Psychology & Neuroscience

ISSN: 1984-3054

https://doi.org/10.1037/pne0000224

Cognitive and Socioeconomic Predictors of Stroop Performance in Children and Developmental Patterns According to Socioeconomic Status and ADHD Subtype

Vanessa Arán-Filippetti[1, 2, 3], María C. Richaud[1, 2, 3], Gabriela Krumm[1, 2, 3], and Waldina Raimondi[3]

Q35

[1] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental (CIIPME)

[2] Centro Interdisciplinario en Ciencias de la Salud y del Comportamiento (CIICSAC), Universidad Adventista del Plata

[3] Facultad de Humanidades, Educación y Ciencias Sociales, Entre Ríos, Universidad Adventista del Plata

Objectives: We conducted three empirical studies with the aim at (a) examining the cognitive predictors (i.e., working memory, inhibition, cognitive flexibility, reading, and intelligence) of each *Stroop Color and Word Test* (SCWT) condition (i.e., Word, Color, and Color–Word) and the convergent and divergent validity among measures, (b) examining the socioeconomic predictors of SCWT performance, further establishing normative values according to socioeconomic status (SES) and age, and (c) analyzing the distinctive patterns of performance according to SES and Attention Deficit and Hyperactivity Disorder (ADHD) subtype. ***Methods***: A large sample of typically developing (TD) children from Middle- (n = 779) and Low- (n = 129) SES and ADHD children (n = 44), inattentive versus combined subtype, was evaluated. Multivariate analysis of variances (MANOVAs), Pearson's correlations, and hierarchical and stepwise regressions analyses were performed. ***Results***: Study 1 results indicated that SCWT conditions are selectively associated with reading speed and executive functions (EFs), and that the former would not depend on child's IQ. Study 2 findings revealed distinct patterns of SCWT performance according to SES and selective associations between socioeconomic indicators and SCWT conditions, being maternal education and housing conditions the main predictors. Finally, Study 3 results revealed distinctive patterns of SCWT performance according to ADHD subtype, with no differences on the interference measure among groups. ***Conclusions***: Our findings support the validity of the SCWT as a measure of inhibition in TD children. However, when the pattern of SCWT performance is different from the typical expected one (i.e., Word score higher than Color score and this, in turn, higher than Color–Word score), the interference measure should be interpreted with caution but without disregarding the relevant and distinctive information provided by each SCWT condition.

Vanessa Arán-Filippetti https://orcid.org/0000-0002-0753-5089; María C. Richaud https://orcid.org/0000-0001-6408-4897; Gabriela Krumm https://orcid.org/0000-0002-2765-9062.

The authors are indebted to the schools and all the children who contributed to this study.

The authors declare no conflict of interest.

Normative data for Argentinean Spanish-Speaking children according to SES and Age.

Correspondence concerning this article should be addressed to Vanessa Arán-Filippetti, PhD, Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental (CIIPME-CONICET). Dirección: Teniente General Juan Domingo Perón 2158 C1040AAH. Buenos Aires, Argentina. Email: vanessaaranf@gmail.com

Public Significance Statements

To the authors' knowledge, there are neither studies establishing differentiated norms according to SES for children nor research examining the cognitive and socioeconomic predictors of SCWT performance and its convergent and divergent validity in children. Besides, no study has examined the SCWT sensitivity with ADHD Spanish-speaking children according to inattentive versus combined subtypes. Having knowledge of those factors that could influence SCWT performance, and further setting normative data adjusted for age and SES do become of great relevance for the neuropsychological assessment of TD children.

Keywords: SCWT, socioeconomic status, ADHD, executive functions, child neuropsychology

Supplemental materials: https://doi.org/10.1037/pne0000224.supp

The *Stroop Color and Word Test* (SCWT) (Golden, 1978) is one of the most commonly used measure of executive function (EF) in clinical and research settings. Its performance requires the ability to inhibit overlearned responses (Gruber et al., 2002) and cognitive flexibility (Golden & Golden, 2002), along with certain reading skills (León-Carrion et al., 2004; Martín et al., 2012). According to Golden (1978), those dimensions assessed by the SCWT that relate to cognitive flexibility and interference control play a central role in various interrelated cognitive processes.

Executive processes underlying SCWT performance, such as response inhibition and conflict resolution, are considered as mediated by the frontal lobe (Adleman et al., 2002). Indeed, Q1 neuroimaging studies using PET and fMRI techniques have found Stroop task-related activation in the anterior cingulate cortex (ACC) (Adleman et al., 2002; Carter et al., 1995; Pardo et al., 1990; van Veen & Carter, 2005), the lateral prefrontal cortex (Adleman et al., 2002), and the frontal polar cortex (Carter et al., 1995). Besides, SCWT performance also elicited activity in parietal brain regions (Adleman et al., 2002; van Veen & Carter, 2005) and in the left thalamus (van Veen & Carter, 2005). Consistently, studies in children have also found activation in the left lateral prefrontal cortex during SCWT performance (Schroeter et al., 2004). This dorsolateral prefrontal cortex activity increases with age, together with an improvement in behavioral performance (Adleman et al., 2002; Schroeter et al., 2004).

On the Cognitive Correlates of Stroop Performance in Children: Executive Functions, Intelligence, and Reading

Although the SCWT is considered an index of EFs (Golden & Golden, 2002), mainly used for the assessment of cognitive inhibition (Arán Filippetti & Richaud, 2017; Roy et al., 2018; Scarpina & Tagani, 2017), few studies have examined its convergent validity in children. Evidence of convergent validity [i.e., the extent to which different instruments capture a common construct (Carlson & Herdman, 2012)] is a "minimal and basic requirement for the validity" of any psychological measure; however, this parameter has been often assumed rather than examined directly (Fiske, 1971, p. 164, as cited Q2 in Duckworth & Kern, 2011). To date, research is mostly restricted to clinical (Bondi et al., 2002) and nonclinical (Johnson et al., 2003; Leverett et al., 2002) adult populations, and to our knowledge, only one psychometric study (see Rodríguez Barreto et al., 2016) has examined the SCWT convergent validity in children (though it also includes an adult sample). Instead, its validity as an EF measure has been mainly supported by confirmatory factor-analytic studies when exploring the dimensional nature of EF in children. Among the psychometric studies on the SCWT is that of Rodríguez Barreto et al. (2016) who analyzed the psychometric properties of the task in a wide sample of subjects between 7 and 80 years of age. The authors found moderate correlations between the three SCWT conditions and the Trail making test (TMT) (range

$r = .35$ to $r = .41$), and a weak correlation between the TMT and the interference score. In Q3 turn, when examining the structure of EF in children, factor analytic studies have consistently demonstrated that SCWT performance relates with other measures of EF, including inhibition, working memory, and cognitive flexibility (see e.g., Arán Filippetti & Richaud, 2017; Wu et al., 2011). Though SCWT performance has also been related to intelligence in children (Arán Filippetti, Krumm et al., 2015; Friedman, 1971; Roy et al., 2018), it is considered a task not based on subjects' intelligence (Golden & Golden, 2002). Finally, considering that SCWT requires good reading skills (Golden & Golden, 2002, León-Carrion et al., 2004; Martín et al., 2012) beyond executive processes (Protopapas et al., 2007), and that the former skills may even affect Stroop construct validity (Cox et al., 1997), it becomes important to contemplate reading speed when examining the SCWT cognitive predictors or its convergent validity.

Age, Sex, and Socioeconomic Status Effects on Inhibition: Evidence From the Stroop Paradigm

Stroop tasks, such as the SCWT (Golden, 1978) or Stroop-like day–night test for children under 7 years of age (Gerstadt et al., 1994), are considered classical tasks to assess response inhibition throughout development. When working with preschoolers, it has been observed that while 3½-years-olds experience some difficulties during Stroop-like tasks, older ones easily overcome them (Gerstadt et al., 1994). Other neurodevelopmental studies in school-aged children have indicated that response inhibition, as measured by the SCWT, develops over the course of childhood from 6 years to the end of adolescence. Among these studies is that of Roy et al. (2018) who found, when working with French-speaking children, a continuous increase in the Word (W) and Color naming (C) conditions performance between 7 and 12 years, and a reduction of the Stroop effect with age. Similar results were reported by Martín et al. (2012) for Spanish children aged 6–12 years who encountered a linear increase in performance on all three conditions of the SCWT (i.e., Word, Color, and Color–Word), and a quadratic relationship for the interference index. Another developmental study, comprising children and adolescences from 6 to 17 years old, found that SCWT performance increases with age during both childhood and adolescence (León-Carrion et al., 2004). Finally, when establishing normative data for the SCWT in Spanish-speaking pediatric populations, Rivera et al. (2017) consistently found that Stroop scores increased linearly with age in Spain and different Latin American countries. Although age clearly influences SCWT performance, studies analyzing sex effects have yielded contradictory results (see e.g., Rivera et al., 2017). Although some studies conducted with children have not found differences between boys and girls (Armengol, 2002; León-Carrion et al., 2004; Roy et al., 2018), a more recent research has reported sex effects on SCWT performance in children from different Spanish-speaking countries (Rivera et al., 2017).

