Development of a Program based on Goal Management Training for Adolescents with Executive Dysfunctions Complaints

Liana Garcia Nunes 🗅 and Alessandra Gotuzo Seabra 💿

Universidade Presbiteriana Mackenzie (Brazil)

Abstract. This study aimed to develop a program based on Goal Management Training (GMT) and to investigate its effectiveness on executive functions, through formal instruments and an ecological task. Participants were 25 adolescents with complaints of executive dysfunctions. They underwent neuropsychological assessment of working memory, inhibitory control, cognitive flexibility, planning, and intellectual ability. Participants also took part in a cooking activity and were evaluated for errors per action, of omission, activity performance time, recipe consultation. After, they were randomly allocated to an active control group (CG), which underwent psychoeducation sessions, and an experimental group (EG), stimulated through GMT in eight sessions. Then participants underwent another assessment and follow-up after 4 weeks. In post-intervention analyses, results showed an improvement in executive functions in EG, in the working memory measurement and time of the ecological activity (g = 1.78 and .93, respectively), IQ (g = -1.01), reasoning (g = -.89), flexibility (g = -2.95), inhibition (g = -5.78) and execution time of the ecological activity (g = .98). Significant interactions between assessment Time x Group revealed EG gains in IQ, scores in reasoning and flexibility. EG also had longer execution time in flexibility and inhibition tests. That is, EG had greater scores and probably was less impulsive in these tests. Furthermore, EG decreased the number of verifications and the time in the ecological task, that is, had a more efficient performance. Results suggest the intervention can be as instrument to promote executive function.

Received 20 November 2019; Revised 26 January 2021; Accepted 12 February 2021

Keywords: adolescents, cognitive intervention, executive functions

Executive functions (EFs) are a set of cognitive processes that refer to executive control, that is, the ability to set goals, make decisions, define strategies, and direct and regulate behavior to achieve specific goals (Diamond, 2013; Friedman & Miyake, 2017). This definition has been widely accepted in the last decade, despite there are different definitions of EFs looking for operationalize and clarify this term any better (e.g., Lezak et al., 2012; Zelazo et al., 2003).

In this paper, Miyake et al. model (2000) and Diamond proposal (2013) will be used as a background for the results interpretation. According to these models, three core skills make up EFs: Working memory, which

Conflicts of Interest: None.

involves the ability to maintain and manipulate information; inhibition or inhibitory control, important for ignoring distracting elements and suppressing attention or behavior directed at something irrelevant; and cognitive flexibility, which involves the ability to thinking creatively in order to achieve a goal (Diamond, 2013; Friedman & Miyake, 2017). According to the theoretical model of Diamond (2013), these three core skills are related to higher-order components of EFs, such as reasoning, problem solving, and planning. This set of processes that make up EFs relate to the ability to prioritize, self-monitor, and organize actions (Friedman & Miyake, 2017).

Other models have been proposed and there are still some aspects to be better understood, such as the definition and the role of inhibition, and other possible skills, as fluency and dual-task coordination

How to cite this article:

Nunes, L. G., & Seabra, A. G. (2021). Development of a program based on goal management training for adolescents with executive dysfunctions complaints. *The Spanish Journal of Psychology*, 24. e20. Doi:10.1017/SJP.2021.17

Correspondence concerning this article should be addressed to Liana Garcia Nunes. Rua Paranaíba, 1165, Centro. 79603–090 Três Lagoas-MS (Brazil).

E-mail: eu_tovar@yahoo.com.br

Funding Statement: This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq Parte superior do formulário (AGS, grant number 310588/2017–0); the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (LGN, grant number 88887154311/2017–00); and the MackPesquisa (AGS, grant number M0001–0000108479).

(Friedman & Miyake, 2017). However, despite the Miyake et al. model (2000) is not comprehensive (because there are other EF components), or it is not even a proposition about elementary processes (since these core skills were chosen to examine unity and diversity of EFs) (Friedman & Miyake, 2017), that models with the three core skills, as well as the complex skills, have been widely accepted in the research in the area (Zelazo et al., 2017) and can facilitate the operationalization of assessment and intervention.

Alterations or underdevelopment of EFs can lead to difficulties in self-regulation of various behaviors in various age groups. In adulthood, they are related to work productivity impairment, alcohol abuse, drug abuse, and motor vehicle collisions; in children, they may cause reading and writing difficulties, learning difficulties in mathematics, and disruptive behaviors (Clements & Sarama, 2019; Meixner et al., 2019; Pope et al., 2016).

Given the relevance of EFs, studies have been conducted with promising results seeking to stimulate such functions, especially in children and adolescents. As examples we could mention studies that sought to verify the cognitive stimulation through dance classes, karate, martial arts, tennis lessons, meditation, digital games, football training program, educational robotics (Harwood et al., 2017; Homer et al., 2018; Kiani et al., 2017; Kullina et al., 2018).

One of the cognitive techniques for rehabilitation of EFs is the Goal Management Training approach (GMT), which was developed by Robertson in 1996 (Levine et al., 2000; Robertson, 1996) from theories about the functions of the prefrontal cortex based on Duncan's goal neglect theory (1986). The GMT approach follows a multi-module protocol in which strategies are trained while a cognitive activity is performed. Its main purpose is to train the individual to stop his or her automatic behavior in order to set goals and to hierarchically plan the steps to achieve them.

Thus, GMT studies follow a protocol of predefined stages with specific topics to be worked on in each session. The stages consist of: (a) Stop and ask: What am I doing?; (b) check the mental blackboard; (c) define the main task or goal to be met; (d) list the steps or stages required to meet a goal; (e) learn the steps; (f) do the steps; (g) check if you did what you planned (Robertson, 1996).

In addition, some studies were analyzed to development this protocol based on GMT (Grant et al., 2012; Levine et al., 2000; Tornås et al., 2016). Because it is a sample of adolescents, the use of metaphor "check the mental blackbord" to the training of working memory was changed to "check the mental tablet" and the presentations were done by Prezi program for give more movement and improve the interesting of the young sample.

