

---

# Executive demands of the Tower of London task in Brazilian teenagers

Natália Martins Dias and Alessandra Gotuzo Seabra

Universidade Presbiteriana Mackenzie, São Paulo, SP, Brazil

## Abstract

Some executive tasks may require diverse abilities to solve, having implications for the comprehension of an individual's performance. This study investigated the processes involved in the resolution of a complex test of executive function (EF): the Tower of London (ToL) task. One hundred twenty-four healthy teenagers, 11–14 years old ( $M = 12.5$ ,  $SD = 1.14$ ) participated in the study. They were tested using the Auditory Working Memory Test, Visual Working Memory Test, Computerized Stroop Test, Semantic Generation Test, Cancellation Attention Test, and Trail Making Test B, in addition to the ToL. Multivariate analysis showed significant effects of age on visual and auditory working memory, interference control, and planning, with a marginal effect on selective attention. A gender effect was found only on visual working memory where boys scored higher than girls. Significant correlations were found between EF measures despite their low and moderate magnitudes. Performance on the ToL task was correlated with some executive ability measures. The most difficult problems that required four and five moves were more correlated with EF measures than the easier problems that required two and three moves. Regression analyses revealed that only auditory working memory integrated the ToL explanation model after controlling for age. However, the explained variability was very modest, suggesting that other abilities not analyzed in this study may participate in the ToL solution. In conclusion, this investigation highlighted the modest role of auditory working memory in ToL scores in a sample of teenagers.

**Keywords:** executive function; assessment; planning; task demands.

Received 28 November 2011; received in revised form 7 January 2012; accepted 16 February 2012. Available online 29 June 2012.

---

## Introduction

Executive function (EF) includes cognitive and metacognitive processes that allow an individual to carry out, control, and regulate their own behavior and cognition, thus permitting involvement in adaptive behaviors, self-organization, and direction toward aims. The processes involved in EF are believed to have different but interconnected operations (Fournier-Vicente, Larigauderie, & Gaonac'h, 2008; Miyake et al., 2000) including working memory, inhibitory control, selective attention, planning, cognitive flexibility, and monitoring (Gazzaniga, Ivry, & Mangun, 2006; Gil, 2002; Lezak, Howieson, & Loring, 2004; Malloy-Diniz, Sedo, Fuentes, & Leite, 2008).

Because EF involves multiple dimensions (Miyake et al., 2000), one may wonder what some of the specific tests conceived to evaluate EF are actually appraising. This matter has been addressed by Strauss, Sherman,

and Spreen (2006) as a problem of impure tasks. An example is the Wisconsin Card Sorting Test (WCST). Although this test is classically used for assessing neuropsychological performance in EF (Rabin, Barr, & Burton, 2005), some studies suggested that its resolution involves planning, abstract concept elaboration, visual processing skills, numerical skills, rules induction, working memory, inhibition, cognitive flexibility, and feedback processing, in addition to a reasonably adequate level of motivation (Silva-Filho, 2007; Strauss et al., 2006). The test, therefore, should be considered a complex performance task (Miyake et al., 2000) taking into account its limitations in isolating skills or specific mental processes (Gazzaniga et al., 2006).

Miyake et al. (2000) investigated the requirements for solving complex executive tests including the WCST, Tower of Hanoi Task (ToH), Random Number Generation (RNG), Operations Span (OS), and Dual Task (DT) performed by healthy adults. Their results showed that flexibility is crucial for perseverative errors in the WCST. Inhibition is fundamental with regard to ToH performance, and inhibition and working memory are both involved in the RNG solution. In the OS task, working memory was also implicated as a major ability involved in participant performance. In the DT, no single model was identified that was statistically

---

Natália Martins Dias and Alessandra Gotuzo Seabra, Developmental Disabilities Program, Universidade Presbiteriana Mackenzie, São Paulo, SP, Brazil. Correspondence regarding this article should be directed to Natália Martins Dias, Est. Municipal Nemésio Dário dos Santos, s/n-Cx. Postal = 45, Moenda, Itatiba, São Paulo, Brasil, CEP 13.252-755. Phone +55 11 9818-7335. E-mail: natalia\_mdias@yahoo.com.br

suitable. Therefore, working memory, inhibition, and flexibility contribute in different ways to performance in complex tests of EF.

Because of the diversity of processes involved in complex executive tests, conducting research to develop or adapt and validate instruments to register specific processes or simple tasks is fundamental. This has already been the goal of studies performed in Brazil (Dias, 2009; Dias, Menezes, & Seabra, 2012; Seabra, Dias, & Trevisan, 2010). Understanding the nature of the specific processes involved in the performance of complex executive tests is also essential. Thus, determining which specific aspects such as working memory, selection, inhibitory control, and cognitive flexibility contribute to the performance of different complex EF assessments is important. This understanding, however, must consider the developmental course of these functions. In children and teenagers, these cognitive functions can be immature; therefore, partial contributions of each cognitive process can be distinct from those observed in adults.

In fact, compared with other cognitive abilities, EF matures later (Malloy-Diniz et al., 2008). Studies suggest that some executive abilities such as inhibitory control and cognitive flexibility appear to reach mature performance levels only in adolescence, whereas others such as working memory do not appear to develop until the beginning of adulthood (Best & Miller, 2010; Davidson, Amso, Anderson, & Diamond, 2006; Huizinga, Dolan, & van der Molen, 2006). These findings also support the notion that different executive abilities develop in different ways (Best, Miller, & Naglieri, 2011; Dias et al., 2012). In terms of planning ability, some evidence indicates that the development of this ability accelerates between the ages of 5 to 9 years and is slower in adolescence (Matute et al., 2008), despite evidence that performance continues to improve until 15 to 17 years of age (Luciana, Collins, Olson, & Schissel, 2009). These findings are corroborated by other reports that noted a relationship between structural changes and the functional pattern of the prefrontal cortex, which are particularly noticeable during adolescence, and the development of cognitive control and EF (Durstun & Casey, 2006; Lamm, Zelazo, & Lewis, 2006). Moreover, the interrelationships between EFs appear to change throughout the developmental process in a manner that indicates that these abilities are more interdependent in children and become more differentiated and noticeable in early adolescence (Cuevas & Alanís, 2008). Evidence also indicates that the associations between EF and other skills such as educational achievement can change during the developmental process. For example, Gathercole, Pickering, Knight, and Stegmann (2004) reported that the association between working memory, a capacity related to EF, and performance on national curriculum assessments changed in different age groups. Working memory ability was significantly associated with achievement in English (reading and writing) in 7-year-olds but not in 14-year-olds.

To understand the nature of the processes involved in the performance of complex tests, considering the age of the participants being tested and the developmental course of EF is necessary. This was the purpose of the present study where we investigated the relative contributions of various EFs (auditory and visual working memory, selective attention/interference control, inhibitory control, cognitive flexibility, and attentional skills) on the performance of children and adolescents in the Tower of London (ToL) task, an instrument that involves various executive processes for its resolution (Lezak et al., 2004).

### ***Planning ability and the Tower of London task***

Planning is a central cognitive component in any problem-solving task that requires the identification and organization of a series of acts and elements aimed to attain a goal (Krikorian, Bartok & Gay, 1994; Lezak et al., 2004). Planning can be considered an action that depends on the cooperative achievement of other functions. Among these functions, working memory, decision-making ability, inhibitory control, flexibility, and sustained attention are some of the basic abilities required for a plan to be elaborated (Lezak et al., 2004; Malloy-Diniz et al., 2008). In other words, the skill of planning can be considered a complex ability that relies on other functions (Miyake et al., 2000).