Finally, earlier studies have found that family socioeconomic status (SES) is associated with SCWT performance, both in English- (Sarsour et al., 2011) and Spanish-speaking pediatric populations (Arán Filippetti & Richaud de Minzi, 2012). Consistently, when working with a large sample of Spanish-speaking children, Rivera et al. (2017) found an association between parental education and SCWT performance, with children from parents with higher education obtaining better scores than children from parents with lower education. Considering that earlier neurodevelopmental studies have even showed some differences on SCWT scores among Middle-SES Spanish-speaking children from different countries (see e.g., León-Carrion et al., 2004; Martín et al., 2012, Rivera et al., 2017; Rodríguez Barreto et al., 2016), it is important to establish normative data according to SES, language, and country of origin (Rivera et al., 2017), as variables associated with the culture or educational system could influence cognitive task performance.

Stroop Paradigm in the Neuropsychological Assessment: Implications With ADHD Children

The Stroop represents one of the most used paradigms to measure frontal lobe functioning (Demakis, 2004; Homack & Riccio, 2004) for both clinical and research purposes (Scarpina & Tagini, 2017). In the child neuropsychology field, its importance has been documented in several earlier studies, which address it as a useful tool for the neuropsychological assessment of children with Attention Deficit and Hyperactivity Disorder (ADHD) (Barkley et al., 1992; López-Villalobos et al., 2010), fetal alcohol syndrome (Mattson et al., 2011), epilepsy (Chevalier et al., 2000), and Turner's syndrome (Temple et al., 1996), among others (see Homack & Riccio, 2004, for a review). In addition, its clinical utility has been proven for the study of attention and executive functioning in children and adolescents with learning disabilities (Golden & Golden, 2002; Lazarus et al., 1984), disruptive behavior (Lavoie & Charlebois, 1994), and mood disorders (Cataldo et al., 2005).

Because ADHD is associated with significant weakness in core EF (Barkley, 1997; Craig et al., 2016; Willcutt et al., 2005), the assessment of these higher order cognitive processes in children with this disorder has received increasing interest in recent years. ADHD is a highly heritable neurodevelopmental condition affecting around 5% of school-aged children (Demontis et al., 2019). It is characterized by an age-inappropriate pattern of hyperactivity/impulsivity and inattention that affects development and interferes with academic and social activities (American Psychiatric Association [APA], 2013). According to the presence of symptoms of inattention and/or hyperactivity–impulsivity, three specific presentations can be distinguished, that is, predominantly Inattentive, predominantly Hyperactive–Impulsive, and Combined presentation (APA, 2013), being the Inattentive (ADHD/I) (Willcutt, 2012) and the Combined subtype (ADHD/C) (Barkley, 1998; Elosúa et al., 2017) the most prevalent presentations. As formerly mentioned, among the most widely used tests for ADHD neuropsychological assessment is the SCWT (Barkley et al., 1992; Borella et al., 2013; Homack & Riccio, 2004) for both studying executive functioning in children with ADHD (see e.g., Elosúa et al., 2017; Nigg et al., 2002) and monitoring clinical pharmacological responses (Langleben et al., 2006). Research has also demonstrated that the Stroop Color–Word measure is a good predictor of hyperactive/impulsive ADHD symptoms mainly due to its cognitive demands (i.e., naming and processing speed and response inhibition) (Rucklidge & Tannock, 2002).

The Present Study

The SCWT is one of the most commonly used measures for assessing inhibition in pediatric clinical and research settings. This has generated increasing interest in the study of different aspects related to the instrument, such as its sensitivity and specificity (see Homack & Riccio, 2004, for a review), its psychometric properties (Rodríguez Barreto et al., 2016), and the development of normative data adjusted for age and sex in children from different countries (see e.g., León-Carrion et al., 2004; Martín et al., 2012; Rivera et al., 2017; Rodríguez Barreto et al., 2016). Despite considerable research in the field, some issues remain unanswered.

First, though theoretically the SCWT measures reading skills and executive processes (Golden & Golden, 2002; León-Carrion et al., 2004; Protopapas et al., 2007; Rivera et al., 2017), to our knowledge, no studies have examined its convergent validity in children and the contribution of EF to each SCWT condition, controlling for the effect of age, reading speed, and comprehension. Evidence on those cognitive processes that differentially influence SCWT measures contributes to the understanding of the patterns of performance in pediatric population and provides support for SCWT convergent and divergent validity. For this reason, Study 1 aims at examining the contribution of EF to each SCWT condition (i.e., Word, Color, and Color–Word), controlling for the effect of age, reading skills, and IQ, further studying the convergent and divergent validity among measures. Based upon previous theoretical and empirical evidence, we first hypothesized that SCWT performance is related to EF and reading skills. However, as it is assumed that each SCWT condition taps distinct cognitive abilities (Golden & Golden, 2002; Rivera et al., 2017), we expected reading fluency to mainly predict Word reading while EFs mainly predict the Incongruent (i.e., Color–Word) and Color naming conditions. In addition, our correlation analysis results will also provide some support for the convergent and divergent validity of the SCWT. Convergent validity is normally evaluated by means of the magnitude of the zero-order correlations between the target measure and other closely related tasks (Carlson & Herdman, 2012), showing moderate to high correlations evidence of convergent validity (Gregory, 2007) and low correlations evidence of divergent validity (Cohen et al., 2010). Therefore, moderate to high correlations between the SCWT and other putative measures of EF would provide evidence for convergent validity, whereas low coefficients with measures of intelligence would provide support for SCWT divergent validity.

Second, although previous studies have established normative data for the SCWT for children from different countries, to our knowledge, there are neither studies establishing norms for Spanish-speaking children according to SES nor research examining the socioeconomic predictors of each SCWT condition. As SES, language, and country-of-origin may affect SCWT performance (Rivera et al., 2017), Study 2 aims at establishing normative data for Argentinean children according to SES, further examining the socioeconomic predictors [i.e., family head profession (FHP), maternal education level (MEL), main source of family income (MSFI), and housing conditions (HCs)] of each SCWT condition. Considering that SES is strongly associated with diverse reading measures (see e.g., Aikens & Barbarin, 2008; Noble et al., 2006), we first hypothesize that there is an effect of SES on all the three SCWT conditions, mainly on Word reading. In turn, we assumed that the Low-SES group would show less interference due to their lower-than-expected Word score. Finally, considering that parents' education level has proven to be an important predictor of children's SCWT performance (see Rivera et al., 2017), we hypothesized that among the socioeconomic indicators under analysis, MEL would be the main predictor of SCWT conditions.

Finally, although the sensitivity of the SCWT has been examined with ADHD children from English-speaking countries (see Homack & Riccio, 2004 for a review), to our knowledge, only two studies has reported its sensitivity with ADHD Spanish-speaking children (Elosúa et al., 2017; López-Villalobos et al., 2010) but from a dissimilar culture to that of the present study and without analyzing the pattern of performance according to inattentive versus combined subtypes. This analysis does become important, because althoughe there would be no differences on SCWT performance between subtypes (see e.g., López-Villalobos et al., 2010), distinctions could indeed arise when contrasting performance of each subtype to that of typically developing (TD) children (see e.g., Houghton et al., 1999). In addition, earlier studies regarding SCWT performance in ADHD have yielded mixed results, with conflicting assumptions regarding their patterns of performance (Bará-Jiménez et al., 2003; Borella et al., 2013; Elosúa et al., 2017). Therefore, further research is needed—specially proceeding from different countries—when considering that differences have been observed on SCWT normative data according to cultural variables (see e.g., Armengol, 2002; Rivera et al., 2017). For these reasons, Study 3 aimed at exploring the sensitivity of the SCWT with Spanish-speaking children with ADHD inattentive and combined subtypes. We hypothesized that, although there would not exist differences on SCWT performance between ADHD subtypes, differences would be indeed observed when comparing the performance of each of the ADHD groups (inattentive vs. combined) to that of the TD control.

In summary, the purposes of the current investigation were (a) to analyze the cognitive predictors of each SCWT condition controlling for the effect of age, reading speed and IQ, and the convergent and divergent validity among measures (Study 1); (b) to examine the effect of SES and age on SCWT performance, further establishing normative values for Argentine children from 7 to 12 years of age (Study 2); and (c) to explore the sensitivity of the SCWT with children with ADHD predominantly inattentive and combined subtypes (Study 3).

Study 1 (S1)

Study 1 analyzed the cognitive predictors of each SCWT condition controlling for the effect of age, reading speed and IQ, and the convergent and divergent validity among measures.