The technique has also been applied to individuals with other conditions or age groups than adults with brain injury. A case study with a patient with schizophrenia sought to ecologically verify the benefits of GMT in meal preparation, laundry, and meeting preparation, noting significant improvements in the questionnaires answered by the caregiver (Levaux et al., 2012). Another recent study was conducted with three young college students without previous diagnosis, and revealed improvements after the intervention in their ability to set goals and plan the steps to achieve them (Rivera et al., 2019). Another study reported the case of an 11-year-old teenager with ADHD symptoms and complaints of executive dysfunctions in the social environment. The results suggested benefits of GMT in attention and executive functions, but when combined with another type of intervention (Nicoló, 2015). There are also published studies with young university students with complaints of executive dysfunction in their social environment (Carstens, 2016) and children with traumatic brain injury (Krasny-Pacini et al., 2014).

Despite advances in the use of the GMT technique, only a still few studies with adolescents are reported, especially without specific diagnoses of neurodevelopmental injuries or disorders. It is essential to think of such a population that, despite the lack of diagnosis, may present with subclinical difficulties in executive functions. Since such functions, as previously described, predict physical health, substance dependence, personal finances, and criminal offending outcomes, their development is essential to the health, wealth, and public safety of a population (Moffitt et al., 2010). Thus, executive function difficulties in adolescence can have important lifetime consequences (Moffitt et al., 2010).

Thus, based on the literature suggesting gains with the application of the GMT technique and highlighting the importance of the development of executive functions in adolescence, the aims of this study were to develop a program based on GMT training and to investigate its effectiveness among adolescents with executive function deficits, through formal instruments and an ecological task. We expected greater gains in the experimental group, submitted to GMT, on formal EFs tests, as well as on an ecological task, in comparison to the active control group.

Method

Participants

Participants were recruited from private and public schools in Três Lagoas, state of Mato Grosso do Sul, Brazil. We contacted 7 coordinators and 35 teachers from 5 different schools to talk about all their 700 students. Students with executive functioning issues according to parents' and teachers' report were selected, as described below.

Participant selection criteria. The inclusion criteria were students between 11 and 17 years, from private and public schools in Três Lagoas, Brazil, with difficulties in executive functions, according to both teachers' and parents' reports, as explained above.

We thoroughly explained the concept of executive functions to the 7 coordinators and 35 teachers in an initial 2-hour meeting. They were given descriptions of EF based on examples that could be observed in the classroom environment.

Then teachers and coordinators were asked - based on those descriptions and coexistence in school context to indicate all students between 11 and 17 years with executive functioning complaints (including working memory, inhibitory control, flexibility, and planning issues). Of the 700 students, 102 were indicated. Then students' teachers and parents responded to the Executive Function, Regulation, and Delay Aversion Inventory (Inventário de Funções Executivas, Regulação e Aversão ao adiamento - IFERA), described in the Assessment Tools section. In the following stage of the study, only those participants who showed IFERA difficulty indices above the expected average were included, according to both teachers' and parents' reports. A total of 53 participants had scores higher than 80 for both the parent and teacher scales (the maximum scale score is 140 and the average score in Brazilian samples is 77.3 points, according to Trevisan, 2014). The 53 participating students corresponded to 7.5% of the initial 700 students and 51% of the 102 students initially indicated by teachers.

Exclusion criteria. The 53 students with scores indicating difficulties above the expected average in IFERA, according to both teachers' and parents' reports, underwent a neuropsychological assessment to analyze the other exclusion criteria. Participants who had previous psychotic symptoms (1 participant), IQ below 70 (26 participants), or any type of brain injury as reported by the school or their guardians (1 participant) were excluded. This selection resulted in a final sample of 25 participants (corresponding to approximately 3.6% of the initial 700 students).

Thus, IQ below 70 was used as exclusion criteria to avoid that participants with difficult in comprehension and realization of activities participated of research. For that, we used the abbreviated measure of IQ, with subtests of vocabulary and matrix reasoning of WASI, as explained in Instruments. In addition, *T*-scores for vocabulary and matrix reasoning subtests were used in analyses of effect of this intervention to check if the GMT would produce improve in these variables.

We established no prior criteria regarding the origin of the difficulties, so that participants with a known or unknown etiology of their difficulties were considered for the study. The subjects were not submitted to any other neuropsychological intervention program before or during our intervention. Four participants had known etiologies, namely: Attention deficit hyperactivity disorder (2 participants), high-functioning autism (1 participant), and oppositional defiant disorder (1 participant), all with medical and neuropsychological diagnoses. Although these participants reported using medication for at least 2 years, the complaints persisted. The remaining 21 participants reported unknown etiology and no medication use to minimize the disorder.

Study Design

The 25 participants were first divided into 2 groups according to age group: G1 (11–14) and G2 (15–17). G1 was then randomly divided (by drawing names) into an experimental group (EG1), which was submitted to GMT intervention, and a control group (CG1) that underwent psychoeducation intervention. G2 was also divided via random draw into experimental (EG2) and control (CG2). Two participants in group EG1 had a diagnosis of attention deficit disorder, 1 participant in group EG2 had a diagnosis of autism spectrum disorder and 1 participant in CG1 had a diagnosis of oppositional defiant disorder. It is noteworthy that parents and teachers were unaware of the group to which the child was allocated (EG or CG). Figure 1 shows a flowchart with the research group division.

Assessment Tools

For IQ assessment, we used the Brazilian Portuguese version of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) using *T*-scores for the variables vocabulary, matrix reasoning, and full scale IQ (test, evidence of validity and reliability in the Brazilian manual by Trentini et al., 2014). Such measures were used as exclusion criteria and also to measure possible changes throughout the three measures. To assess executive functions, we used neuropsychological tests, parent/teacher reports, and an ecological task, as described below.