Tests based on tower paradigms have been classically used to assess planning ability in clinical and experimental settings (Kaller, Unterrainer, & Stahl, 2011; Lezak et al., 2004; Strauss et al., 2006; Sullivan, Riccio, & Castillo, 2009). However, despite the similarities between the tower tasks, evidence suggests that these tests, such as the ToL task and ToH task, do not assess identical abilities. The latter imposes a demand on the ability to inhibit a response in the case of a goal/subgoal conflict situation, making the ToH task more complex than the ToL task (Lezak et al., 2004). Because of its reduced complexity, the ToL task, developed by Shallice and McCarthy, has been especially recommended for assessing children (Krikorian et al., 1994) despite evidence that test performance is multifaceted (Berg, Byrd, McNamara, & Case, 2010). This has been corroborated by the notion that individuals with various disorders demonstrate impaired performance on the test (Sullivan et al., 2009). One could ask which specific impairments underlie performance on the test.

Studies have been conducted to investigate the abilities required for the ToL solution. The findings have been considerably divergent possibly because of differences in the sample ages and ToL versions. Huizinga et al. (2006) encountered problems establishing a clear model in their sample of children and teenagers (ages of the groups were 7, 11, and 15 years). However, performance on the Stroop Test, which measures inhibitory control, was a strong predictor of a perfect ToL solution rate in 21-year-old participants. However, another study proved that working memory and inhibitory control could together account for up to 55% of the variance in the performance of young adults

in a revised version of the ToL that contained problems that ranged in move length from four to six moves. Only the score measure was used (i.e., one point was given if the participant solved the problem in only one attempt; Welsh, Satterlee-Cartmell, & Stine, 1999). However, these authors included in their study only a measurement of visuospatial working memory. Another important aspect to be considered is that one of the two inhibition measures imposed a greater demand on cognitive flexibility.

In contrast to the majority of the findings, Unterrainer et al. (2004) noted the absence of a significant contribution of working memory to ToL performance in college participants (average age 25.5 years;  $SD$  4.6 years). These authors used a ToL version with three to seven moves, and the score used was the weighted performance score defined as the summed number of moves for the correctly solved trials. Fluid intelligence ( $Gf$ ) measured by a computerized version of Raven's Standard Progressive Matrices was the only important predictor of ToL performance ( $R^2_{adj} = 13.6\%$ ) despite the fact that the study considered verbal and visuospatial working memory measures. The authors concluded that ToL performance could not be sufficiently explained by other cognitive areas. A factor that may have altered the outcome was the measures that were included as possible predictors. In this case a test of fluid intelligence ( $Gf$ ), which is possibly more complex than pure EF measures, was used. The executive ability measures were already partially included in the evaluation, which may have reduced the predictive power of these measures. In fact, the correlation between Raven's Standard Progressive Matrices and working memory tests has been consistently demonstrated and attributed to the need to hold a certain number of items in both tasks or control attention (for review, see Heitz, Unsworth, & Engle, 2004).

In terms of ToL execution time, Pulos and Denzine (2005) noted that this measure was related to visual working memory in undergraduate participants ( $r = -.41, p < .001$ ) and auditory working memory ( $r = -.26, p < .05$ ), without significant correlations with measures on the Stroop Test and Trail Making Test-B. The authors concluded that individuals with increased working memory capacity could solve ToL tasks more rapidly. Moreover, the lack of a relationship between performance on the ToL task and Stroop Test and Trail Making Test-B can be attributed to the nature of the planning involved. Thus, when online planning may occur, the importance of other executive abilities is minimized. The ToL, however, may have different predictors of execution time and score, which may explain the different results observed.

Therefore, inhibitory control and working memory appear to be the best predictors of ToL performance in adults. Studies conducted with children and adolescents, although somewhat less conclusive, tended to show similar results with regard to predictors of ToL performance. For example, in a heterogeneous sample of children aged 8 years to adults aged 30 years, inhibitory control, assessed using the antisaccade task with eye movement

data, was the best predictor of ToL performance in terms of the number of moves; however, other measures such as working memory and rapid processing, were also associated with test performance (Asato, Sweeney, & Luna, 2006). These correlations were more evident in the more complex items of the ToL task, particularly for items that represented problems that could be solved in four or five moves. According to the authors, separating the analyses by age group did not alter the noted correlations, but the association between ToL performance and inhibitory control was especially applicable to children. The results also indicated that ToL performance matured during adolescence, an idea supported by the parallel development of other cognitive processes such as inhibitory control and working memory through childhood until adolescence. Notably, these abilities were assessed by eye movement tracing techniques and, again, no pure auditory working memory measurement was available.

Data from Luciana et al. (2009) also showed that response inhibition (Go-No-go task) and working memory (digit span backwards) were related, although with a low magnitude, to ToL performance in a heterogeneous sample of 9- to 20-year-old individuals. After controlling for intelligence quotient (IQ) and age, only working memory contributed to accuracy in ToL performance (number of perfect solutions and planning time  $\times$  accuracy index), but the authors did not assess a wide range of EFs such as flexibility, selective attention or, again, different subsystems of working memory.

Reviews by Batista, Adda, Miotto, Lucia, and Scaff (2007) and Unterrainer and Owen (2006) corroborated the absence of a consensus with regard to which measures, among memory, working memory, inhibitory control, intelligence, and cognitive flexibility, can best predict ToL performance. Despite the divergence, inhibitory control and working memory tended to be the best predictors of ToL performance. These findings are reasonable because the task required in this test depends on establishing a plan of action, keeping it in memory while the movements are sequentially executed and avoiding impulsive movements and prepotent answers. In some cases, momentarily averting the desired solution and achieving it within the rules of the task may be necessary (Asato et al., 2006).

Another nonconsensual aspect described by Unterrainer and Owen (2006) is the fact that the literature does not report any consensus about the effects of the auditory and visuospatial subsystems of working memory on task performance. These authors have highlighted that the divergences noted between the above findings may be attributable to the large number of versions or adaptations of the ToL and differences in the instructions and measurements taken. They also argued for the need to use the original version of the test, thus allowing a more reliable comparison between investigation results. To overcome this problem, the present study used the original ToL version proposed by Shallice and McCarthy and systematically presented by Krikorian et al. (1994).

Notably, the majority of the studies summarized in this introduction have investigated processes related to

the solution of complex tests using samples composed of adult participants. In studies that evaluated children and adolescents, the results appear to be similar to the results obtained in adults. However, the samples were very heterogeneous so the precise executive demands involved in the ToL task remain unclear. Owing to the shortage of investigations on younger participants, the objective of this work was to investigate the relative contributions of various EFs to ToL performance using a sample composed of children and adolescents. We expected that inhibitory control and working memory measures would emerge as the best predictors of performance in terms of ToL scores.