Method

Participants

Study 1 follows a cross-sectional, correlational design. The sampling comprised 118 Spanish-speaking children (54 girls and 64 boys) from 8 to 12 years of age (M = 10.03, SD = 1.42), of middle SES families living in Argentina. Parents' educational level was categorized by means of a 5-point scale, as follows: (a) Primary level, (b) Secondary level, (c) More education than secondary school, but less than a university degree, (d) University degree, and (f) Master's degree or higher education. The average of fathers' educational level was 3.00 (0.78), whereas the average of mothers' educational level was 2.64 (0.73). The Department of Education proposes a socioeconomic coefficient that is based on family income by means of a scale that goes from *very good* to *deficient* (source: Computer System of the Department of Education of the Province of Santa Fe, Argentina). The socioeconomic coefficient of the schools was *good*. According to this classification, good refers to children whose parents have remunerated jobs and fixed incomes. From the data obtained in the schools, inclusion criteria were as follows: (a) children with no known history of neurological or psychiatric treatment; (b) who attend school regularly; and (c) without school repetition. Prior to the administration of cognitive tasks, the Kaufman Brief Intelligence Test (K-BIT test) (Kaufman & Kaufman, 2000) was used to establish that children had a performance within the expected range for their age group. Intellectual functioning was found to be within the expected range for children (M = 91.58, SD = 10.36).

Instruments

Kaufman Brief Intelligence Test (Kaufman & Kaufman, 2000)

It offers a measure of crystallized (Gc) and fluid (Gf) intelligence and consists of two subtests: (a) vocabulary (verbal/crystallized/knowledge), which includes part A to assess expressive vocabulary and part B to assess definitions and (b) matrices (manipulative/fluid/mental processing). The internal consistency analyzed with the two halves method for the Vocabulary and Matrices subtests are .98 and .97, respectively. In turn, the test–retest stability coefficients for these subtasks are .94 and .86, respectively, (Kaufman & Kaufman, 2000).

Stroop Color and Word Test (Golden, 2001)

It provides a measure of interference control and response inhibition. It consists of three conditions: (a) Word, (b) Color, and (c) Color–Word. In the first page (i.e., Word condition), participants must read aloud the words "red," "green," and "blue" which are printed in black ink and randomly organized into five columns. In the second page (i.e., Color naming condition), participants are asked to name the color of each element
Q4 (XXXX) which are also randomly organized and printed in blue, red, or green ink. In the last page (i.e., Color–Word condition), participants have to name the color of the ink of words that are printed in nonmatching ink. Thus, this last condition offers a measure of interference, because subjects must inhibit the reading of the word in order to name the color of the ink. The dependent variables consisted of the number of items properly named in 45 s in the Word, Color, and Color–Word conditions and the interference score, which was then calculated according to the scoring method proposed by Golden (2001): $IS = WC - [(W \times C)/(W + C)]$. In addition to considering these scores, the manual guidelines suggest analyzing the pattern of performance across conditions, as it has been indicated that from 7 years on, the CW score should be lower than the C score and this is in turn, lower than the W score (i.e., Word > Color > Color–Word) (Golden, 2001). Examining the pattern of Stroop performance (i.e., relationship between the three Stroop scores) has been useful for the diagnosis of brain dysfunction and psychiatric disorders (see Golden, 2001). The reliability of the SCWT using the test–retest method for the Word, Color, and Color–Word variables is .86, .82, and .73 respectively (Golden, 1975, cited in Golden, 2001). Q5

NEPSY Knock-Tap Battery (Korkman et al., 1998) Q6

It assesses self-regulation and inhibition capacity. Specifically, the task requires suppressing one motor action to produce another conflicting motor response. In the first part of the task (i.e., items 1–15), when the examiner taps on the table, the participant must knock, and when the examiner knocks on the table, the participant must tap. In the second section (i.e., items 16–30), the participant must tap with the side of his/her fist when the examiner knocks with the knuckles and must knocks with the knuckles when the examiner taps with the side of his/her fist. However, the participant does not must respond when the examiner taps on the table. This task has been used as a measure of inhibitory capacity in English- (Pratt et al., 2014), French- (Mainville et al., 2015), and Spanish-speaking (Aguilar-Alonso & Moreno-González, 2012; Arán Filippetti & Krumm, 2020) children.

d2 Attention Test (Brickenkamp, 2004) Q7

It offers a measure of processing speed, selective attention, and mental focus, through selective searching for relevant stimuli. It is composed of 658 items, ordered into 14 lines, each containing 47 letters. The target stimuli are the letters “d” and “p” with one or two dashes. The participant must search across the lines in order to identify and cross all “d’s” with two dashes that can be located either above or below the letter, during 20 s per line. Earlier research has found that the d2 task loaded with other well-known measure of inhibition as the SCWT (Brickenkamp & Zillmer, 1998). It demonstrates a high internal consistency ($r > .90$), regardless of the statistics (two-halves and odd-even methods) and the sample used (Brickenkamp, 2004). We used the TN–E variable (i.e., total number of items processed minus errors made (omissions + commissions), which offers a measure of attentional and inhibitory control (Brickenkamp, 2004).

WISC-IV Working Memory Index (Wechsler Intelligence Scale for Children—Fourth Edition) (Wechsler, 2010)

It provides a working memory measure. It includes two main subtests: (a) Digits, which provides a measure of immediate retention when evaluated with Digits Forward (DF), and maintenance and manipulation of information when using Digits Backward (DB) and (b) Letters and Numbers sequencing (LNS). In this last case, participants are instructed to recall a series of numbers and letters read by the examiner, while ordering the numbers from lowest to highest and the letters in alphabetical order.

Q8 Wisconsin Card Sorting Test (Heaton et al., 1993)

It provides a measure of EF, particularly cognitive flexibility and categorization ability. First, four key cards (i.e., one red triangle, two green stars, three yellow crosses, and four blue circles) are placed in front of the participants. Then, they receive a stack of 128 additional response cards in order to match each card to one of the key cards. Examiners inform participants about whether their answers are right or wrong, as they matched the different cards; however, categories are not provided to participants while classifying.

Trail Making Test (Reitan & Wolfson, 1992)

It comprises two subtests, Part A and Part B. It allows obtaining a measure of attention, visual search, and mental flexibility (Spreen & Strauss, 1998). It consists of two forms, Part A (TMT-A) and Part B (TMT-B). In the TMT-A, participants must draw lines connecting 15 encircled numbers in order, which are randomly dispersed on a page. In the TMT-B, participants have to alternate between numbers and letters (e.g., 1 with A; 2 with B, and so on). For both forms, A and B, the time and number of errors are recorded. The test–retest reliability coefficient ranges from .60 to .90 (Spreen & Strauss, 1998).

Semantic Verbal Fluency and Phonological Verbal Fluency (FAS Fluency Test; Benton & Hamsher, 1989) Q9

Participants are asked to say as many words as possible within 60 s, belonging to a certain category (i.e., fruits and animals) and starting with a particular letter (i.e., F, A, and S). VF tasks have developmental norms for Spanish-speaking children (Arán Filippetti & Allegri, 2011).

Five-Point Test (Regard et al., 1982)

It offers a measure of nonverbal fluency or spontaneous flexibility. It consists of a page that contains 35 identical squares organized into 5 columns and 7 rows. Each square includes five symmetrically organized dots. Participants are requested to make as several diverse designs as possible in a 3-min period by connecting two or more dots with straight lines. The test–retest stability coefficient for the number of unique designs is .77 (Tucha et al., 2012).

Porteus Maze Test (Porteus, 2006)

It consists of twelve mazes of increasing complexity that enable to assess the ability to concrete a plan. In each maze, participants must trace the way from an initial point to an exit and must avoid blind alleys and dead ends, with no backtracking allowed. Correct performance is an indicator of adequate planning and impulsivity control.

Pyramid of México, Subtest of ENI (Neuropsychological Assessment of Children) Battery (Matute et al., 2007)

ENI is a battery that allows a neuropsychological comprehensive evaluation in children between the ages of 5 and 16 years (Rosselli-Cock et al., 2004). We used the Pyramid of México subtest that offers a measure of planning and organization. Participants must use three wooden blocks of different colors (red, green, and white) and sizes (small, medium, and big) under certain restrictions, in order to build a series of designs that are offered as a model.

Reading Comprehension Subtest of ENI (Neuropsychological Assessment of Children) Battery (Matute et al., 2007)

We used the reading comprehension subtest of the ENI battery, where children must mentally read a text to then answer questions related to its content. It allows to assess comprehension and reading speed. The present study included both scores.

Ethics Procedure

An interview was requested to school principals in order to clarify the research characteristics. Next, children' parents or legal guardians were sent a note requesting authorization. It was also explained that children's participation was voluntary and anonymous. Finally, parents and legal guardians' consent was obtained prior to assessment.

Statistics Procedures

Pearson's correlation coefficients were obtained to analyze the associations among the SCWT, age, IQ, EF, and reading skills. To adjust for multiple comparisons, we used Bonferroni correction ($\alpha < .01$). To examine the contribution of EF to each SCWT condition, controlling for age and reading speed, hierarchical regression analysis was conducted. The variance inflation factor (VIF) measure was used to detect the presence of collinearity. Due to the sample size, the entry criterion was set at the $p \leq .01$ level. All analyses were performed with SPSS for Windows, Version 20.0.

