Training-induced cognitive and neural plasticity. *Frontiers in Human Neuroscience*, *7*(FEB), 1–2. https://doi.org/10.3389/fnhum.2013.00048

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in Human Neuroscience*, *8*, 242. https://doi.org/10.3389/fnhum.2014.00242

Katz, B., Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2018). The effect of monetary compensation on cognitive training outcomes. *Learning and Motivation*, *63*, 77–90. https://doi.org/10.1016/j.lmot.2017.12.002

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2018). How phonological awareness mediates the relation between working memory and word reading efficiency in children with dyslexia. *Dyslexia*, *24*(2), 156–169. https://doi.org/10.1002/dys.1583

Landsberg, C. R., Astwood, R. S., Van Buskirk, W. L., Townsend, L. N., Steinhauser, N. B., & Mercado, A. D. (2012). Review of adaptive training system techniques. *Military Psychology*, *24*(2), 96–113. https://doi.org/10.1080/08995605.2012.672903

Little, D. R., Lewandowsky, S., & Craig, S. (2014). Working memory capacity and fluid abilities: The more difficult the item, the more more is better. *Frontiers in Psychology*, *5*(MAR), 239. https://doi.org/10.3389/fpsyg.2014.00239

Lopes-Silva, J. B., Moura, R., Júlio-Costa, A., Haase, V. G., & Wood, G. (2014). Phonemic awareness as a pathway to number transcoding. *Frontiers in Psychology*, *5*(JAN), 1–9. https://doi.org/10.3389/fpsyg.2014.00013

Mansur-Alves, M., Flores-Mendoza, C., & Tierra-Criollo, C. J. (2013). Preliminary evidence of effectiveness in cognitive training to improve school children intelligence. *Psicologia: Reflexao e. Critica*, *26*(3), 423–434. https://doi.org/10.1590/s0102-79722013000300001

Mansur-Alves, M., & Saldanha-Silva, R. (2017). Treinar memória de trabalho promove mudanças em inteligência fluida? *Temas Em Psicologia*, *25*(2), 787–807. https://doi.org/10.9788/tp2017.2-19pt

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*(4), 512–534. https://doi.org/10.1177/1745691616635612

Muris, P., Meesters, C., & Van den Berg, F. (2003). The strengths and difficulties questionnaire (SDQ) further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry*, *12*(1), 1–8. https://doi.org/10.1007/s00787-003-0298-2

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Peijnenborgh, J. C. A. W., Hurks, P. M., Aldenkamp, A. P., Vles, J. S. H., & Hendriksen, J. G. M. (2016). Efficacy of working memory training in children and adolescents with learning disabilities: A review study and meta-analysis. *Neuropsychological rehabilitation*, *26*(5–6), 645–672). Routledge. https://doi.org/10.1080/09602011.2015.1026356

Pergher, V., Shalchy, M. A., Pahor, A., Van Hulle, M. M., Jaeggi, S. M., & Seitz, A. R. (2019). Divergent Research Methods Limit Understanding of Working Memory Training. *Journal of Cognitive Enhancement*, *4*(1), 100–120. https://doi.org/10.1007/s41465-019-00134-7

Pergher, V., Shalchy, M. A., Pahor, A., Van Hulle, M. M., Jaeggi, S. M., & Seitz, A. R. (2020). Divergent research methods limit understanding of working memory training. *Journal of Cognitive Enhancement*, *4*(1), 100–120. https://doi.org/10.1007/s41465-019-00134-7

Ramos, D. K., & Melo, H. M. (2016). Jogos digitais e desenvolvimento cognitivo: um estudo com crianças do ensino fundamental. *Neuropsicologia Latino Americana*, *8*(3), 22–32. https://doi.org/10.5579/rnl.2016.0324

Redick, T. S. (2015). Working memory training and interpreting interactions in intelligence interventions. *Intelligence*, *50*, 14–20. https://doi.org/10.1016/j.intell.2015.01.014