## Methods

### Participants

Initially, 159 children and teenagers aged 10–17 years and recruited at a public school in the city of São Paulo, Brazil, were tested. Because many of the participants had discrepancies between their ages and school grades, we excluded participants who had either repeated grades or had started school late. Additionally, because of the limited number of 10-year-old participants ( $n = 11$ ), we opted to exclude them. As a result, the final sample comprised 124 children and teenagers aged 11–14 years ( $M = 12.5$  years,  $SD = 1.14$  years) who attended junior high school. Among these, 53 (42.7%) were males. Although no formal testing was performed to rule out intellectual disability beyond the screening for discrepancies between age and school grade (grade retention), no students with mental disabilities or uncorrected sensory disorders were included in the sample based on records and information provided by the participating schools. Table 1 shows the distribution of participants as a function of age and gender and the

minimum and maximum ages in months per gender in each age group. All participants were from middle-low and low socioeconomic status.

### Instruments

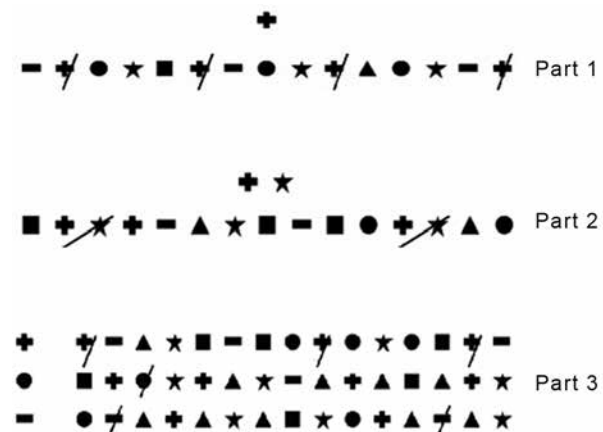
*Cancellation Attention Test (CAiT; Montiel & Seabra, 2009a, 2012a):* The CAiT assesses selective attention in a test of visual search. The test is conducted collectively and consists of three matrices with a target and 300 stimuli in each matrix, always geometric shapes. Similar to the classic cancellation paradigms, the participant should identify and mark the stimuli that are similar to the target. Each matrix has a fixed time of 1 min. The total number of correct answers was used as the score, which could vary from 0 to 132 points. Figure 1 shows examples of the CAiT task.

*Auditory Working Memory (AWM) test (Primi, 2002):* The AWM test is a computerized test that evaluates the auditory working memory of the subject. The AWM test is conducted individually with unlimited time to solve the task. The software emits digitalized voice sequences that consist of two to 10 words and numbers. The task is to first repeat the words in the sequence and then the numbers in increasing order. For example, if the stimulus is “three - bird - one - rain,” then the correct answer is “bird - rain - one - three.” The score is based on the number of correct sequences and can vary from 0 to 27 points.

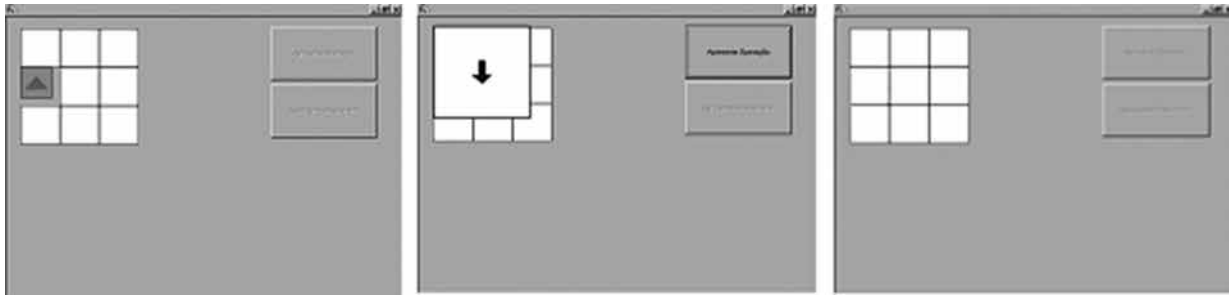
*Visual Working Memory (VWM) test (Primi, 2002):* The VWM test is a computerized test that assesses visual working memory. The test is administered individually with unlimited time to solve the task. On the computer screen, one  $3 \times 3$  matrix is shown that contains a stimulus (a geometric shape) in some cell. Spatial manipulations represented by arrows that indicate the direction of each move are then shown on the screen. The participant should imagine the stimulus moving on the matrix and select its final position. The task has increasing difficulty until four matrices are shown on the same screen. The score is the number of correct answers and may vary from 0 to

**Table 1.** Sample characteristics by age group

Age group	Gender	<i>n</i>	%	<i>p</i>	Age (months)		<i>p</i>
					Min	Max	
11	M	18	52.9	.732	132	143	.644
	F	16	47.1		132	142	
	Total	34	100.0		132	143	
12	M	10	41.7	.414	144	155	.449
	F	14	58.3		144	155	
	Total	24	100.0		144	155	
13	M	14	38.9	.182	156	167	.850
	F	22	61.1		156	167	
	Total	36	100.0		156	167	
14	M	11	36.7	.144	168	176	.497
	F	19	63.3		168	177	
	Total	30	100.0		168	177	



**Figure 1.** Examples of the required tasks in each part of the CAiT.



**Figure 2.** Sequence of the screens presented on one item of the VWM test with one matrix task.

26 points. Figure 2 shows a sequence of screens of a VWM item, with one matrix task.

*Trail Making Test (TMT;* Gil, 2002): The TMT assesses cognitive flexibility. The test is conducted collectively and the Brazilian adaptation by Montiel and Seabra (2009b, 2012b) was used. The task is composed of items (i.e., letters and numbers) that must be connected in sequence but alternately during a 1-min period. The score in sequences (i.e., the number of items connected correctly) is used. The maximum score is 24 points.

*Semantic Generation Test (SeGeT;* Seabra, Cozza, Capovilla, Macedo, & Dias, 2009): The SeGeT is a computerized test that measures inhibitory control ability. The SeGeT is conducted individually with unlimited time to solve the task. The test was developed based on the work of Thompson-Schill, Swick, Farah, D'Exposito, Kan, and Knight (1998). The subject is asked to name an action that is semantically associated with drawings (i.e., nouns) that appear on a computer screen. The nouns can be low-selection (60 items), meaning that they evoke just one word (e.g., *scissors* is usually associated with the verb *cut*) or high-selection (60 items), meaning that they can evoke a list of words (e.g., *string*, which can be associated with the words *lace*, *tie*, *leap*, or *roll up*). The differential analysis of performance based on the low- and high-selection conditions (performance on high-selection condition minus performance on low-selection condition) provides a score and reaction time interference index. Therefore, the measures used are the interference effects for score and reaction time. A greater interference effect for reaction time and lower effect for score are associated with the participant being more susceptible to performance interference. The test has no maximum or minimum measure for the effect of interference on reaction time. With regard to the score, the measure of interference can vary from -60 to 60. The interference measure would be expected to be negative because the high-selection condition is theoretically more difficult than the low-selection condition (Thompson-Schill et al., 1998).

*Computerized Stroop (Stroop-Comp) test* (Seabra, Macedo, & Dias, in press): The Stroop-Comp test measures selective attention and the ability to control interference. The test is conducted individually with unlimited time to solve the task, despite the stimulus presentation time of 40 ms. This computerized version

of the Stroop Test was adapted from the Victoria Version (Regard, 1981). In the first part, the participant must read, as quickly as possible, the items on a screen that correspond to the words *yellow*, *blue*, *green*, and *red*, all printed in black ink. This part is used to ascertain the reading ability of the participant. In the second part, the participant must name the color of the circles that appear on the screen, with four possible colors: yellow, blue, green, and red. In the third part, the participant must name the color of the written words that correspond to the four colors above, but they are printed in another color that disagrees with their meaning (for example, the word *green* is written in blue). The interference effect for word vs. color, or the Stroop effect (performance on part 3 minus performance on part 2), is analyzed in terms of score and reaction time. Thus, the measures used in the present study were the interference effects for score and reaction time. A greater interference effect for reaction time and a lower effect for score are associated with the participant being more susceptible to performance interference. The test has no maximum or minimum measure for the effect of interference on reaction time. With regard to the score, the measure of interference can vary from -24 to 24. Because part 3 is theoretically more difficult than part 2, the interference measure is expected to be negative.