For the selection of participants, we also applied the Executive Function, Regulation, and Delay Aversion Inventory (*Inventário de Funções Executivas, Regulação e Aversão ao adiamento* – IFERA, Trevisan, 2014) to parents and teachers to assess their perceptions of students' executive functioning in everyday and classroom environments. This inventory aims to assess executive functioning in an ecological manner and includes items that assess state regulation and delay aversion (some examples include "When you are in the middle of an activity, you often forget what you were doing" or "Do things without thinking first about what might happen"). Answers are on a 5-point likert scale ranging from *never*

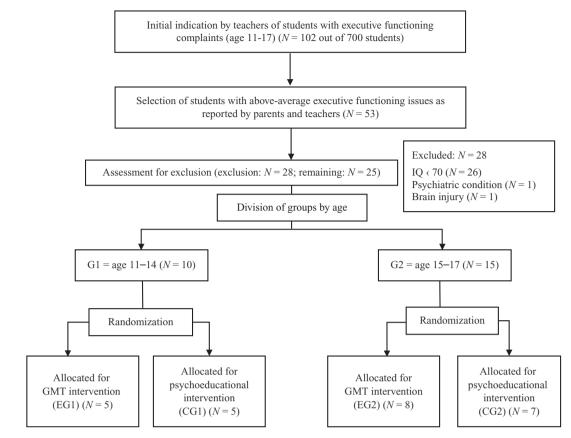


Figure 1. Flowchart with Research Group Division

(1) to always (5) where responses with higher scores indicate greater reported difficulties. The measure used was the total raw score of the inventory. As problems with executive functioning are not expected to decrease with age, since the comparison is made with the expected average for age, the use of raw scores is justified despite the participants' age difference. There are evidences of validity and reliability for Brazilian population (Trevisan, 2014; Trevisan, et al., in press). The inventory has evidences of content validity, as assessed by experts with experience in clinical neuropsychology and psychometrics; evidences for confirmatory factor analysis, with acceptable five factors model; evidences for relation to other instruments (SNAP-IV and CHEXI); discrimination between children with and without Attentional-deficit/hyperactivity disorder; and good reliability indexes (Trevisan, 2014; Trevisan et al., in press).

Among the neuropsychological tests, we used the Five Digit Test – FDT to assess participants' cognitive flexibility and inhibitory control (test, evidence of validity and reliability in Sedó et al., 2015). The test has evidences of validity, reliability and normative scores for Brazilian population and presents a numerical Stroop paradigm in four parts. Parts 1 and 2 assess automatic processes: The examinee is supposed to read Arabic numerals in a screen (1, 2, 3, 4, 5) and count asterisks (1-5 asterisks). In Part 3, the examinee should identify the amount of numbers (instead of reading the numbers). In Part 4, the task changes between the rules given on Part 1 and Part 3, depending on an explicit clue. There are two executive function measures: Inhibition (Part 3 minus Part 1) and flexibility (Part 4 minus Part 2), calculated in terms of score and time to task completion. Percentile was used for score measurements and, for time measurement, we converted the results to *t* score (Strauss et al., 2006).

To assess working memory we used the WISC–IV Digits subtest only in the backward digit sequence (test, Brazilian standards, evidence of validity and reliability in Wechsler, 2013). In this test, the researcher read a sequence of numbers and the participant had to verbally reproduce these digits indirectly. Results were converted to *t*-score.

The Tower of London test – ToL (Brazilian standards, evidence of validity and reliability in Menezes et al., 2012) assessed participants' planning ability. The test consists of a board with three vertical rods on which four different colored beads of the same size are placed. The task is to move the colored beads to reproduce the position shown in a stimulus book in as few moves as possible. Planning skill is considered a complex EF

because the individual needs to use several basic EFs to solve the task, such as setting the goal and the steps to achieve it, staying active during task resolution – inhibiting distracting stimuli – and using cognitive flexibility when each card presented requires a new way of thinking about the goal. To measure the results of the Tower of London test we used the standard score according to the Brazilian norms (with a mean of 100 and a standard deviation of 15).

The ecological task to assess executive functioning consisted of a cooking activity, which was performed in a kitchen in the clinic where the research was conducted. The students were supposed to make a cookie (no baking required, only the proper mixing of ingredients). To do so, they were given a recipe without the "ingredients" section. Therefore, the ingredients were not described separately, but throughout the recipe execution steps. During the activity, all the necessary ingredients were available on the counter where students worked, as well as other unnecessary ingredients that were added as distractors. The participant was instructed to follow the recipe and make it as quickly as possible. The following variables were measured: total execution time, total errors (errors per action, such as using ingredients not needed or in a larger quantity; and errors of omission, such as not adding required ingredients to the recipe or using them in smaller quantities), number of actions (number of total actions taken by the participant from the beginning to end), and recipe verification (number of times the participant reviews the recipe). There are no normative data that quantitatively measure the gains of this intervention through this activity. For this reason, we analyzed changes in the number of times participants consulted the recipe (action considered indicative of working memory use), number of errors (considered indicative of inhibition), and time (management of actions, which was an analysis used in a previous study conducted by Levaux et al., 2012).

Intervention Tools

GMT technique. The experimental groups (EG1 and EG2) underwent 8 intervention sessions. The development of the activities was based on GMT principles (Levine et al., 2000). For the animations and visual examples we used PREZI computer presentations. The first session consisted of general information about the training and identification of situations that place high demands on EF in daily tasks. In the second session, participants worked on the concept of autopilot. The concept refers to behaviors that emerge naturally, automatically, without conscious control over them. In the third session, we explored the concept of working memory through examples of daily life using the "mental tablet" metaphor. In the fourth

session, participants were instructed to set goals, keep them in their working memory (mental tablet) and resort to them whenever necessary. In the fifth session, we addressed the issue of decision making, which is necessary when we have two competing goals that have the same weight of obligation. In the sixth and seventh sessions, we explored complex goals, meaning in these sessions participants were trained to divide a complex task into several smaller tasks to cope with complex goals. Finally, in the eighth session, we proposed a complex ecological activity, with the objective of making participants apply the strategies learned in an Internet search in order to get a response for the task. The task consisted of a fictitious situation in which participants simulated purchases for their relatives with a certain amount of play money and with the help of the researcher. Internet searches were based on tips on what family members needed and liked. The tasks involved: 1) Finding a gift for a newly married older sister who loves blue shoes as well as cooking; 2) finding a gift for a brother who is very fond of chocolates but has started a diet and needs to practice sports to lose weight; 3) finding a gift to the twin brothers who have just come out of diapers and will go to their first day of kindergarten; 4) finding something for mom who likes photography and reading, but currently has no space to place more photos at home; 5) finding something for father who is a truck driver - therefore, who is almost never home - and loves his family.