Results

Pattern of SCWT Performance for the Total Sample of Children

Results show that scores and the pattern of performance of the TD children are consistent with those expected for children older than 7 years, being Word score ($M = 71.29$, $SD = 16.73$) higher than the Color score ($M = 47.22$, $SD = 9.52$), which in turn is higher than Color–Word score ($M = 25.11$, $SD = 6.41$). Besides, the differences among the correctly read items between the Color and Color–Word condition and between Color and Word variables are 24 and 22 items, respectively.

Relationship Between the SCWT and Performance-Based Measures of EF, Age, Reading Skills, and IQ

Significant correlations were found between the performance in the SCWT conditions and all EF measures under analysis. Specifically, performance on the Word, Color, and Color–Word conditions was associated with better performance on tasks that value inhibition (knock and tap and d2) (range from $r = .356$ to $r = .535$), WM (range from $r = .304$ to $r = .608$), spontaneous (VF and phonological verbal fluency [PVF]) (range from $r = .364$ to $r = .556$) and reactive (WCST and TMT) (range from $r = .278$ to $r = .641$) cognitive flexibility and planning (Porteus Maze and pyramid of México) (range from $r = .269$ to $r = .462$). However, the interference score was only associated with the knock and tap ($r = -.299$, $p < .01$) and letter–number sequencing of the WISC-IV ($r = -.314$, $p < .01$). Age was associated with the Word ($r = .675$), Color ($r = .629$), and Color–Word ($r = .533$) conditions, but not with IS ($r = -.110$). Besides, the Word, Color, and Color–Word conditions and the IS were related to reading speed (range from $r = .256$ to $r = .611$). No significant correlations were found between SCWT conditions, and neither with reading comprehension (Word = .190; Color = .165; Color–Word = 079; IS = −.114) nor with general intelligence (Word = .077; Color = .170; Color–Word = 138; IS = .009) (see Table 1).

Hierarchical Regression Analysis

To explore the contribution of EF to SCWT performance, controlling for the effect of age and reading speed, hierarchical regression analysis was performed. None of the variation inflation factors (VIFs) indicated multicollinearity, even with a strict cutoff (VIF > 4). The first model that Q10

Table 1
Pearson's Correlations Coefficients and Significance Levels Between SCWT Conditions, EFs, IQ, and Reading Skills

Variables	Word	Color	Color–Word	Interference
Age	**.675****	**.629****	**.533****	−.110
Intelligence				
Gc	.041	.129	.093	.003
Gf	.094	.147	.135	.012
General IQ	.077	.170	.138	.009
Executive Functions				
KT	**.356****	**.368****	.089	**−.299****
TN−E d2	**.429****	**.529****	**.535****	.078
DF WISC-IV	**.321****	**.304****	**.311****	−.003
DB WISC-IV	**.502****	**.547****	**.390****	−.142
LNS WISC-IV	**.608****	**.550****	**.286****	**−.314****
TMT-A	**−.641****	**−.595****	**−.416****	.215*
TMT-B	**−.460****	**−.487****	**−.413****	.057
CC-WCST	.178	**.278****	.206*	−.023
SVF	**.504****	**.475****	**.404****	−.080
PVF	**.429****	**.438****	**.409****	−.017
FPT	**.556****	**.475****	**.364****	−.156
Porteus mazes	**.314****	**.462****	**.305****	−.091
Pyramid of Mexico	.198*	**.269****	.076	−.184*
Reading				
Reading speed	**.611****	.457**	**.291****	**−.256****
Reading comprehension	.190*	.165	.079	−.114

Note. Gc = crystalized intelligence; Gf = fluid intelligence; KT = Knock and Tap; TN−E d2 = Attentional and Inhibitory control d2; DF = Digit Forward of the WISC-IV; DB = Digit Backward of the WISC-IV; LNS = Letter–number sequencing of the WISC-IV; TMT-A = Trail making test form A; TMT-B = Trail making test form B; CC-WCST = Number of complete categories of the WSCT; SVF = Semantic Verbal Fluency; PVF = Phonological Verbal Fluency; FPT = Five-Point Test; SCWT = Stroop Color and Word Test; EF = executive function.
In bold ink the significant correlations are indicated according to the Bonferroni correction ($\alpha < .01$).
* $p < .05$. ** $p < .01$.

analyzes the contribution of EF to W measure included the following blocks: (a) age and reading speed and (b) EF that showed moderate or a fair correlation (i.e., greater than .40, see Akoglu, 2018) with the Word condition. The model explained 67% of the variance. Specifically, age and reading speed accounted for 57% of the variance, whereas EF explained 10% over the variance explained by age and reading fluency. When including EF in block 2, age ($\beta = .328$, $p < .001$) and reading speed ($\beta = .237$, $p = .001$) continue to explain performance on Word condition, although among the attention and executive measures included in the analysis, only attentional control and processing speed (TMT-A) ($\beta = -.276$, $p < .001$) predicted its performance. The second model that analyzes the contribution of EF to the Color condition included the following blocks: (a) age and reading speed and (b) EF that showed moderate or fair correlation with Color variable. The model explained 60% of the variance. Specifically, age and reading speed explained 43% of the variance while EF accounted for 17% over the variance explained by age and reading speed. When including EF in block 2, neither age ($p = .156$) nor reading speed ($p = .317$) explained Color performance, although of the attention and executive measures included in the analysis, only TMT-A ($\beta = -.252$, $p = .003$) and Porteus task (i.e., impulsivity control) ($\beta = .189$, $p = .009$) predicted this condition performance. Finally, the third model that analyzes the contribution of EF to the Color–Word condition included the following blocks: (a) age and reading speed and (b) Word and Color conditions and EF that showed moderate or fair correlation with the Color–Word condition. The model explained 49% of variance. Specifically, age explained 29% of the variance while EF accounted for 20% over the variance explained by age. When including EFs in block 2, age ($p = .260$) did not predict performance on the Color–Word condition, whereas of the attention and executive measures included in the analysis, only attentional and inhibitory control (TN–E d2) (β = .207, p = .029) and the Color condition (β = .449, p < .001) predicted performance on the Color–Word variable (see Table 2).

Conclusion S1

Study 1 results showed significant correlations between EF and the Word, Color, and Color–Word conditions, except with the interference score that only had a significant relationship with Knock and tap and with LNS. These findings are in line with those of Rodríguez Barreto et al. (2016) who also observed a stronger correlation between the three SCWT conditions and the TMT-B (range r = .35 to r = .41), than between the interference score and the EF task. Consistently, an earlier study in school-aged children also reported that interference score was not related to other EF measures (Wu et al., 2011). Overall, these findings suggest that the

Table 2

Complete Results of the Hierarchical Regression Analysis of Word, Color, and Color–Word Conditions Performance in TD Children

Dependent	Predictor	R^2	ΔR^2	β	p
Word	Block 1				
	Age	.57	.57	.498	<.001
	Reading speed			.377	<.001
	Block 2				
	Age	.67	.10	.328	<.001
	Reading speed			.237	.001
	TN−E d2			−.088	*ns*
	DB WISC-IV			.032	*ns*
	LNS WISC-IV			.062	*ns*
	TMT-A			−.276	<.001
	TMT-B			.035	*ns*
	Verbal fluency			.124	*ns*
	FPT			.107	*ns*
Color	Block 1				
	Age	.43	.43	.532	<.001
	Reading speed			.206	.011
	Block 2				
	Age	.60	.17	.140	*ns*
	Reading speed			.079	*ns*
	TN−E d2			.141	*ns*
	DB WISC-IV			.156	*ns*
	LNS WISC-IV			−.013	*ns*
	TMT-A			−.252	.003
	TMT-B			−.018	*ns*
	Verbal fluency			.095	*ns*
	FPT			.049	*ns*
	Porteus Mazes			.189	.009
Color–Word	Block 1				
	Age	.29	.29	.509	<.001
	Reading speed			.052	*ns*
	Block 2				
	Age	.49	.20	.125	*ns*
	Reading speed			−.044	*ns*
	Stroop W			−.128	*ns*
	Stroop C			.449	<.001
	TN−E d2			.207	.029
	TMT-A			−.004	*ns*
	TMT-B			−.076	*ns*
	Verbal fluency			.138	*ns*

Note. TN−E d2 = Attentional and Inhibitory control d2; DB = Digit Backward of the WISC-IV; LNS = Letter–number sequencing of the WISC-IV; TMT-A = Trail making test form A; TMT-B = Trail making test form B; FPT = Five-Point Test; TD = typically developing.

three SCWT conditions would be better indicators of EF and attention skills than the IS. However, as expected, correlations coefficients were in the low-to-moderate range demonstrating, as literature proposes, the multidimensional nature of the EF construct both in children (Arán Filippetti, 2013; Lehto et al., 2003) and adult populations (Miyake et al., 2000). As regards IQ and reading skills, our findings indicate that SCWT conditions and the IS were correlated with reading speed but not with reading comprehension or intelligence. Previous studies in children have also reported that Stroop task performance would mainly demand reading speed (during its word reading condition), but neither reading comprehension (Protopapas et al., 2007) nor intelligence (Golden & Golden, 2002; Protopapas et al., 2007).