*Tower of London (ToL) task* (Krikorian et al., 1994; Brazilian version of Seabra, Dias, Berberian, Assef & Cozza, 2012): This test evaluates planning ability. The task is composed of a wooden platform with three vertical rods and three colored spheres (red, green, and blue). The participant must transpose the three spheres, starting from an initial position, to reach various final positions. The task comprises 12 problems (i.e., target positions), the level of difficulty of which increases as a function of the number of moves necessary to reach the final position, with a variation of two to five moves, with unlimited time to solve the task. The measure used is total score. For each problem, the participant earns three points if a solution is reached in only one attempt, two points if the solution is reached in two attempts, one point if the solution is reached in three attempts and zero points if the solution is not reached in three attempts. Therefore, the total score can range from 0 to 36 points.

All of the above-described tests were selected for this study based on their psychometric properties as

validated in samples of Brazilian children and teenagers in studies by Assef, Capovilla, and Capovilla (2007), Capovilla, Dias, Trevisan, Montiel, and Andrade (2008), and Dias et al. (2012). Furthermore, we calculated Cronbachs' Alpha reliability coefficients for our sample for the SeGeT and Stroop-Comp test. We found .65 for score and .72 for reaction time in the low-selection condition on the SeGeT and .89 for score and .84 for reaction time in the high-selection condition on the same instrument. In the Stroop-Comp test, we found .40 for score and .76 for reaction time in the second part of the test, and .70 for score and .80 for reaction time in the third part of the test. We used the original version of the ToL task as systematized by Krikorian et al. (1994) in an attempt to minimize the challenge inherent in the use of diverse test versions. Certain computerized tests were also used including the SeGeT and Stroop-Comp test. The computerized tests allow for additional quantitative measurements such as reaction time in addition to score registration, which may be the most sensitive of the cognitive processes under analysis (Assef et al., 2007).

### Procedure

The project was approved by the Ethics Research Committee. A consent form was sent to the School Board of Directors and students' parents to obtain their consent to carry out the study. Application of the instruments occurred in a classroom made available by the school during regular school hours. Data collection began with administration of the collective application instruments, initially involving the CAiT and TMT, in the classroom over two sessions. The other instruments were applied in individual sessions for a total of five evaluation sessions for each participant, with an average time of 20 min for each session. The order of the application of the tests was counterbalanced, and only one test per day was administered to each participant, with an interval of 2 or 3 days between sessions with the same child.

### Statistical analyses

Descriptive statistics and a multivariate analysis with age and gender as independent variables were performed. Pearson correlation analyses were performed between the EF measures and between the EF measures and performance on the total ToL score, the two and three movement tasks (ToL2-3), and the four and five movement tasks (ToL4-5). Additionally, a multiple Linear Regression Analysis was performed to determine which EF variables predicted ToL performance, controlling for age. The independent or predictive variables were performance on the CAiT, AWM test, VWM test, SeGeT, TMT, and Stroop-Comp test. The Enter method was used to explore the contributions of each variable.

### Results

Descriptive statistics and results from the multivariate analysis of variance (MANOVA) are shown in Table

2. Age had a significant effect on performance on the AWM test ( $F [3,98] = 6.015, p = .001$ ), performance on the VWM test ( $F [3,98] = 3.205, p = .027$ ), interference reaction time on the Stroop-Comp test ( $F [3,98] = 8.408, p \leq .001$ ), and ToL task ( $F [3,98] = 3.371, p = .022$ ), with a marginally significant effect on performance on the CAiT ( $F [3,98] = 2.579, p = .058$ ). A tendency toward increased performance was found with increasing age on the AWM test, VWM test, ToL task, and CAiT and a tendency toward decreased reaction time interference on the Stroop-Comp test in the age groups. Only one significant effect of gender was found on the VWM measure ( $F [1,98] = 4.741, p = .032$ ) where boys outperformed girls. No significant group  $\times$  gender interaction was found.

Pearson analysis between EF measures is presented in Table 3, and the analysis between EF measures and ToL performance is shown in Table 4. Some significant correlations with low and moderate magnitude were found between the EF measures. Performance on the easier ToL problems requiring two and three moves correlated with the auditory working memory measure but only to a low degree. Performance on more demanding problems that required four and five moves correlated with cognitive flexibility and auditory and visual working memory measures, despite the low magnitude of these correlations. Total performance on the ToL task was positively and significantly correlated with auditory and visual working memory, flexibility, and selective attention, in this order. Significant correlations are indicated in bold in Table 3.

A multiple linear regression analysis was conducted. The first step was to control for age. The model adjustment was significant ( $p = .004$ ). In fact, age had a significant effect on ToL performance, explaining up to 7.2% of the variance in test performance. This finding indicated that controlling for age was indeed necessary. All of the EF measures were then included in the regression (Enter mode of selection). The model adjustment was significant ( $p = .008$ ), and the inclusion of EFs improved the explanatory power of the model to 13.6%. However, only the auditory working memory measure significantly contributed to ToL performance ( $\beta = .33, p = .005$ ) after controlling for age. In fact, after the EF measures were included with the AWM contributions to ToL performance, the effect of age lost statistical significance. No other EF measure manifested a significant regression coefficient. The summary of the models and regression coefficients is presented in Table 5. The statistics related to the residues and multicollinearity diagnostics exhibited adequate values, indicating that the results of the analysis can be considered relatively reliable.

### Discussion

The MANOVA showed that performance on some EF tests did not differ between the 11- and 14-year-old age groups, whereas others showed an improvement with increasing age. Thus, auditory and visual working memory, interference control (in terms of reaction time

**Table 2.** Mean scores and standard deviations according to age group and gender and multivariate analysis with *p* statistics