In addition to group activities, participants were given homework assignments in each section. They were to report everyday situations in which session techniques were used. Such reports should be sent to the researcher's cell phone via text messages, videos, or voice messages. Table 1 presents a summary of the sessions.

Psychoeducation. Both control groups (CG1 and CG2) underwent a psychoeducational intervention that corresponded to the time of administration and the collective application of the GMT technique. All participants had access to computer-presented educational materials and group discussions on some age-appropriate topics. Thus, topics for discussions about everyday issues by adolescents were addressed, such as bullying, self-awareness, vandalism, inclusive education, technology overwhelm, gender equality, and teenage pregnancy. Similar to EGs, homework was assigned after each session. Participants were to report on the subjects that were covered in the sessions and send them to the researcher's cell phone via text messages, videos, or voice messages. Table 1 presents a summary of the sessions.

Procedure

After approval of the project by the Research Ethics Committee of our University, and after authorization of responsible and consent of the adolescents using the

Table 1. Plan of the Sessions for GMT and Control Groups

Session			GMT Group		Control Group				
	Subject	Goal	Activities	Homework	Subject	Goal	Activities	Homework	
1	Introduction/ risk's situations	Recognize high- risk situations	Card categorization	Report high-risk situations on a daily basis	Introduction/ Bulling	Reflection on bullying and maximize the capacity for empathy	Presentation of the movie "bullying"	Film your daily empathy practices	
2	Automatic pilot	Be able to stop routine attitudes	The master ordered / categorization	Report situations where and when you practiced stopping the autopilot	Self knowledge	Reflection about self image	Movie: "the mask you live in"/ pencil-paper	Film inclusion behaviors in the social environment	
3	Mental tablet	Frequently check working memory	Categorization of letters and text activities with restrictions	Report situations where and when you practiced a mental tablet	Vandalism	Identify vandalism behaviors	Movie/ pencil- paper	Filming places on the street that have been vandalized	
4	Set a goal	Set a goal	Complex activities	Report situations using the strategies taught	Inclusive education	Maximize empathy about inequality	Experience of limitations / movie: "the fundamental caring"	Film support practice or support inclusion	
5	Make decisions	Be able to make decisions in the face of two options	Find the missing part of the problem picture / situation	Filming everyday situations using everyday decision-making skills	Excess of technology	Identify excess technology daily	Identify street games	Shoot situations when they are having fun away from the computer	
6	Divide work	Be able to divide a complex task into subtasks	Subtask activities	Film everyday situations using the ability to split complex tasks into smaller ones	Gender equality	Recognize gender equality	Internet search about gender	Film your daily actions related to gender tendencies	
7	Check if reach your Goal	Be able to check that you are on the right track	Subtask activity	Filming everyday situations using the learned strategies	Teenage pregnancy	Identify the difficulties and responsibility of pregnancy	Search activity on questions about sexuality	Film the care with a native tree seedling	
8	Final activity	Use all strategies in a final activity	Ecological activity	No activity between sessions	Final activity	Be able to apply all the issues discussed in an activity	Country creation activity	No activity between sessions	

informed consent forms, the baseline assessment was initially performed. Then, the G1E and G2E groups received GMT intervention in 8 sessions, 60-minute each, which were held in three weeks. Concomitantly, groups G1C and G2C underwent interventions of 8 psychoeducation sessions, 60-minute each. After the intervention period, all participants were reevaluated with the same measures used in the pre-intervention. After a period of 4 weeks of the intervention period, all participants went through the same follow-up process again. The evaluations and interventions were performed in a private psychological clinic in the city of Três Lagoas - MS.

Data Analysis

In the study design, there were a between-subject variable (group: EG and CG, with grouping by age) and a within-subject variable (assessment time: Preintervention, post-intervention, and follow-up). Our dependent variable were the measures in the tasks, as described in Instruments section.

Considering the small sample size (n = 25) and the arbitrary distribution of the data, we conducted a series of nonparametric mixed analysis of variance (ANOVA), based on rank estimates. This approach enables to compare between- and within-groups temporal effects, and their interaction even with small samples sizes or presence of outliers, due to its rank-based robust statistics (Noguchi et al., 2012). The ANOVAs were implemented through nparLD (nonparametric longitudinal designs; Noguchi et al., 2012) package, at the R statistical environment, which allow Fisher-liked hypothesis testing with an adjusted estimator for model's degrees of freedom. In this analysis, we focused on the interaction between group and assessment time in order to verify if the EG would have a different progress than the CG.

Additionally, differences for the different measures were calculated, considering Difference 1 (post-test result minus pretest result) and Difference 2 (follow-up result minus pretest result). Due to the small number of subjects in the analyzes, we used Hegdes' *g* to appropriate measure the effect size of the gain scores in order to compare the differences in EG with the differences in CG. Effect sizes were interpreted as follows up *g* from .20 = no effect; *g* from .20 to .30 = small effect; *g* from .40 to .70 = medium effect; *g* greater than or equal to .80 = large effect (Kline, 2004). It is noteworthy that, in our analysis, negative g values were indicative that the difference was greater – in terms of raw scores – for EG than for CG.

Results

Table 2 summarizes the results obtained, with the values of Differences 1 and 2 for the control (C) and

experimental (E) groups, and the effect size (*g*) of the comparison of such gain scores between the groups. The table also shows the results of the Assessment Time x Group Interaction obtained with mixed ANOVA (*F*-value, *p*-value).

Table 3 summarizes the results obtained.

ANOVA results revealed significant effects of the intervention in terms of interaction between time and group in five of the 14 measures assessed. Analyzing the effect size of the differences, we observed medium or large effect sizes in several measures, namely, seven measures for Difference 1 and nine measures for Difference 2.

Regarding WASI scores, we observed a significant time-group interaction for the estimated IQ. The EG showed greater gains than the CG soon after the intervention and also at follow-up, with a large effect size for Difference 1 and medium effect size for Difference 2, according to Hedges' g. These gains were leveraged especially by Matrix Reasoning scores in Difference 1, where the EG had a higher gain than the CG (3.85 versus .50, respectively), and by Vocabulary, in difference 2, where the EG had higher gain than the CG (.62 versus -.67, respectively). For the estimated IQ, the EG gain in Difference 2 was lower than in Difference 1, suggesting that changes soon after the intervention decreased over time. There was significant Time x Group for IQ and Matrix Reasoning, with higher gains for EG in both measures.