Redick, T. S. (2019). The hype cycle of working memory training. *Current Directions in Psychological Science*, *28*(5), 423–429. https://doi.org/10.1177/0963721419848668

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra Psychology*, *5*(1), 18. https://doi.org/10.1525/collabra.203

Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology*, *53*(4), 671–685. https://doi.org/10.1037/dev0000265

Sala, G., & Gobet, F. (2020). Working memory training in typically developing children: A multilevel meta-analysis. *Psychonomic Bulletin & Review*, *27*(3), 423–434. https://doi.org/10.3758/s13423-019-01681-y

Salles, J. F., Fonseca, R. P., Parente, M. A. M., Cruz-Rodrigues, C., Mello, C. B., Barbosa, T., & Miranda, M. C. (2015). Instrumento de Avaliação Neuropsicológica Breve Infantil: NEUPSILIN-Inf - Manual. Vetor.

Sánchez-Pérez, N., Castillo, A., López-López, J. A., Pina, V., Puga, J. L., Campoy, G., González-Salinas, C., & Fuentes, L. J. (2018). Computer-based training in math and working memory improves cognitive skills and academic achievement in primary school children: Behavioral results. *Frontiers in Psychology* Frontiers Media S.A., *8*, 1–12. https://doi.org/10.3389/fpsyg.2017.02327

Saur, A. M., & Loureiro, S. R. (2012). Qualidades psicométricas do Questionário de Capacidades e Dificuldades: revisão da literatura. *Estudos de Psicologia (Campinas)*, *29*(4), 619–629. https://doi.org/10.1590/s0103-166x2012000400016

Snowling, M., & Hulme, C. (1994). The Development of Phonological Skills. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *346* (1315), 21–27. https://doi.org/10.1098/rstb.1994.0124

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics*. 7th ed. Pearson.

Trentini, C. M., Yates, D. B., & Heck, V. S. (2014). *Escala de inteligência Wechsler Abreviada (WASI)*. Manual profissional.

Uehara, E., & Landeira-Fernandez, J. (2010). Um panorama sobre o desenvolvimento da memória de trabalho e seus prejuízos no aprendizado escolar. *Ciências & Cognição*, *15*(2), 31–41. http://www.cienciasecognicao.org/revista/index.php/cec/article/view/375

Vernucci, S., Canet Juric, L., Introzzi, I., & Richard's, M. M. (2020). Working memory training in children: A review of basic methodological criteria. *Psychological Reports*, *123*(3), 605–632. SAGE Publications Inc. https://doi.org/10.1177/0033294119832978

Wagner, R. K., & Torgesen, J. K. (1987). The Nature of Phonological Processing and Its Causal Role in the Acquisition of Reading Skills. *Psychological Bulletin*, *101*(2), 192–212. https://doi.org/10.1037/0033-2909.101.2.192

Yang, J., Peng, J., Zhang, D., Zheng, L., & Mo, L. (2017). Specific effects of working memory training on the reading skills of Chinese children with developmental dyslexia. *PLOS ONE*, *12*(11), 1–20. https://doi.org/10.1371/journal.pone.0186114

Zhang, H., Chang, L., Chen, X., Ma, L., & Zhou, R. (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*, *12*, 154. https://doi.org/10.3389/fnhum.2018.00154

Zhao, X., Wang, Y. X., Liu, D. W., & Zhou, R. L. (2011). Effect of updating training on fluid intelligence in children. *Chinese Science Bulletin*, *56*(21), 2202–2205. https://doi.org/10.1007/s11434-011-4553-5

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Barbosa-Pereira, D., Martins, P. S. R., Ferreira-Junior, L. A., da Costa, M. F. A., Paixão, J. P. C., Costa, R. R., Saldanha-Silva, R., & Mansur-Alves, M. (2022). Is working memory training efficient? Effects on IQ and school performance in Brazilian children. *Applied Cognitive Psychology*, *36*(2), 332–345. https://doi.org/10.1002/acp.3921