Groups		Performance—M (SD)								
		TMT	CAcT	AWM	VWM	SeGeT (score)	SeGeT (RT)	Stroop-Comp (score)	Stroop-Comp (RT)	ToL
11	M	14.53 (6.12)	79.60 (27.24)	9.27 (2.43)	7.33 (3.02)	-2.67 (4.03)	1.46 (1.18)	-1.53 (2.42)	.58 (.15)	27.13 (2.67)
	F	12.67 (2.96)	71.33 (23.01)	8.33 (3.68)	5.17 (2.29)	-1.08 (2.11)	1.16 (.73)	-2.75 (2.77)	.73 (.18)	28.00 (3.1)
	Total	13.70 (4.98)	75.93 (25.32)	8.85 (3.02)	6.37 (2.88)	-1.96 (3.36)	1.32 (1.0)	-2.07 (2.6)	.65 (.18)	27.52 (2.86)
12	M	16.00 (2.96)	87.33 (10.21)	10.78 (3.35)	7.89 (2.15)	.00 (2.69)	.73 (.78)	-3.11 (2.32)	.64 (.24)	28.89 (2.21)
	F	14.92 (6.03)	87.15 (16.99)	10.38 (3.04)	8.00 (2.24)	-3.69 (4.84)	1.47 (1.68)	-1.54 (2.22)	.52 (.18)	28.77 (3.68)
	Total	15.36 (4.94)	87.23 (14.30)	10.55 (3.1)	7.95 (2.15)	-2.18 (4.43)	1.17 (1.41)	-2.18 (2.34)	.57 (.21)	28.82 (3.1)
13	M	16.90 (2.03)	82.20 (13.82)	12.90 (2.93)	9.10 (4.61)	-1.80 (1.87)	1.03 (.81)	-1.60 (2.92)	.50 (.28)	30.0 (3.34)
	F	15.29 (7.74)	91.64 (14.6)	11.64 (2.47)	8.00 (3.16)	-1.43 (3.23)	1.14 (.56)	-.57 (.94)	.42 (.18)	30.79 (3.53)
	Total	15.96 (6.01)	87.71 (14.75)	12.17 (2.68)	8.46 (3.78)	-1.58 (2.70)	1.1 (.66)	-1.00 (2.02)	.46 (.22)	30.46 (3.40)
14	M	16.22 (5.74)	90.33 (22.9)	11.00 (3.32)	10.00 (2.55)	-2.00 (3.00)	.75 (.45)	-1.78 (2.28)	.44 (.12)	30.11 (3.72)
	F	18.06 (5.79)	89.13 (21.42)	11.75 (3.11)	7.44 (4.12)	-1.94 (3.47)	.69 (.45)	-.69 (1.08)	.41 (.13)	29.69 (4.22)
	Total	17.40 (5.72)	89.56 (21.50)	11.48 (3.14)	8.36 (3.79)	-1.96 (3.25)	.71 (.44)	-1.08 (1.66)	.42 (.13)	29.84 (3.98)
Total	M	15.74 (4.71)	84.07 (20.66)	10.79 (3.15)	8.42 (3.29)	-1.77 (3.2)	1.06 (.93)	-1.93 (2.48)	.54 (.20)	28.79 (3.17)
	F	15.44 (6.12)	85.42 (20.24)	10.65 (3.28)	7.22 (3.25)	-2.04 (3.6)	1.1 (.97)	-1.31 (1.98)	.51 (.21)	29.38 (3.74)
Total	Total	15.57 (5.52)	84.83 (20.33)	10.71 (3.21)	7.74 (3.30)	-1.92 (3.41)	1.08 (.95)	-1.58 (2.22)	.53 (.20)	29.12 (3.5)
Age differences ( <i>p</i> )		> .05	.058	<b>&lt;.001</b>	<b>.027</b>	> .05	> .05	> .05	<b>&lt;.001</b>	<b>.022</b>
Bonferroni <i>post hoc</i> analysis		—	—	11<13,14	—	—	—	—	11>13, 14 12>14	11<13
Gender differences ( <i>p</i> )		> .05	> .05	> .05	<b>.032</b> M > F	> .05	> .05	> .05	> .05	> .05
Age × gender interaction ( <i>p</i> )		> .05	> .05	> .05	> .05	> .05	> .05	> .05	> .05	> .05

(Legends as in Table 3)

Significant results in bold.

**Table 3.** Correlations between executive function (EF) measurements

		TMT	CAtT	AWM	VWM	SeGeT (score)	SeGeT (RT)	Stroop-Comp (score)
CAtT	<i>r</i>	<b>.18</b>						
	<i>P</i>	<b>.050</b>						
AWM	<i>r</i>	.17	.13					
	<i>P</i>	.072	.161					
VWM	<i>r</i>	<b>.22</b>	<b>.23</b>	<b>.40</b>				
	<i>P</i>	<b>.024</b>	<b>.015</b>	<b>&lt;.001</b>				
SeGeT (score)	<i>r</i>	.12	.06	.09	.12			
	<i>p</i>	.229	.557	.312	.204			
SeGeT (RT)	<i>r</i>	-.07	.054	-.11	.01	<b>-.28</b>		
	<i>p</i>	.468	.583	.255	.915	<b>.003</b>		
Stroop-Comp (score)	<i>r</i>	.18	.17	<b>.32</b>	<b>.27</b>	-.02	-.02	
	<i>p</i>	.059	.081	<b>&lt;.001</b>	<b>.003</b>	.778	.843	
Stroop-Comp (RT)	<i>r</i>	<b>-.21</b>	-.18	<b>-.23</b>	<b>-.36</b>	-.003	.10	<b>-.46</b>
	<i>p</i>	<b>.024</b>	.061	<b>.013</b>	<b>&lt;.001</b>	.974	.290	<b>&lt;.001</b>

TMT, sequence score on the Trail Making Test; CAtT, total number of correct answers on the Cancellation Attention Test; AWM, total number of correct sequences on the Auditory Working Memory Test; VWM, total number of correct items on the Visual Working Memory Test; SeGeT (score), interference score (performance on high-selection trials minus performance on low-selection trials) on the Semantic Generation Test; SeGeT (RT), interference reaction time in seconds (reaction time on high-selection trials minus reaction time on low-selection trials) on the Semantic Generation Test; Stroop-Comp (score), interference score (performance on incongruent trials minus performance on congruent trials) on the Computerized Stroop Test; Stroop-Comp (RT), interference reaction time in seconds (reaction time on incongruent trials minus reaction time on congruent trials) on the Computerized Stroop Test; ToL2-3, score on two and three movement tasks on the ToL task; ToL4-5, score on four and five movement tasks on the ToL task; ToL, total score on the ToL task. Significant results in bold.

**Table 4.** Correlations between EF measurements and ToL performance

		ToL2-3	ToL4-5	ToL
TMT	<i>r</i>	.14	<b>.24</b>	<b>.26</b>
	<i>p</i>	.123	<b>.008</b>	<b>.005</b>
CAtT	<i>r</i>	.15	.16	<b>.19</b>
	<i>p</i>	.098	.093	<b>.041</b>
AWM	<i>r</i>	<b>.18</b>	<b>.29</b>	<b>.32</b>
	<i>p</i>	<b>.045</b>	<b>.001</b>	<b>&lt;.001</b>
VWM	<i>r</i>	.03	<b>.27</b>	<b>.24</b>
	<i>p</i>	.765	<b>.003</b>	<b>.008</b>
SeGeT (score)	<i>r</i>	-.03	-.07	-.07
	<i>p</i>	.739	.450	.447
SeGeT (RT)	<i>r</i>	-.02	-.06	-.06
	<i>p</i>	.865	.530	.552
Stroop-Comp (score)	<i>r</i>	.08	.15	.15
	<i>p</i>	.413	.111	.095
Stroop-Comp (RT)	<i>r</i>	-.10	-.03	-.06
	<i>p</i>	.287	.779	.538

EF, executive function; ToL, Tower of London.

ToL2-3, score on two and three movement tasks on the ToL task; ToL4-5, score on four and five movement tasks on the ToL task; ToL, total score on the ToL task. Significant results in bold.