In the measurement of working memory, as assessed by the WISC scale in the backward digit sequence, we observed no significant interaction between assessment time and group. However, there was an average effect size in favor of the EG in the first difference, suggesting that, right after the intervention, the EG showed higher gains than the CG. On the other hand, the opposite occurred in the second difference, that is, the CG revealed higher gains. Thus, there seems to have been no significant long-term differences in this measure after the intervention.

The FDT revealed important results: Out of the eight measures, four had large effect sizes and three had medium effect sizes. There was also significant Time x Group interactions for time measures in flexibility and inhibition, and for score in flexibility. In terms of means, EG showed greater gains in flexibility (Difference 1) and inhibition scores (Differences 1 and 2), and CG exhibited greater gain in flexibility score (Difference 2). Still in relation to the means, the EG showed greater time gain (that is, participants took longer to respond) in all FDT measurements. It is noteworthy that, regarding time, there was significant interaction between time and group. That is, the EG started to exhibit a longer execution time, which was accompanied by more accurate performances, especially in flexibility.

Table 2. Results obtained, with the Values of Differences 1 (Post-test – Pre-test) and 2 (Follow-up – Pretest) and 2 for the Control (C) and Experimental (E) Groups, Effect Size (g) of the Comparison of the Gain Scores, and Results of the Time x Group Interaction after Mixed Non-Parametric ANOVA

		Gain 1		Gain 2 (Follow up – pre test)				
	(Pos	t test – pre test))				Interaction Time x Group	
Group	С	Е	g	С	Е	g	F	р
Vocabular	y - T-score							
М	25	.54	33	67	.62	5	.72	.435
(SD)	(1.53)	(2.82)		(2.01)	(2.79)			
Matrix reas	soning - T-score							
М	.50\(2.02)	3.85	89	1.00	1.62	13	3.08	.05
(SD)		(2.71)		(2.71)	(5.54)			
IQ		. ,						
М	.08	3.69	-1.01	.08	1.92	57	4.28	.017
(SD)	(.76)	(4.71)		(.86)	(4.21)			
Indirect di	gits - T-score							
M	1.08	1.62	22	.67	46	.30	.91	.388
(SD)	(1.93)	(2.68)		(1.55)	(4.67)			
· · ·	bility (Time) - T-	()						
М	33	5.23	-1.21	58	9.00	-2.95	26.48	< .001
(SD)	(1.37)	(5.96)		(1.44)	(4.11)			
· · ·	bility (Score) - Pe	()			(
М	1.67	14.62	66	3.75	-1.92	.72	3.60	.048
(SD)	(5.53)	(27.42)		(8.45)	(6.66)			
()	ition (Time) - T-	()		(0.20)	(0.00)			
М	.17	10.77	-3.11	.67	15.92	-5.78	78.82	< .001
(SD)	(2.27)	(4.12)	0.111	(2.53)	(2.59)	0110		1001
· · ·	vition (Score) – P	· · ·		(2.00)	(=:::))			
M	.00	3.46	24	.00	3.08	57	.61	.509
(SD)	(.00)	(18.85)	.21	(.00)	(7.22)	.07	.01	.507
. ,	ondon – T-score	()		(.00)	(7.22)			
M	-1.83	.31	15	-6.08	69	43	.69	.481
(SD)	13.97	(12.01)	.15	(12.23)	(12.02)	.10	.09	.101
()	logical task (Acti	· · · ·		(12.20)	(12.02)			
M	0	.02	02	5	31	21	.60	.542
(SD)	(.58)	(1.15)	02	(.87)	(.82)	— . ∠⊥	.00	.042
· · ·	logical task (Om	. ,		(.07)	(.02)			
M	–.25	23	03	25	38	.18	.61	.535
(SD)	23	23 (.58)	05	23	38 (.49)	.10	.01	.555
· · ·	(.60) logical task (Tim	· · /		(.03)	(.±2)			
M	10.25	-74.69	.93	10.33	-73.92	.98	6.33	.002
	(110.25)		.93			.90	0.33	.002
(SD)	· · · ·	(59.23)	\ \	(78.79)	(85.52)			
-	logical task (Rec	1		1.05	0.01	45	10.62	< 0.01
M (SD)	.33	-2.38	1.78	-1.25	-2.31	.45	10.63	< .001
(SD)	(1.70)	(1.21)		(2.74)	(1.64)			

Note. The cells in bold refer to significant values

The Tower of London, which measured planning ability, revealed no significant interaction between assessment time and group, but showed an average effect size in the Hedges' analysis on Difference 2. The effect size had been insignificant for the Tower of London shortly after the intervention, which may suggest that participants may have continued to use planning strategies and, consequently, the effect size increased over time, detected at follow-up.

In the cooking activity, which aimed to verify executive functions in a real-life situation, we observed a significant interaction between assessment time and group for time measure in the execution of the recipe and number of times participants referred to the recipe.

Construct	Assessment	Interaction time x group (p < = .05)	Difference	No Effect (≤ .2)	Small Effect (.2, .3)	Medium Effect (.4, .7)	Large Effect (≥ .8)
Neuropsychological assessmet	44	•					
Vocabulary	WASI Vocabulary		1		33		
Vocabulary	Which vocubulary		2		.00	5	
Reasoning	WASI Matrix Reasoning	.053	1			10	≥.8
8	0		2	≤ .2			
Inteligence	WASI IQ	.017	1				≥.8
C			2			57	
Operational memory	WISC Backward Digits		1		22		
			2		30		
Cognitive flexibility	FDT Flexibility Time	< .001	1				≥.8
			2				≥.8
	FDT Flexibility Score	.048	1			66	
			2			.72	
Inhibition	FDT Inhibition Time	< .001	1				≥.8
			2				≥.8
	FDT Inhibition Score		1		24		
			2	-		57	
Planing	Tower of London		1	≤ .2		10	
			2			43	
Ecological task							
Inhibition (action)	Recipe task - Action errors		1	≤ .2			
	1		2		21		
Inhibition (omission)	Recipe task - Omission errors		1	≤ .2			
			2	≤ .2			
Action management time	Recipe task -Time	.002	1				≥.8
			2				≥.8
Operational memory	Recipe task -Recipe verification	< .001	1				≥.8
			2			.45	

Table 3. Summary of Non-parametric Anova Results (Interaction between Time and Group) and Size Effects for Differences 1 (Post-test – Pre-test) and 2 (Follow-up – Pretest)

Note. Gray cells in effect sizes indicate measures in which CG gains were greater than EG gains

9

We observed that the EG decreased the number of recipe verification, which reduced the total time to perform the activity, that is, the performance has become more efficient. Moreover, regarding Hedges' analysis, the effect sizes were insignificant or small for errors (per action or omission), but with a medium to large effect size for the other measures. That is, EG participants were more effective in memorizing the recipe before executing it, which reduced execution time. However, we observed no differences in CG in terms of behavior inhibition, measured by errors per action and of omission.