When examining the contribution of EF to each SCWT condition, after controlling for age and reading speed, we found that SCWT measures impose different demands on attention and executive processes. Specifically, processing speed and visual search (TMT-A) predict performance on the Word condition, whereas processing speed and impulsivity control (Porteus mazes) predict performance on the Color component. In turn, only attentional and inhibitory control predict the performance on the Color–Word condition. It is further noticed that when including attention and executive measures in block 2, age and reading speed predicted performance only on the Word condition, but not on the Color and Color–Word ones. In this latter case, performance was also influenced by naming speed (as valued with the Color condition), but not by Word reading. This suggest that color naming imposes greater demands on executive processes than word naming, and that performance on the Color–Word condition would mainly depend on color naming and attentional and inhibitory control. Together, our findings support the validity of the SCWT as a measure of inhibition in children and suggest that each SCWT condition would provide relevant and distinctive information on attention and executive processes and reading speed.

Study 2 (S2)

Study 2 aimed at examining the effect of SES and age on SCWT performance further establishing normative values for Argentinean children from 7 to 12 years of age.

Method

Participants

The intentional sample consisted of 730 Argentine children aged 7–12 years from several regions of the country (i.e., Buenos Aires, Santa Fe, and Entre Ríos). The sample was subdivided into two groups considering school's characteristics (socioeconomic coefficient) and sample's neighborhood, that is, (a) Low SES (low-SES) group: Consisting of 129 boys, of both sexes (56.6% girls) aged between 7 and 12 years, attending an urban-marginal school and residing in peripheral neighborhoods. Based on family's income, the school's socioeconomic coefficient was deficient (Computer System of the Ministry of Education of the Province of Santa Fe, Argentina). According to this classification, *deficient* refers to those families that do not have salaried jobs or fixed incomes. Most parents in this category are unemployed or unqualified workers (e.g., laboring as street vendors). The neighborhoods where these families resided had a high concentration of low-income populations with diverse housing needs. Public services (i.e., sewer, water supply network, and natural gas) were not provided. (b) Medium SES (middle-SES) group consisting of 601 boys and girls (49.8% girls) from 7 to 12 years old attending urban schools and residing in middle class neighborhoods. With the aim at establishing normative data, we included our S1 sample of 118 children within this group. The socioeconomic coefficient of the schools (obtained, depending on the region, from the Computer System of the Department of Education of Santa Fe or from the school institutions) was good. From information obtained from the educational establishments, the inclusion criteria were as follows: (a) children without a neurological or psychiatric history; (b) who attend school regularly; and (c) without school repetition. Prior to the administration of cognitive tasks, the K-BIT test (Kaufman & Kaufman, 2000) was used to establish that children had a performance within the expected range for their age group. Intellectual functioning was found to be within the expected range for children ($M = 92.88$, $SD = 10.07$). To analyze the effect of age, the sample was subdivided within each group (middle-SES and low-SES) into three subgroups: Group 1: children aged 7 and 8 years; Group 2: children 9 and 10 years old, and Group 3: children aged 11 and 12 years. Finally, Graffar's modified scale (Méndez-Castellano & de Méndez, 1994) was used to examine SES predictors of SCWT performance. This scale offers a measure of four socioeconomic indicators; that is, FHP, MEL, MSFI, and HCs. For every variable, higher scores corresponded with higher poverty.

Instruments

Stroop Color and Word Test (Golden, 2001)

See S1 description.

Kaufman Brief Intelligence Test (Kaufman & Kaufman, 2000)

See S1 description.

Ethic Procedure

At first, contact was made with school principals in order to request authorization to perform the investigation. Then, children's parents or legal guardians received a note explaining the work's objectives and the task to be carried out. It was further clarified that participation was voluntary and anonymous. Finally, their permission and written consent was requested and obtained before assessment.

Statistics Procedure

Bifactorial multivariate analysis of variance (MANOVA) was used to analyze SES (low-SES vs. middle-SES) and age effects on SCWT performance. SES (low-SES and middle-SES) and age group (7–8, 9–10 and 11–12 years) were entered as between-subjects factors and the four SCWT scores (W, C, CW, and IS) as dependent variables. For multiple comparisons, Bonferroni correction was used ($p < .01$). To examine which socioeconomic indicator predicts performance on each SCWT condition, stepwise regression analysis, using the stepwise method (entry criterion $p < .05$, removal criterion $p > .10$), was used. To detect the presence of collinearity, the VIF measure was used. All analyses were performed with SPSS for Windows, Version 20.0.

Results

IQ and Sex Effects on SCWT Performance

There was not IQ or sex effects on SCWT performance (W, C, CW, and IS) neither in the Middle-SES group (IQ Hotelling's $F(12, 1766) = 1.534$; $p = .105$, $\eta p^2 = .010$, Sex Hotelling's $F(4, 590) = 1.027$; $p = .392$, $\eta p^2 = .007$) nor in the Low-SES group (IQ Hotelling's $F(4, 122) = .423$; $p = .792$, $\eta p^2 = .014$, Sex Hotelling's $F(4, 122) = .112$; $p = .978$, $\eta p^2 = .014$).

Age and SES Effects on SCWT Performance

A significant main effect of age, Hotelling's $F(8, 1440) = 26.724$; $p < .001$, $\eta p^2 = .13$, SES, Hotelling's $F(4, 721) = 73.295$; $p < .001$, $\eta p^2 = .29$, and for the interactions age × SES, Hotelling's $F(8, 1440) = 1.953$; $p = .049$, $\eta p^2 = .01$ was found. The effect of age was observed for the Word, $F(2, 724) = 93.784$; $p < .001$, $\eta p^2 = .21$, Color, $F(2, 724) = 51.082$; $p < .001$, $\eta p^2 = .12$, and Color–Word conditions, $F(2, 724) = 42.287$; $p < .001$, $\eta p^2 = .11$, but nor for the IS, $F(2, 724) = 1.069$; $p = .344$, $\eta p^2 = .003$. SES effect was observed for the Word condition, $F(1, 724) = 251.727$, $p < .001$, $\eta p^2 = .26$, Color, $F(1, 724) = 97.064$, $p < .001$, $\eta p^2 = .12$, and Color–Word variables, $F(1, 724) = 88.876$, $p < .001$ $\eta p^2 = .11$, but not for the IS, $F(1, 724) = 1.239$; $p = .266$, $\eta p^2 = .002$. Finally, the age × SES effect was only observed for the Color–Word condition, $F(2, 724) = 3.969$; $p = .019$, $\eta p^2 = .011$. Post hoc contrasts indicated significant differences (Bonferroni correction, $p < .01$) for the Word, Color, and Color–Word conditions among all age groups (i.e., 7–8 with 9–10, 7–8 with 11–12, and 9–10 with 11–12), but not for the interference score.

Normative data for the SCWT stratified by age and SES are available in the Supplementary material. Tables S1–3 provide percentiles for SCWT variables stratified by age and sex and descriptive statistics (i.e., mean, *SD*, minimum and maximum, skewness, and kurtosis) in middle-SES children, whereas Tables S4–6 provide percentiles for SCWT variables stratified by age and sex and descriptive statistics (i.e., mean, *SD*, minimum and maximum, skewness, and kurtosis) in low-SES children. Skew and kurtosis values less than 2 and less than 7, respectively, can be considered appropriate to assume normality (Kim, 2013).

Stepwise Regressions

To identify the socioeconomic predictors of SCWT performance, stepwise regression analysis was conducted. None of the VIF indicated multicollinearity, even with a strict cutoff (VIF > 4). Note that the higher the score in each socioeconomic indicator, the higher the poverty level. For the Word condition, MEL ($\beta = -.341$, $p = .001$) accounted for 27% of the variance while HCs ($\beta = -.212$, $p = .031$) explained an additional 1%. For the Color condition, only HCs ($\beta = -.331$, $p < .001$) explained 11% of the variance while for Color–Word condition only MEL ($\beta = -.327$, $p < .001$) accounted for 11% of the variance (see Table 3).

Conclusion S2

Our S2 results revealed there was a significant effect of age and SES, but no of sex, on SCWT performance. First, SCWT performance was found to improve with age. Specifically, in the middle-SES group, we observed that from 7–8 to 11–12 years old, the Word, Color, and Color–Word scores increased linearly as children gets older. These results are in agreement with those of earlier studies also conducted with Spanish-speaking children, but from different Q11 countries (see e.g., Martín et al., 1993; Rivera et al., 2017). The pattern of SCWT performance is also consistent with what would be expected for TD children; that is, a Color–Word score lower than that of the Color condition, which in turn is lower than the Word score (Golden, 2001). Low-SES children, in contrast, performed worse than their Middle-SES peers across all three SCWT conditions (i.e., Word, Color, and Color–Word). Besides, contrary to expected outcomes, interference increases linearly as child gets older following a distinct pattern to that observed in the Middle-SES group (see Figure 1).