**Table 5.** Standardized regression coefficients ( $\beta$ ),  $t$  and  $p$  statistics, and  $R^2$  and adjusted  $R^2$  for each measurement included in the regression models

Model		Standardized regression coefficient ( $\beta$ )	$t$	$p$	$R^2$	Adjusted $R^2$
1	(Constant)		4.915		.081	.072
	<b>Age (years)</b>	<b>.285</b>	<b>2.913</b>	<b>.004</b>		
2	(Constant)		4.632	<b>.001</b>	<b>.216</b>	<b>.136</b>
	Age (years)	.118	1.096	.276		
	TMT	.125	1.231	.222		
	CAAt	.045	.450	.654		
	<b>AWM</b>	<b>.326</b>	<b>2.877</b>	<b>.005</b>		
	VWM	.051	.467	.642		
	SeGeT (score)	-.116	-1.161	.249		
	SeGeT (RT)	-.050	-.488	.627		
	Stroop-Comp (score)	-.056	-.500	.618		
	Stroop-Comp (RT)	-.049	-.467	.642		

Dependent variable: ToL

Predictors: Age (years); TMT, sequence score on the Trail Making Test; CAAt, total number of correct answers on the Cancellation Attention Test; AWM, total number of correct sequences on the Auditory Working Memory Test; VWM, total number of correct items on the Visual Working Memory Test; SeGeT (score), interference score (performance on high-selection trials minus performance on low-selection trials) on the Semantic Generation Test; SeGeT (RT), interference reaction time in seconds (reaction time on high-selection trials minus reaction time on low-selection trials) on the Semantic Generation Test; Stroop-Comp (score), interference score (performance on incongruent trials minus performance on congruent trials) on the Computerized Stroop Test; Stroop-Comp (RT), interference reaction time in seconds (reaction time on incongruent trials minus reaction time on congruent trials) on the Computerized Stroop Test. Significant results in bold.

on the Stroop-Comp test), planning, and, marginally, selective attention ability showed an increase with increasing age, whereas cognitive flexibility and inhibitory control tended not to be different among age groups. This result may be attributable to different factors such as the complexity of the tests, but findings from other studies suggest that different executive abilities develop in different ways (Best et al., 2011; Dias et al., 2012). In fact, studies suggest that some abilities such as inhibitory control and cognitive flexibility may reach a mature level of performance only in adolescence, whereas others such as planning appear to develop at 15 or 17 years of age. Working memory appears to develop at the beginning of adulthood (Best & Miller, 2010; Davidson et al., 2006; Huizinga et al., 2006; Luciana et al., 2009). A more detailed discussion of the development of EF can be found in the study by Dias et al. (2012). An effect of gender was found only for one measure (i.e., VWM score) in which boys outperformed girls. This finding was expected because higher scores among males vs. females are usually observed in visuospatial processing tasks (Lippa, Collaer, & Peters, 2010).

Some correlations were also found between EF measures. Generally, the correlations were low and moderate, similar to other studies (e.g., Assef, 2005; Dias, 2009; Seabra et al., 2010). The correlations were positive between performance on each test and negative

with the reaction time measures. Teenagers who had shorter reaction time interference were more efficient and faster in solving the conflict in the Stroop-Comp test and SeGeT and tended to present better performance on the other measures. The magnitude of these correlations agrees with the notion that EF involves different but interconnected operations (Fournier-Vicente et al., 2008; Miyake et al., 2000).

The second correlation analysis showed that ToL performance was associated with the executive measures of cognitive flexibility, selective attention, and auditory and visual working memory. Despite the low magnitude, these correlations are theoretically consistent because they illustrate that executive measures have something in common with the ToL measure and, to some extent, participate in the resolution of this test. Pulos and Denzine (2005) found a correlation between ToL performance and measures of auditory and visual working memory, but these authors used execution time as a measure of ToL performance. In contrast, we used the total score on the ToL task, which may explain why previous findings differed from those of the present study. For example, the authors found no significant correlations with the other executive measures used. In our study, in contrast, significant correlations were found between ToL performance and the flexibility and selective attention measures.

However, our sample consisted of children and adolescents, so the pattern of correlations observed may

differ from studies with adults because of developmental differences. This fact may support the hypothesis that the strategies used to solve the problems on the ToL task in our sample of children and adolescents may be different from those used by adults. Thus, because executive skills are still developing at this age (Best & Miller, 2010; Davidson et al., 2006; Dias et al., 2012; Huizinga et al., 2006), different skills could be recruited to perform the task. However, variations in the measures used between studies limits the conclusions that can be made about this possibility, and future studies should clarify whether indeed different skills are involved in solving the ToL task and other complex tests as a function of the level of development.

Contrary to our expectations, no significant correlation was found between ToL performance and measures of inhibition and interference control. Pulos and Denzine (2005) also found no such relationship in their sample of undergraduates. However, the relationship between these skills has been reported in other studies, especially in children (Asato et al., 2006). This discrepancy may have several explanations including the specific ToL measures used, composition of the samples in these studies, and hence differences in development. Additionally, the differences may be attributable to the specific demands of inhibitory or interference control that are evaluated using different tests. For example, our tests of inhibition and interference control require cognitive control because they do not involve any motor component with the exception of speech for response verbalization. Other studies have used measures (e.g., those supported in a Go-No-go paradigm) that require motor control or measures that involve eye movements (e.g., Asato et al., 2006; Luciana et al., 2009). These studies included children and adolescents in their samples and found relationships between inhibitory control and ToL measures.

Thus, measures that require motor control may be more discriminating in younger subjects because of evidence that different aspects of inhibitory control develop at different ages, with earlier development of motor control relative to cognitive control (Bodrova & Leong, 2007; Nigg, 2001). In samples of adults, this pattern could be reversed. This would explain the findings of Huizinga et al. (2006) where performance on the Stroop Test, a measure of cognitive interference control, was associated only with ToL performance in adult subjects but not in groups of children and adolescents. In our sample, the tests used (SeGeT and Stroop-Comp test) may not have been the most appropriate or the skills assessed by these tests because they are still developing and are not required for ToL resolution.

With regard to the correlations found among the indices formed by the simplest problems that require two and three moves and the most complex problems that require four and five moves, the largest number of associations was found between executive measures and the more complex ToL items. The simplest items

were correlated with only the auditory working memory measure, and the most complex items were more discriminative and correlated with both measures of working memory and performance on the cognitive flexibility test. This pattern was reported by Asato et al. (2006) and is theoretically consistent because the more complex problems may impose a greater demand on the various executive abilities.

Considering the results of the regression analysis after controlling for age, only the auditory working memory measure was a predictor of planning performance in the 11- to 14-year-old sample. The other abilities did not significantly influence test performance in terms of total score.

This result, to some extent, supports our initial hypothesis that working memory contributes to ToL resolution. However, the results also diverged from what was expected because of the lack of evidence of any contribution of inhibitory control to the test solution. Unexpectedly, only auditory working memory, not visual memory, was integrated into the regression model.

The non-inclusion of a measure of inhibition or interference control in the model can be linked to the type of work required and age of the sample (previously discussed). In turn, the contribution of auditory working memory, but not visual memory, to ToL resolution is an important point based on the findings of the present study and extends the results of previous research because many of the prior studies generally did not cover both measures of working memory subsystems (e.g., Asato et al., 2006; Luciana et al., 2009; Welsh et al., 1999). These studies thus maintained the generic notion that working memory or, as a more obvious inference, visual working memory capacity participates in solving the ToL task. In fact, this gap was noted previously by Unterreiner and Owen (2006) who highlighted the fact that the literature does not report any consensus about the effects of the auditory and visuospatial subsystems of working memory performance on ToL performance.