Discussion

The aim of the present study was to evaluate the effectiveness of Goal Management Training in adolescents with executive functioning issues through formal instruments and an ecological task. The results revealed gains for the experimental group in the IQ, flexibility, inhibition, and ecological task tests, regarding significant assessment Time x Group Interaction. Moreover, measurement of all tasks revealed medium or large effect sizes.

Thus, initially, after the intervention (Difference 1), by analyzing medium or large effect sizes, EG participants had greater gains in matrix reasoning skills, estimated IQ, longer time spent on flexibility and inhibition on FDT, higher flexibility scores on FDT, shorter performance time, and fewer verification at the recipe task. The analysis of Difference 2 revealed the maintenance of the gains after the intervention was completed. Considering medium or large effect sizes, the EG had greater gains in vocabulary, estimated IQ, longer time spent on flexibility and inhibition on FDT, higher inhibition scores on FDT, higher Tower of London score, shorter performance time, and fewer recipe consultations. It is noteworthy that only one measure revealed a relevant effect size in favor of the control group, namely, the flexibility score on FDT at Difference 2. In addition, there was significant interaction Time x Group for estimate IQ, reasoning, flexibility (time and score), inhibitory control (time), recipe verification and time in ecological task.

Therefore, overall, the results suggest that the 8-session GMT training was able to promote greater gains in these skills than the psychoeducation performed with the control group. Our findings were consistent with previous studies that reported greater benefits of inhibitory control, sustained attention (Novakovic-Agopian et al., 2018), planning (Tornås et al., 2016), and time (measured by the Tower Test; Stubberud et al., 2014).

We observed important changes in various inhibition skills. Such changes probably occurred because

inhibitory control ability is already trained in early GMT sessions, when the concept of "turning off the autopilot" in cognitive activities and day-to-day tasks is presented to participants. In the present study, it is noteworthy that higher EG gains in inhibition score measures were accompanied by longer execution time in the activity, which may suggest that participants began to think more critically in the face of unusual situations that required voluntary responses. That is, in this case, since the tasks were difficult and not automated, the increase in time can be considered a gain based on the intervention procedure used, which has as one of the key principles to turn off the autopilot (Levine et al., 2000). Several other studies also showed improvements in participants' inhibition (Cuberos-Urbano et al., 2018; Levaux et al., 2012).

Regarding cognitive flexibility, the present study revealed greater score gains in the EG only for Difference 1, while the CG revealed greater gains in Difference 2. Despite CG had greater score gains than EG in one of the differences (flexibility score at Difference 2), there was significant interaction between moment and group for flexibility scores and for flexibility time. That is, the score gain in the EG for Difference 1 was accompanied by longer execution time in both differences, which may suggest that participants began to think more critically in situations that required cognitive flexibility, as occurred in relation to inhibition. Gains in flexibility were found in some previous studies, but with some inconsistence. For example, although Cuberos-Urbano et al. (2018) showed benefits of GMT in the flexibility of its participants, some methodological limitations were observed in that study, such as the very small sample size of only 15 participants and the non-use of a control group. On the other hand, although the study performed by Tornås et al. (2016) had a larger sample, it showed no improvement in flexibility skills. We hypothesize that, although the core of GMT is in fact inhibition, because this ability is fostered since the beginning of the intervention and because it is present in all GMT sessions, the flexibility was essential in our procedure to divide a complex task into smaller tasks is worked out, supporting the development of higher mental functions related to planning and problem-solving.

About working memory, we observed no significant differences in the results between the groups in the traditional digit test, in which there was no significant interaction between assessment time and group, with only small effect size in favor of the EG in Difference 1 and in favor of the CG in Difference 2. However, we observed a gain and significative effect in working memory proposed scores for the ecological activity in relation to the number of times participants needed refer to the recipe for its execution. Thus, the EG revealed a greater reduction than the CG in the number of recipe verification, both in Difference 1 (with a large effect size) and Difference 2 (with a medium effect size) and significant interactions between Time x Group. Thus, we can hypothesize that the techniques stimulated in the intervention, such as the use of the mental tablet, may have favored such performance in the EG compared to the CG.

The planning activity, assessed by the Tower of London test, is characterized as complex EF as it requires the joint use of core EFs (Diamond, 2013). In this task, although we observed no significant interaction between assessment time and group, and the effect size was non-significant in Difference 1, there was an average effect size in Difference 2 in favor of the EG. That is, EG participants had greater gain scores than the CG in the Tower of London test from pretest to follow-up. It is noteworthy that, according to Diamond's (2013) executive function theory, the so-called core executive functions - which include inhibition, working memory, and cognitive flexibility - when working together, enable the execution of complex executive functions, such as planning. Based on the results presented here, we can hypothesize that the change in effect size (from non-significant in Difference 1 to medium in Difference 2) is related to the complexity of the task, that is, it is possible that, because it is a more complex skill, planning is developed after the most basic skills.

It is important to note that the benefits generated by the more traditional measures (inhibition and flexibility) were not verified in the recipe ecological task. Thus, our findings are in line with the research by Carstens (2016), who also sought to verify the use of the technique in young individuals only with complaints of executive dysfunction. It may be that in non-chronic individuals, behavioral changes may be milder and barely noticeable.