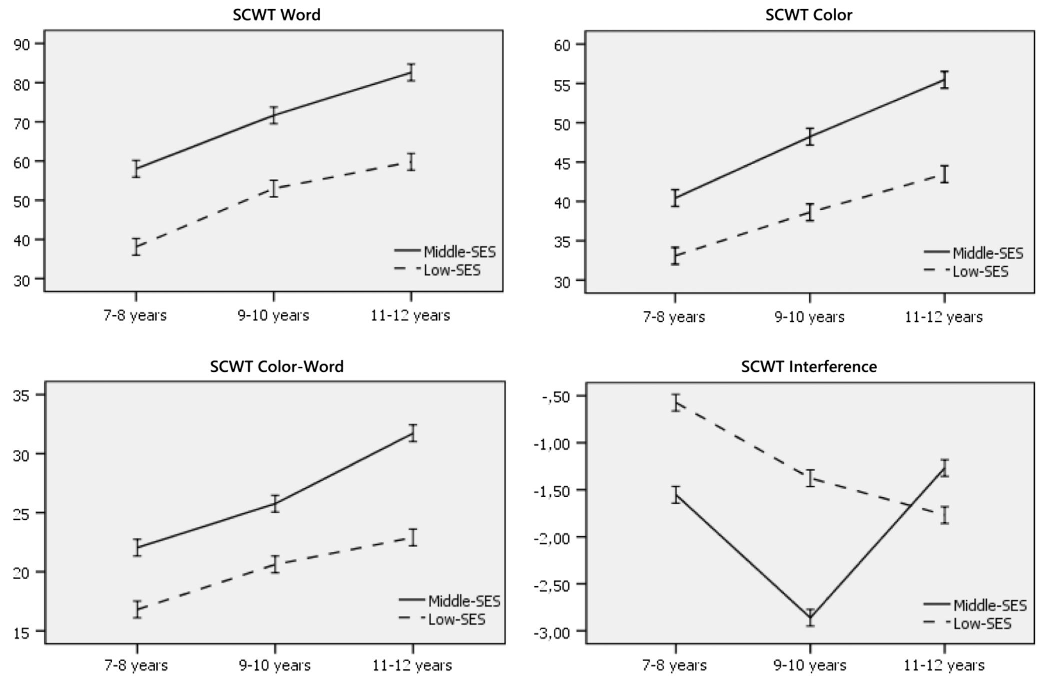
Finally, it should be noted that differences were observed between groups regarding the relationship among SCWT measures (i.e., Word > Color > Color–Word), with minor distance between scores in comparison to that observed in the Middle-SES children, due to the marked difference in the Word score. According to our results, differences in Word reading would be explained by MEL and HCs, whereas differences in Color naming skills and Color–Word condition would be explained by HCs and MEL, respectively. This association between parents' educational level and cognitive performance has been observed for both tasks that value memory and attention (Arán Filippetti, 2012; Matute et al., 2009) and EF (Ardila et al., 2005; Klenberg et al., 2001; Piccolo et al., 2016). Poor sanitary conditions at home (Bradley & Corwyn, 2002; Guo & Harris, 2000) and in the neighborhood (Santos et al., 2008) have also been proposed as significant factors that influence children's cognitive development. This suggests that reading and naming skills are influenced by SES, which leads to a less interference that resembles what is observed in children with learning disabilities (see e.g., Golden & Golden, 2002). This would be due to the fact that, as occurs with prereading children (Martín et al., 2012) or with children with poor reading skills (Protopapas et al., 2007), the poor reading proficiency noticed in Low-SES children during the Word condition would decrease the effect of reading dominance over the presence of Stroop interference. These findings emphasize the importance of considering not only the reading ability when

Table 3
Summary of Stepwise Regression Predicting SCWT Conditions

Step	Predictor	Condition developmental pattern of interference scores according to groupon	R	R^2	R^2 change	β	t	p
Step 1	MEL	Word	.518	.269	.269	−.518	−9.623	<.001
Step 2	MEL		.531	.282	.013	−.341	−3.488	.001
	HCs					−.212	−2.164	.031
Step 1	HCs	Color	.331	.109	.109	−.331	−5.564	<.001
Step 1	MEL	Color–Word	.327	.107	.107	−.327	−5.500	<.001

Note. MEL = maternal education level; HCs = housing conditions; SCWT = Stroop Color and Word Test.

Figure 1
Developmental Pattern of Stroop Color and Word Test (SCWT) Performance According to Socioeconomic Status (SES) and Age



using the Stroop color and word paradigm but also the SES effects.

Study 3 (S3)

Study 3 aimed at exploring the sensitivity of the SCWT with ADHD Spanish-speaking children, predominantly inattentive and combined subtypes.

Method

Participants

The sampling consisted of 104 children (41 girls and 63 boys) divided as follows: (a) TD children: 60 children between 7 and 12 years of age, of medium socioeconomic level, living in Argentina. Like previous studies, parents or legal guardians' consents was obtained prior to assessment, considering the same inclusion criteria; (b) ADHD children: Archived data collected from 44 children with ADHD combined ($n = 30$) and inattentive subtypes ($n = 14$) from 7 to 12 years old of both sexes. The children were clinically diagnosed by different professionals (i.e., pediatric neurologists and neuropsychologists) based Q12
on DSM-IV (American Psychiatric Association [APA], 1994), the abbreviated Conners' Rating Scales for parents (CPRS-HI) (Conners, 1990) and diagnostics interviews with their parents and teachers. DSM-V maintains the same DSM-IV 18 core symptoms and ADHD symptom domains, indicating that the DSM-IV criteria for ADHD have largely withstood the test of time (Epstein & Loren, 2013). The presence of ADHD clinical symptoms was further verified using the Conner's Continuous Performance Test II (CPT). The CPT is one of the most widely used tests for the neuropsychological evaluation of children with ADHD (Nichols & Waschbusch, 2004; Riccio et al., 2001). CPT paradigms efficiently discriminate between children with ADHD from TD (Nichols & Waschbusch, 2004) and predict the presence of ADHD symptoms (Epstein et al., 2003). Most children obtained results that are consistent with and ADHD clinical profile according to the test reports (i.e., ADHD combined group: CPT omissions: $M = 29.77$, $SD = 15.09$;

CPT commissions $M = 28.50$, $SD = 4.41$; CPT mean response time for all target responses: $M = 453.88$, $SD = 80.45$; ADHD inattentive group: CPT omissions: $M = 33.79$, $SD = 27.71$; CPT commissions $M = 25.71$, $SD = 6.07$; CPT mean response time for all target responses: $M = 512.96$, $SD = 99.05$). Exclusion criteria were (a) IQ below 75, (b) history of other neurological disorders, sensory, or motor impairment, (c) illiterate children or children with reading disabilities, and (d) children with other DSM-IV diagnoses. Though some children presented emotional or behavioral symptomatology, they do not fully meet criteria for another comorbid diagnosis. Prior to analyzing the differences on task performance according to subtype, it was verified through ANOVA that there were no significant differences between the three groups in terms of age ($p = .384$) and IQ levels ($p = .492$). None of the children were under stimulant medication at the time of assessment.

Instruments

Stroop Color–Word Test (Golden, 2001)

See S1 for description.

Kaufman Brief intelligence Test (Kaufman & Kaufman, 2000)

See S1 for description.

Conners' Continuous Performance Test II (Conners, 2000)

The CPT is a widely used task of sustained attention and inhibitory control. Examiners are asked to press the space bar every time any letter appears, except for the letter "X." Main indicators are (a) omissions, (b) commissions, and (c) hit response time.

Statistical Analyses

For group comparisons, MANOVA was performed using the Bonferroni correction for multiple comparisons ($p < .01$). Group (ADHD combined, ADHD inattentive, and TD) was entered as the between-subjects factor and the four SCWT scores (Word, Color, Color–Word, and IS) as dependent variables. Effect sizes were calculated using Cohen's *d*.

Results

SCWT Task Performance According to Clinical Versus NonClinical Sample

A significant main effect of group was found Hotelling's $F(8, 194) = 5.305$; $p < .001$, $\eta p^2 = .18$. Specifically, these differences were observed for the Word, $F(2, 101) = 6.169$; $p = .003$, $\eta p^2 = .11$, Color, $F(2, 101) = 4.107$; $p = .019$, $\eta p^2 = .08$, and Color–Word, $F(2, 101) = 4.949$; $p = .009$, $\eta p^2 = .09$ conditions. For the Word condition, the inattentive group performed significantly poorer than TD (Bonferroni correction, $p < .01$) while for the Color and Color–Word variables, the combined subtype performed significantly poorer than their TD peers (Bonferroni correction, $p < .01$). No significant differences between groups were found for the interference measure ($p = .611$). A medium effect size was observed for the W, C, and CW conditions for the ADHD combined subtype, while a medium effect size for the W condition and a small effect size for the C and CW conditions were observed for the ADHD inattentive group (see Table 4).