The results of this inquiry are best supported by Pulos and Denzine (2005) and Luciana et al. (2009). The latter study also reported a modest role of working memory in ToL performance, and Pulos and Denzine (2005) found correlations between ToL performance and working memory measures. Although the correlation with verbal working memory was low, the correlation with visual working memory was moderate. However, in contrast, the auditory subsystem in this study was the best and only predictor of planning ability. These authors concluded that individuals can recruit distinct strategies to solve the problems that originate from the ToL and that task problems may differentially require either or both working memory subsystems. Based on our finding and expanding this idea beyond visual strategies, one may hypothesize that people create mental instructions or procedures for themselves to execute tasks (e.g., "first, green ball on the right; then, the blue one in the center..."). This understanding clarifies the role of auditory working memory in ToL problem solutions.

Recalling Unterreiner and Owen (2006), the role of working memory subsystems in ToL performance has not been clearly defined in the literature. Our results may help answer this question.

Despite the theoretical consistency, the predictive power of auditory working memory was rather modest (13.6%) and other abilities were not integrated into the regression model. Pulos and Denzine (2005) also noted the lack of correlations between ToL performance and other executive measures. They concluded that when online planning is allowed, the importance of executive abilities decreases, an explanation that can also be applied to the present results. Future studies should clarify this matter. For example, the instructions could clearly explain that the examinee should initially plan the entire action needed to solve the question and only afterward initiate the resolution. Measurements of planning and execution times can improve the measures gathered and better clarify the task requirements.

The level of EF development must also be emphasized. As previously mentioned, the differences between the present study and previous studies may be attributable to the ages of the participants. These differences, however, could be less important for the ToL task because studies have suggested mature performance on the ToL task at an earlier age, possibly at the beginning of adolescence. For example, Asato et al. (2006) noted that ToL performance was already mature in teenagers, and Matute et al. (2008) suggested that planning ability mainly develops up to 9 years of age. However, other executive abilities could still be developing such as different aspects of inhibitory control, which may explain why inhibitory control is a predictor of ToL performance in adults but not in children or adolescents when this parameter is measured using tests that require cognitive (not motor) control such as the SeGeT and Stroop Test.

Nonetheless, working memory, the sole ability that contributed to ToL performance in the present study, exhibited higher correlation or regression coefficients in other studies (e.g., Pulos & Denzine, 1995; Welsh et al., 1999) relative to ToL performance than those noted herein, possibly indicating that working memory in children and teenagers (Best & Miller, 2010; Davidson et al., 2006; Huizinga et al., 2006) plays a smaller role in ToL performance compared with the contribution of the same ability to adult performance.

The distinct versions, instructions, and measures of the ToL task including the revised (Welsh et al., 1999) and computerized (Asato et al., 2006) versions may influence the results and impose differential requirements on task solution. This makes comparisons between studies difficult and can explain the diversity of the findings (Batista et al., 2007; Unterreiner & Owen, 2006). Importantly, studies should use standardized versions and measures (e.g., Krikorian et al., 1994) to facilitate comparisons between results.

Resorting to one of these measures made available by Krikorian et al. (1994), our results support the

perception of the ToL task as a complex test (Miyake et al., 2000; Strauss et al., 2006) because the explanatory power of auditory working memory was quite moderate. Therefore, other abilities may influence performance, an idea supported by Lezak et al. (2004). Further studies are required to investigate this possibility.

### *Study limitations and suggestions for future research*

Limitations of the present study include the absence of adult participants and the low number of participants in each age group, precluding separate analyses according to age. Another limitation is the loss of subjects because of the large number of students with age and class discrepancies. Future studies should attempt to overcome these shortcomings by examining the requirements involved in these and other complex tests and including adult subjects in whom these abilities are more developed. Furthermore, future inquiries could refine our findings to investigate the procedures involved in the solution of complex tests in different age groups to determine whether the development of these abilities in parallel with advancing age can alter the strategies applied to solve the tasks.

Additionally, another limitation is the absence of reliability data for some of our tests, despite evidence of their validity. Calculating Spearman–Brown or Cronbach’s Alpha coefficients is not possible or suitable in some tests (e.g., TMT, CAiT, ToL, AWM test, and VWM test) because of the way they are performed or scored. Moreover, there is a problem related to test–retest reliability in EF tests. Strauss et al. (2006) indicated that test–retest reliability is not suitable for EF measures because in the second assessment the test is no longer a new situation for the participant who tends to apply the same strategy used previously (i.e., executive demand is reduced). For these authors, this may be the reason why test–retest reliability coefficients for EF tests are generally situated below what is considered acceptable in clinical practice. Studies are being conducted to increase available data on psychometric features for the tests used in this study, including reliability data, in an attempt to narrow this gap.

### *Final views*

The results of the present study revealed that different executive abilities can contribute in different ways to performance on complex tasks. In our study, only auditory working memory contributed to planning task performance after controlling for age. This result may reflect a typical pattern of performance in our sample age group where other abilities are not involved in ToL resolution because of their developmental course, with or without interactions with our specific executive measures. These findings have direct implications for the neuropsychological assessment of EF because knowing which aspects the test is specifically measuring is pertinent when using a complex test as a general measure

of EF. These results do not exhaust task requirements. Other executive and nonexecutive abilities are likely to be involved in performance, and future studies should delve into this matter and investigate differential performance patterns as a function of development.

## Acknowledgements

The authors acknowledge financial support by Fapesp (N.M. Dias) and CNPq (A.G. Seabra)