Besides the gains evidenced in relation to executive function measurements, we also highlight a significant interaction between assessment time and group for the estimated IQ. Moreover, the EG showed significantly greater gains than the CG. This result is interesting because, a priori, it is not expected that the intervention in EF significantly changes IQ. However, we highlight the important relationship between EF and intelligence that can be outlined, for example, from Primi's (2002) research. According to that study, intelligence is related to the ability to find a goal and subsequently inhibit irrelevant stimuli. Thus, the ability to inhibit, which is fostered in GMT, would play an important role in the resolution of intelligence tests. Additionally, it is noteworthy that previous studies have also suggested the relationship between fluid intelligence and working memory, and that working memory stimulation, one of the components of EF, is associated with increased performance in fluid intelligence tests (Au et al., 2016). Thus, stimulating the basic components of EF can have an important effect on intelligence test performances.

In summary, the main results showed evidence of the effectiveness of GMT treatment for adolescents with complaints of executive dysfunction according to parents' and teachers' reports. However, limitations of this study should be highlighted. Firstly, our study sample was small, which generates some sample bias or may contribute to the absence of significant effects on ANOVA, although mean trends have been identified. The second limitation is that although the study was randomized in a second stage, it was not a double-blind study in its assessment and intervention phases. A third limitation is the short follow-up time for the study, as the 4-week period may not be sufficient to predict the prolonged gains of the intervention. Additionally, two participants in this research, who were randomly included in the experimental group, were using methylphenidate throughout the intervention process. This is an important variable in the stimulation process, as the gains from medication use could be enhanced with the GMT stimulation.

Thus, future studies could include more sessions and with longer duration, a larger number of participants, training tasks closer to activities of daily living that include executive functions, and a longer follow-up. It is noteworthy that, despite the differences in cognitive abilities measured here - related to working memory, inhibitory control, planning, and flexibility - there is no predictive benefit in the ability to make decisions in adolescents, as the socio-emotional aspects related to peer influence, emotional control, and the reward process were not the focus of this study. Since decision making has been treated as the product of cognitive and socio-emotional aspects, and its deficits associated with risky behavioral patterns in adolescence - such as high mortality rates, toxic abuse, and mental disorders interventions that focus both on cognitive and socioemotional skills could be beneficial to this population (Defoe et al., 2015; Moffitt et al., 2010; Zelazo & Carlson, 2012).

Despite the limitations, the study allowed the preliminary development and investigation of an intervention proposal that aims to improve working memory, inhibitory control, flexibility, and planning skills in adolescents. The results, albeit modest, seem to indicate a promising path for further research in this area.

References

Au, J., Buschkuehl, M., Duncan, G. J., & Jaeggi, S. M. (2016). There is no convincing evidence that working memory training is not effective: A reply to Melby-Lervåg and Hulme (2015). *Psychonomic Bulletin & Review*, 23(1), 331–337. https://doi.org/10.3758/s13423-015-0967-4

- Carstens, J. (2016). Goal management training in undergraduate students: The effects on executive functioning skills and academic self-efficacy (No. 85805) [Published doctoral dissertation, University of Windsor, Toronto, Canada]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/5805
- Clements, D. H., & Sarama, J. (2019). Executive function and early mathematical learning difficulties. In A. Fritz, V. Haase, & P. Räsänen (Eds.), *International handbook of mathematical learning difficulties* (pp. 755–771). Springer. http://doi.org/ 10.1007/978-3-319-97148-3_43
- Cuberos-Urbano, G., Caracuel, A., Valls-Serrano, C., García-Mochón, L., Gracey, F., & Verdejo-García, A. (2018). A pilot investigation of the potential for incorporating lifelog technology into executive function rehabilitation for enhanced transfer of self-regulation skills to everyday life. *Neuropsychological Rehabilitation*, 28(4), 589–601. http:// doi.org/10.1080/09602011.2016.1187630
- Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2015). A meta-analysis on age differences in risky decisionmaking: Adolescents versus children and adults. *Psychological Bulletin*, 141(1), 48–84. http://doi.org/10.1037/ a0038088
- Diamond, A. (2013). Executive functions. Annual Review of Psychology, 64, 135–168. http://doi.org/10.1146/annurevpsych-113011-143750
- Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3(3), 271–290. https://doi.org/10.1080/02643298608253360
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204. https://doi.org/ 10.1016/j.cortex.2016.04.023
- Grant, M., Ponsford, J., & Bennett, P. C. (2012). The application of Goal Management Training to aspects of financial management in individuals with traumatic brain injury. *Neuropsychological Rehabilitation*, 22(6), 852–873. https://doi.org/10.1080/09602011.2012.693455
- Harwood, A., Lavidor, M., & Rassovsky, Y. (2017). Reducing aggression with martial arts: A meta-analysis of child and youth studies. *Aggression and Violent Behavior*, 34, 96–101. https://doi.org/10.1016/j.avb.2017.03.001
- Homer, B. D., Plass, J. L., Raffaele, C., Ober, T. M., & Ali, A. (2018). Improving high school students' executive functions through digital game play. *Computers & Education*, 117, 50–58. https://doi.org/10.1016/j.compedu.2017.09.011
- Krasny-Pacini, A., Limond, J., Evans, J., Hiebel, J., Bendjelida, K., & Chevignard, M. (2014). Context-sensitive goal management training for everyday executive dysfunction in children after severe traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, 29(5), E49–E64. https://doi.org/10.1097/HTR.000000000000015
- Kiani, B., Hadianfard, H., & Mitchell, J. T. (2017). The impact of mindfulness meditation training on executive functions and emotion dysregulation in an Iranian sample of female adolescents with elevated attention-deficit/hyperactivity disorder symptoms. *Australian Journal of Psychology*, 69(4), 273–282. https://doi.org/10.1111/ajpy.12148

Kline, R. B. (2004). Beyond significance testing: Reforming data analysis methods in behavioral research. American Psychological Association.