Conclusion S3

S3 results revealed distinct patterns of SCWT performance according to ADHD subtypes. Specifically, the ADHD combined group performed worse on the Color and Color–Word conditions than their TD peers. However, the Word > Color > Color–Word condition is met in a similar proportion than that observed in TD children (see Table 4 for the similar differences between scores; i.e., *TD*: Color to Word = 23 items; Color–Word to Color = 21 items; *ADHD combined*: Color to Word = 20 items; Color–Word to Color = 19 items). As a result, their interference score was not significantly reduced, due to their normal pattern of scores (although lower-than-expected) across conditions. In contrast, children with ADHD inattentive subtype scored significantly lower than TD children only on the Word condition. Besides, the Word > Color > Color–Word condition is met but in a different proportion than that of their TD peers (i.e., *ADHD inattentive*: Color to Word = 11; Color–Word to Color = 21). Thus, the markedly reduced Word score in the inattentive subtype would produce a false negative on the interference

Table 4
SCWT Performance According to ADHD Subtype

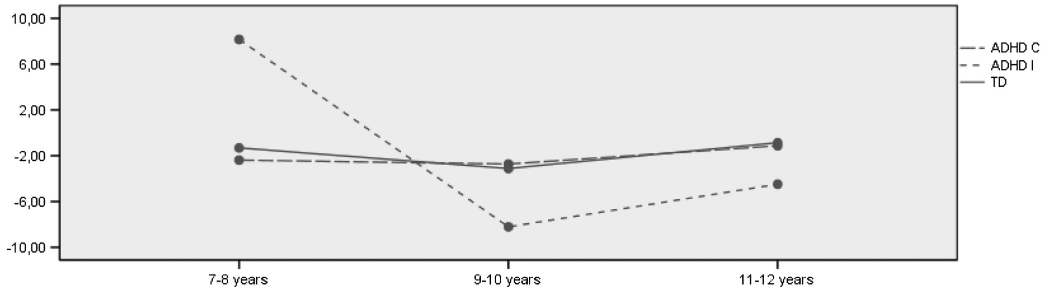
	ADHD C		ADHD I		TD								Cohen's *d*	
	M_1	SD_1	M_2	SD_2	M_3	SD_3	F 2,101	p	M_1–M_2	M_1–M_3	M_2–M_3	Contrasts between groups	ADHD C	ADHD I
Word	61.00	14.42	54.57	23.97	70.40	16.95	6.169	.003	.766	.052	.008	ADHD I < TD	−0.60**	−0.76**
Color	40.87	10.15	43.79	11.25	47.15	9.49	4.107	.019	1.00	.017	.768	ADHD C < TD	−0.64**	−0.32*
Color–Word	21.90	6.42	23.00	8.39	26.32	6.23	4.949	.009	1.00	.010	.281	ADHD C < TD	−.070**	−0.45*
Interference	−2.24	5.47	−.399	11.96	−1.76	3.20	.495	.611	.974	1.00	1.00	*ns*	.011	.016

Note. Bonferroni correction, $p < .01$. Effect sizes: $^{*}d > 0.2$; $^{**}d > 0.5$. SCWT = Stroop Color and Word Test; TD = Typically Developing; ADHD = Attention Deficit and Hyperactivity Disorder.

effect. As a result, only in the ADHD combined group, interference follows a developmental pattern similar to that of TD children (see Figure 2). Considering that according to our S1 results, the Word condition demands not only reading speed but also attentional shift and processing speed, it could be hypothesized, as suggested in previous studies (see Golden & Golden, 2002), that in the inattentive group, the slower Word score would be mainly due to attentional and EF problems (i.e., slow cognitive tempo and initiation problems).

Examining previous studies that have analyzed SCWT performance in children with ADHD, there are similarities among scores with a consistent worse performance in ADHD children as compared to healthy controls. For instance, López-Villalobos et al. (2010) observed in Spanish children with ADHD from 7 to 11 years of age an average of 62.12, 41.28, and 22.80 for the Word, Color, and Color–Word conditions, respectively. In English-speaking children ($M_{\text{age}} = 9.88$), a mean of 68.1 has been noticed for the Word variable, 49.3 for the Color condition, and 26.5 for the Color–Word sheet, with significant differences when comparing the performance with that of TD children only in the Color–Word condition (Golden & Golden, 2002). However, higher values have been reported for children with ADHD from other countries. For instance, in Israeli children aged 9–16 years, an average of 74.52 has been reported for the Word condition, 55.90 for Color, and 28.31 for the Color–Word variable (Lufi et al., 1990). In Spanish children, Elosúa et al. (2017) observed a mean of 99.15, 72.00, and 43 for the Word, Color, and Color–Word conditions, respectively. Thus, our scores regarding the ADHD combined group are consistent with those reported by López-Villalobos et al. (2010) in Spanish children with ADHD and emphasize the importance of establishing developmental norms for children from different countries. As regards the patterns of performance according to ADHD subtypes, earlier studies have yielded mixed results. Although some authors did not find differences on SCWT conditions according to ADHD subtypes (López-Villalobos et al., 2010; Nigg et al., 2002), other studies have reported distinct patterns of performance (Bará-Jiménez et al., 2003; Houghton et al., 1999), as findings from this study.

Figure 2
Developmental Pattern of Interference Scores According to Group



Finally, our results indicate a medium effect size for all SCWT conditions for the ADHD combined subtype and a low effect size for the Color and Color–Word conditions for the ADHD inattentive subtype. Consistently, most of previous studies have found a medium effect size for the three SCWT conditions, indicating that these measures are moderately sensitive to executive deficit in children with ADHD (Homack & Riccio, 2004). However, the effect size for the interference score is small and would not allow differentiating children with ADHD from other disorders (van Mourik et al., 2005).

General Discussion

The present study aimed at (a) examining the convergent and divergent validity of the SCWT and its cognitive predictors in TD children, (b) analyzing the effect of SES and age on SCWT performance, further establishing normative data for Spanish-speaking children, and (c) exploring the sensitivity of the SCWT for the identification of EF deficits in children with ADHD predominantly combined and inattentive subtypes.

Regarding the SCWT convergent and divergent validity, results from S1 showed moderate correlations between SCWT conditions and attention and EF measures. Besides, our results are in line with previous reports (Golden & Golden, 2002; Protopapas et al., 2007) suggesting that SCWT performance would not depend on IQ levels. When examining SCWT cognitive predictors, regression analyses revealed that reading speed and age predicted performance on the Word condition while processing speed and impulsivity predict performance on the Color measure. In turn, performance on the Color–Word condition was explained by both color-naming ability (i.e., Color condition) and attentional and inhibitory control. Overall, our data suggest that the SCWT would be a valid measure for assessing children's EF, and that each condition would offer distinctive and complementary information related mainly to inhibition and attentional processes. Taken together, our results offer additional support to existing research (León-Carrion et al., 2004; Protopapas et al., 2007) which propose that the Stroop values inhibitory processes and reading speed (mainly in its Word condition), and it would not depend on children's intelligence (Golden & Golden, 2002).

S2 results first revealed an effect of age, on the Word, Color naming, and incongruent (i.e., Color–Word) conditions of the SCWT. These findings are in line with previous studies showing an improvement in reading speed, naming skills, and the ability to inhibit with age, without differences according to sex (León-Carrion et al., 2004; Roy et al., 2018). However, there was no age effect on the interference measure. Previous studies have also reported stability between 7 and 10 years old in interference capacity (Martín et al., 2012) and a similar pattern of performance characterized by less interference at age 7–8 years, an increase at 9–10 years, and a new decrease at 11–12 years. This nonlinear developmental pattern (see e.g., Comalli et al., 1962) has been explained by age differences in the automaticity of reading (Schiller, 1966) and the development of inhibitory function with age (León-Carrion et al., 2004) associated with the maturation of the frontal regions (Adleman et al., 2002; Schroeter et al., 2004). Thus, at age 7–8 years, a lesser automaticity of reading would produce less interference on the CW condition, whereas at 9–10 years of age, reading dominance would lead to a greater interference. In turn, at age

11–12 years, as reading is equally automated, it could be hypothesized that the decrease in the interference effect is due to the fact that children have better strategies to resist interference with maturation of EFs.

Second, and in line with previous studies (Arán Filippetti & Richaud de Minzi, 2012; Sarsour et al., 2011), we found a SES effect on SCWT performance with a markedly reduced Word score ($M = 52.85$, $SD = 15.12$) in the low-SES group, especially at younger ages (i.e., 7–10 years). Accordingly, and as expected, the Low-SES children showed less interference, mainly due to their lower Word score. It has been proposed that the negative SES impact on cognitive and emotional development would be the result of the coexistence of multiple environmental risk agents, rather than the presence of a single factor (Evans, 2004; Sameroff, 1998). Indeed, although maternal educational level and withdrawal conditions have demonstrated to be important predictors of cognitive performance (Arán Filippetti & Richaud de Minzi, 2012), cognitive stimulation at home, parental practices, child health, and home environment would mediate the relationship between the poverty and cognitive development (Guo & Harris, 2000). In accordance to previous studies suggestions (see e.g., Rivera et al., 2017; Rodríguez Barreto et al., 2016), these findings do remark the importance of establishing differentiated normative data for the SCWT according to SES.