## References

- Asato, M., Sweeney, J., & Luna, B. (2006). Cognitive processes in the development of TOL performance. *Neuropsychologia*, *44*, 2259–2269.
- Assef, E. C. (2005). *Avaliação das funções executivas em crianças com Transtorno de Déficit de Atenção e Hiperatividade*. Master's dissertation, Psychology Post-graduate Program, São Francisco University, Itatiba, São Paulo, Brazil.
- Assef, E. C. S., Capovilla, A. G. S., & Capovilla, F. C. (2007). Computerized Stroop Test to assess selective attention in children with attention deficit hyperactivity disorder. *Spanish Journal of Psychology*, *10*, 33–40.
- Batista, A. X., Adda, C. C., Miotto, E. C., Lucia, M. C. S., & Scaff, M. (2007). Torre de Londres e Torre de Hanói: contribuições distintas para a avaliação do funcionamento executivo. *Jornal Brasileiro de Psiquiatria*, *56*, 134–139.
- Berg, W. K., Byrd, D. L., McNamara, J. P., & Case, K. (2010). Deconstructing the tower: Parameters and predictors of problem difficulty on the Tower of London task. *Brain and Cognition*, *72*, 472–482.
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641–1660.
- Best, J. R., Miller, P., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*, 327–336.
- Bodrova, E., & Leong, D. J. (2007). *Tools of the mind: The Vygotskian approach to early childhood education*. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.
- Capovilla, A. G. S., Dias, N. M., Trevisan, B. T., Montiel, J. M., & Andrade, A. M. (2008). Validity of neuropsychological instruments for attention assessment in Brazilian children. *Journal of the International Neuropsychological Society*, *14*, 7–8.
- Cuevas, R. R., & Alanis, B. T. (2008). Exploring the interrelation of executive functions during child development. *Journal of the International Neuropsychological Society*, *14*, 56.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037–2078.
- Dias, N. M. (2009). *Avaliação neuropsicológica das funções executivas: tendências desenvolvimentais e evidências de validade de instrumentos*. Master's dissertation. Developmental Disabilities Program, Mackenzie Presbyterian University, SP, Brazil. ([http://mx.mackenzie.com.br/tede/tde\\_busca/arquivo.php?codArquivo=1748](http://mx.mackenzie.com.br/tede/tde_busca/arquivo.php?codArquivo=1748); accessed January 5, 2012).
- Dias, N. M., Menezes, A., & Seabra, A. G. (2012). Age differences in executive functions within a sample of Brazilian children and adolescents. *Spanish Journal of Psychology*, in press.
- Durston, S., & Casey, B. J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia*, *44*, 2149–2157.
- Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: a comprehensive latent-variable analysis. *Acta Psychologica*, *129*, 32–48.
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from National Curriculum Assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*, 1–16.
- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2006). *Neurociência cognitiva a biologia da mente*. Porto Alegre: Artmed.
- Gil, R. (2002). *Neuropsicologia*. São Paulo: Editora Santos.
- Heitz, R. P., Unsworth, N., & Engle, R. W. (2004). Working memory capacity, attention control, and fluid intelligence. In O. Wilhelm, & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 61–77). Thousand Oaks, CA: Sage.
- Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017–2036.
- Kaller, C. P., Unterrainer, J. M., & Stahl, C. (2011). Assessing planning ability with the Tower of London task: psychometric properties of a structurally balanced problem set. *Psychological Assessment*, in press.
- Kane, M. J. (2005). Full frontal fluidity? Looking in on the neuroimaging of reasoning and intelligence. In O. Wilhelm, & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 141–163). Thousand Oaks, CA: Sage.
- Krikorian, R., Bartok, J., & Gay, N. (1994). Tower of London procedure: a standard method and developmental data. *Journal of Clinical and Experimental Neuropsychology*, *16*, 40–50.
- Lamm, C., Zelazo, P. D., & Lewis, M. D. (2006). Neural correlates of cognitive control in childhood and adolescence: disentangling the contributions of age and executive function. *Neuropsychologia*, *44*, 2139–2148.
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment*, 4<sup>th</sup> edition. New York: Oxford University Press.
- Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior*, *39*, 900–997.
- Luciana, M., Collins, P. F., Olson, E. A., & Schissel, A. M. (2009). Tower of London performance in healthy adolescents: the development of planning skills and associations with self-reported inattention and impulsivity. *Developmental Neuropsychology*, *34*, 461–475.
- Malloy-Diniz, L. F., Sedo, M., Fuentes, D., & Leite, W. B. (2008). Neuropsicologia das funções executivas. In D. Fuentes, L. Malloy-Diniz, C. H. Camargo, & R. M. Cosenza (Eds.), *Neuropsicologia: teoria e prática* (pp. 187–206). Porto Alegre: Artmed.
- Matute, E., Chamorro, Y., Inozemtseva, O., Barrios, O., Rosselli, M., & Ardila, A. (2008). Efecto de la edad en una tarea de planificación y organización ('Pirámide de México') en escolares. *Revista de Neurología*, *47*, 61–70.
- 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*, 49–100.
- Montiel, J. M., & Seabra, A. G. (2009a). Teste de Atenção por Cancelamento. In A. G. Seabra, & F. C. Capovilla (Eds.), *Teoria e pesquisa em avaliação neuropsicológica* (pp. 119–124). São Paulo: Memnon.
- Montiel, J., & Seabra, A. G. (2009b). Teste de Trilhas: parte B. In A. Seabra, & F. Capovilla (Eds.), *Teoria e pesquisa em avaliação neuropsicológica* (pp. 94–95). São Paulo: Memnon.
- Montiel, J., & Seabra, A. G. (2012a). Teste Atenção por Cancelamento. Em: A. G. Seabra & N. M. Dias (Orgs.), *Avaliação Neuropsicológica cognitiva: atenção e funções executivas*. Vol. 1 (pp. 57–56). São Paulo: Memnon.
- Montiel, J., & Seabra, A. G. (2012b). Teste de Trilhas – partes A e B. In: A. G. Seabra & N. M. Dias (Orgs.), *Avaliação Neuropsicológica cognitiva: atenção e funções executivas*. Vol. 1 (pp. 79–85). São Paulo: Memnon.
- Nigg, J. T. (2011). Is ADHD a disinhibitory disorder? *Psychological Bulletin*, *127*, 571–598.
- Primi, R. (2002). *Bateria Informatizada de Capacidades Cognitivas [Software]*. Itatiba, São Paulo: São Francisco University.
- Pulos, S., & Denzine, G. (2005). Individual differences in planning behavior and working memory: a study of the Tower of London. *Individual Differences Research*, *3*, 99–104.
- Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: a survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*, 33–65.
- Regard, M. (1991). Cognitive rigidity and flexibility: a neuropsychological study. In O. Spreen, & E. Strauss (Eds.), *A compendium of neuropsychological tests: administration, norms, and commentary*. New York: Oxford University Press.
- Seabra, A. G., Cozza, H., Capovilla, F. C., Macedo, E. C., & Dias, N. M. (2009). Avaliação de controle inibitório em crianças: Teste de Geração Semântica. In A. G. Seabra, & F. Capovilla (Eds.), *Teoria e pesquisa em avaliação neuropsicológica*, 2nd edition. São Paulo: Memnon.

- Seabra, A. G., Dias, N. M., Berberian, A. A., Assef, E. C. S., & Cozza, H. F. (2012). Teste da Torre de Londres. In A. G. Seabra & N. M. Dias (Orgs.), *Avaliação Neuropsicológica Cognitiva: atenção e funções executivas*. Vol. 1. (pp. 109-132). São Paulo: Memnon.
- Seabra, A. G., Dias, N. M., & Trevisan, B. T. (2010). Avaliação neuropsicológica das funções executivas na infância e na adolescência. In C. Hutz (Ed.), *Avaliação psicológica na infância e na adolescência*. Rio Grande do Sul: ArtMed.
- Seabra, A. G., Macedo, E. C., & Dias, N. M. (in press). *Teste de Stroop Computadorizado: Manual e CD ROM* (in press). São Paulo: Mackenzie Presbyterian University.
- Silva Filho, J. H. (2007). *Validade e normas do Wisconsin Card Sorting Test em adultos da região de Ribeirão Preto*. Doctoral thesis. Psychology Postgraduate Program, FFCLRP—Psychology and Education Department, São Paulo University, Ribeirão Preto, SP, Brazil.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary*, 3rd edition. New York: Oxford University Press.
- Sullivan, J. R., Riccio, C. A., & Castillo, C. L. (2009). Concurrent validity of the tower tasks as measures of executive function in adults: a meta-analysis. *Applied Neuropsychology*, *16*, 62–75.
- Thompson-Schill, S., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 15855–15860.
- Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: from neuropsychology to functional neuroimaging. *Journal of Physiology (Paris)*, *99*, 308–317.
- Unterrainer, J. M., Rahm, B., Kaller, C. P., Leonhart, R., Quiske, K., Hoppe-Seyler, K., ... Halsband, U. (2004). Planning abilities and the Tower of London: is this task measuring a discrete cognitive function? *Journal of Clinical and Experimental Neuropsychology*, *26*, 846–856.
- Welsh, M. C., Satterlee-Cartmell, T., & Stine, M. (1999). Towers of Hanoi and London: contribution of working memory and inhibition to performance. *Brain and Cognition*, *41*, 231–242.