Kulinna, P. H., Stylianou, M., Dyson, B., Banville, D., Dryden, C., & Colby, R. (2018). The effect of an authentic acute physical education session of dance on elementary students' selective attention. *BioMed Research International.2018*, 1–8. https://doi.org/10.1155/2018/ 8790283

- Levaux, M.-N., Larøi, F., Malmedier, M., Offerlin-Meyer, I., Danion, J.-M., & van der Linden, M. (2012). Rehabilitation of executive functions in a real-life setting: Goal management training applied to a person with schizophrenia. *Case Reports in Psychiatry*, *10*, *Article ID* 503023. https://doi.org/10.1155/ 2012/503023
- Levine, B., Robertson, I. H., Clare, L., Carter, G., Hong, J., Wilson, B. A., & Stuss, D. T. (2000). Rehabilitation of executive functioning: An experimental–clinical validation of Goal Management Training. *Journal of the International Neuropsychological Society*, 6(3), 299–312. https://doi.org/ 10.1017/S1355617700633052
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). Neuropsychological assessment (5th Ed.). Oxford University Press.
- Meixner, J. M., Warner, G. J., Lensing, N., Schiefele, U., & Elsner, B. (2019). The relation between executive functions and reading comprehension in primary-school students: A cross-lagged-panel analysis. *Early Childhood Research Quarterly*, 46, 62–74. https://doi.org/10.1016/ j.ecresq.2018.04.010
- Menezes, A., Dias, N. M., & Seabra, A. G. (2012). Evidências de validade do Teste da Torre de Londres [Evidence of validity of the Tower of London Test]. In A. G. Seabra & N. M. Dias (Eds.), Avaliação neuropsicológica cognitiva: Atenção e funções executivas [Cognitive neuropsychological assessment: attention and executive functions] (pp. 101–105). Memnon.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. https://doi.org/10.1006/ cogp.1999.0734
- Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G., & Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine*, 40(6), 899–909. https://doi.org/10.1017/S0033291709991036
- Nicoló, C. (2015). Exploring the implementation of attention and goal management training interventions with a boy diagnosed with attentional difficulties in a single case design [Published master's thesis University of Cape Town, Cape Town, South Africa]. http://www.psychology.uct.ac.za/sites/default/files/ image_tool/images/117/Logos/thesis/Candice% 20Nicolo.pdf
- Novakovic-Agopian, T., Kornblith, E., Abrams, G., Burciaga-Rosales, J., Loya, F., D'Esposito, M., & Chen, A. J. W. (2018). Training in goal-oriented attention self-regulation improves executive functioning in veterans with chronic traumatic

brain injury. *Journal of Neurotrauma*, 35(23), 2784–2795. https://doi.org/10.1089/neu.2017.5529

- Noguchi, K., Gel, Y. R., Brunner, E., & Konietschke, F. (2012). nparLD: An R Software Package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal* of Statistical Software, 50(12), 1–23. http://doi.org/10.18637/ jss.v050.i12
- Primi, R. (2002). Inteligência fluida: Definição fatorial, cognitiva e neuropsicológica [Fluid intelligence: Factorial, cognitive and neuropsychological definition]. *Paidéia*, 12(23), 57–75. https://doi.org/10.1590/S0103-863X2002000005
- Pope, C. N., Ross, L. A., & Stavrinos, D. (2016). Association between executive function and problematic adolescent driving. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 37(9), 702–711. https://doi.org/10.1097/ DBP.000000000000353
- Rivera, C. J., Wood, C. L., James, M., & Williams, S. (2019). Improving study outcomes for college students with executive functioning challenges. *Career Development and Transition for Exceptional Individuals*, 42(3), 139–147. https:// doi.org/10.1177/2165143418758256
- Robertson, I. H. (1996). Goal management training: A clinical manual. PsyConsult.
- Sedó, M., de Paula, J. J., & Malloy-Diniz, L. F. (2015). FDT-Teste dos Cinco Dígitos [Five Digit Test]. Hogrefe.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). A compendium of neuropsychological tests: Administration, norms, and commentary. Oxford University Press.
- Stubberud, J., Langenbahn, D., Levine, B., Stanghelle, J., & Schanke, A-K. (2014). Goal management training improves everyday executive functioning for persons with spina bifida: Self-and informant reports six months posttraining. *Neuropsychological Rehabilitation*, 24(1), 26–60. https://doi.org/10.1080/09602011.2013.847847
- Trentini, C. M., Yates, D. B., & Heck, V. S. (2014). Escala de Inteligência Wechsler Abreviada (WASI): Manual profissional [Wechsler Abbreviated Scale of Intelligence: Professional manual]. Casa do Psicólogo.
- Trevisan, B. T. (2014). Adaptação e desenvolvimento de instrumentos para avaliação de crianças e adolescentes com TDAH, análise de perfil neuropsicológico e relação com o desempenho escolar [Adaptation and development of

instruments for the assessment of children and adolescents with ADHD, analysis of neuropsychological profile and relationship with school performance]. [Doctoral dissertation, Universidade Presbiteriana Mackenzie, São Paulo, Brazil]. http://tede.mackenzie.br/jspui/handle/ tede/1781

- Trevisan, B. T., Berberian, A. A., Dias, N. M., Alves, R. J. R., & Seabra, A. G. (in press). Development and psychometric properties of the Difficulties in Executive Functions, Regulation and Delay Aversion Inventory (IFERA–I). *Psico-USF.*
- Tornås, S., Løvstad, M., Solbakk, A. K., Evans, J., Endestad, T., Hol, P. K., Schanke, A.-K., & Stubberud, J. (2016). Rehabilitation of executive functions in patients with chronic acquired brain injury with goal management training, external cuing, and emotional regulation: A randomized controlled trial. *Journal of the International Neuropsychological Society*, 22(4), 436–452. https://doi.org/ 10.1017/S1355617715001344
- Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2017). *Executive function: Implications for education* (NCER 2017– 2000). National Center for Education Research, Institute of Education Sciences. https://ies.ed.gov/ncer/pubs/ 20172000/pdf/20172000.pdf
- Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x
- Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood: I. The development of executive function. *Monographs of the Society for Research in Child Development*, 68(3), 11–27. http://doi.org/10.1111/j.1540-5834.2003. 06803002.x
- Wechsler, D. (1999). Weschler Abbreviated Scale of Intelligence (WASI). Psychological Corporation.
- Wechsler, D. (2013). Escala Wechsler de Inteligência para Crianças: WISC-IV. Manual técnico. Tradução do manual original Maria de Lourdes Duprat (4th Ed.) [Wechsler intelligence scale for children: WISC-IV. Technical Manual. Translation from original manual Maria de Lourdes Duprat (4th Ed.)]. Casa do Psicólogo