Finally, S3 results indicated that children with ADHD performed more poorly than controls in all three SCWT conditions, with a differential pattern of performance between subtypes when comparing to the TD group. These results are in agreement with previous research that has indicated that children with ADHD usually show poorer performance on all three SCWT conditions as compared to TD children (Homack & Riccio, 2004; van Mourik et al., 2005) with no differences in terms of interference control (Golden & Golden, 2002; van Mourik et al., 2005). Poor Stroop performance across all three measures has been associated with the reversed asymmetry of the caudate (Semrud-Clikeman et al., 2000). Because the caudate is part of the frontal–striatal system underlying response inhibition, caudate asymmetry would support Barkley's (1997) hypothesis of disinhibition as a core feature of ADHD (Semrud-Clikeman et al., 2000). Besides, neuroimaging studies analyzing brain differences between children with ADHD and healthy controls have also reported size reduction and differences in the symmetry of the prefrontal regions, age-dependent abnormalities of the caudate nucleus, smaller cerebellar hemispheric volume, and a reduction of gray and white matter in the prefrontal cortex regions (for a review, see Krain & Castellanos, 2006). Thus, the slower retrieval of color names and reading speed could be an indicator of abnormalities in the brain structure in ADHD (van Mourik et al., 2005). However, when SCWT performance is impaired, it is important to examine the pattern of scores across the three conditions, as different profiles have been observed depending on the child condition (e.g., ADHD, learning disabilities, etc.) (Golden & Golden, 2002). Indeed, as previously suggested (Bará-Jiménez et al., 2003; Houghton et al., 1999), our results indicate that the ADHD inattentive group may show impairment on the Word measure mainly due to attentional problems, whereas the combined group may show poor performance on all three SCWT measures due to deficient inhibitory control.

Before discussing the implications of the results, we must address some limitations. First, our sample was intentional and was limited to Argentine 7–12 years old children. Thus, results cannot be generalized to children from different countries or from distinct cultural contexts. However, the sample size used to establish normative data was considerably larger in relation to other normative studies on EF measures, and included three important regions of Argentina (i.e., Buenos Aires, Santa Fe, and Entre Ríos). Second, we did not include a measure of reading speed for the sample of children with ADHD. However, we controlled that children with ADHD did not have a previous diagnosis of reading disability and there were no differences between groups regarding general intelligence. Finally, a relatively small number of ADHD children that met criteria for ADHD, combined, and inattentive subtypes were included in Study 3. Therefore, future studies would benefit from assessing larger samples also including children with ADHD predominantly Hyperactive–Impulsive.

This study has substantial clinical and educational implications for the assessment of EF in child populations. The SCWT is a short, easy-to-administer test that provides a wealth of information (Golden & Golden, 2002). Specifically, and according to our results, it is a valid measure to assess reading speed, naming function (NF), and EFs (interference control) in TD children, and it is not based on their intellectual level. However, when interpreting SCWT results, it is important to consider the variables age, reading skills, and SES, so as not to obtain a false negative on inhibition assessment. First, to interpret SCWT performance, it is important to examine the existence of the direct lexical reading route as a function of age. Considering that the NF involves the reaudit of a visual stimulus, a 9–10-year-old child without automatized reading skills (as shown in a low performance on the Word condition) could perform well when reaudit of the XXXX (i.e., the Color condition) but not in the Word measure, as the word (e.g., BLUE) is not constituted for the child in a visual Gestalt, that is, the child has not constituted that visual stimulus (as a whole) for reaudit. Furthermore, our findings Q13 also emphasize the importance of analyzing SCWT performance in light of normative data according to SES, as in the low-SES group, our results not only showed lower scores than the expected mean values for their age but also demonstrated that the relationship between the three conditions was nonequivalent to that of Middle-SES children. Thus, knowledge on those factors that could influence SCWT performance and further setting normative values adjusted for age and SES do become of great relevance for the neuropsychological assessment of TD children. Finally, our findings also provide useful data for the assessment of children with executive dysfunction. In this regard, it would be necessary to cautiously consider the interference measure under the following conditions: (a) a low performance on the Word condition, but a good one on the Color and Color–Word conditions. This distinctive pattern would reflect children with lack of reading dominance (e.g., low-SES children) or attentional problems (e.g., ADHD inattentive subtype) with a consequent good performance on the Color–Word condition and a false negative on interference effect, as it would be more natural or easier for them to naming the colors than reading the words and (b) a good performance on the Word condition, but a low one in Color and Color–Word variables. This pattern would reflect difficulties during the Color–Word condition by not being able to update (quickly and accurately) the color of the word due to a NF alteration or a deficient inhibitory control (e.g., ADHD combined subtype). Thus, although the interference score would be strictly valid after a good performance on the Word and Color conditions, it is also important to analyze the pattern of performance across all three SCWT conditions (examining the relative distance between scores), as they could provide relevant information about different underlying cognitive processes.

Regarding the utility of the SCWT with ADHD, although the task has already proven to be suitable for monitoring the clinical response to pharmacological treatment with methylphenidate (Langleben et al., 2006), it has also demonstrated moderate sensitivity in detecting executive deficits in ADHD and other disorders (see Homack & Riccio, 2004). For this reason, in the neuropsychological field, it should be used in conjunction with other executive tasks (van Mourik et al., 2005) to increase its predictive validity (Perugini et al., 2000), without losing sight of the importance of considering the study of the proportional relationship between SCWT conditions when interpreting the interference effect.

References

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A. L. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage*, *16*(1), 61–75. https://doi.org/10.1006/nimg.2001.1046

Aguilar-Alonso, Á., & Moreno-González, V. (2012). Neuropsychological differences between samples of dyslexic and reader children by means of NEPSY. *The UB. The Journal of Psychology*, *42*, 35–50.

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, *100*(2), 235–251. https://doi.org/10.1037/0022-0663.100.2.235

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Press.

Arán Filippetti, V. (2012). Estrato socioeconómico y habilidades cognitivas en niños escolarizados: Variables predictoras y mediadoras. *Psykhe (Santiago)*, *21*(1), 3–20. https://doi.org/10.4067/S0718-22282012000100001

Arán Filippetti, V. (2013). Structure and invariance of executive functioning tasks across socioeconomic status: Evidence from Spanish-speaking children. *The Spanish Journal of Psychology*, *16*, Article E101. https://doi.org/10.1017/sjp.2013.102

Q14

Arán Filippetti, V., & Allegri, R. F. (2011). Verbal fluency in Spanish-speaking children: Analysis model according to task type, clustering, and switching strategies and performance over time. *The Clinical Neuropsychologist*, *25*, 413–436. https://doi.org/10.1080/13854046.2011.559481

Arán Filippetti, V., & Richaud de Minzi, M. C. (2012). A structural analysis of executive functions and socioeconomic status in school-age children: Cognitive factors as effect mediators. *The Journal of Genetic Psychology*, *173*(4), 393–416. https://doi.org/10.1080/00221325.2011.602374

Arán Filippetti, V., & Richaud, M. C. (2017). A structural equation modeling of executive functions, IQ and mathematical skills in primary students: Differential effects on number production, mental calculus and arithmetical problems. *Child Neuropsychology*, *23*, 864–888. https://doi.org/10.1080/09297049.2016.1199665

Arán Filippetti, V., & Krumm, G. (2020). A hierarchical model of cognitive flexibility in children: Extending the relationship between flexibility, creativity and academic achievement. *Child Neuropsychology*, *26*, 770–800. https://doi.org/10.1080/09297049.2019.1711034

Arán Filippetti, V., Krumm, G., & Raimondi, W. (2015). Funciones Ejecutivas y sus correlatos con Inteligencia Cristalizada y Fluida: Un estudio en Niños y Adolescentes. *Neuropsicologia Latinoamericana*, *7*(2), 24–33.

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, *28*, 539–560. https://doi.org/10.1207/s15326942dn2801_5

Armengol, C. G. (2002). Stroop test in Spanish: Children's norms. *The Clinical Neuropsychologist*, *16*(1), 67–80. https://doi.org/10.1076/clin.16.1.67.8337

Bará-Jiménez, S., Vicuña, P., Pineda, D. A., & Henao, G. C. (2003). Perfiles neuropsicológicos y conductuales de niños con trastorno por déficit de atención/hiperactividad de Cali, Colombia. *Revista de Neurología*, *37*(7), 608–615. https://doi.org/10.33588/rn.3707.2003189

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. https://psycnet.apa.org/doi/10.1037/0033-2909.121.1.65. https://doi.org/10.1037/0033-2909.121.1.65

Barkley, R. A. (1998). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). Guilford Press. https://doi.org/10.1038/scientificamerican0998-66

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of Abnormal Child Psychology*, *20*(2), 163–188. https://doi.org/10.1007/BF00916547

Benton, A. L., & Hamsher, K. (1989). *Multilingual Aphasia Examination manual*. AJA Associates.

Bondi, M. W., Serody, A. B., Chan, A. S., Eberson-Shumate, S. C., Delis, D. C., Hansen, L. A., & Salmon, D. P. (2002). Cognitive and neuropathologic correlates of Stroop Color-Word Test performance in Alzheimer's disease. *Neuropsychology*, *16*(3), 335–343. https://doi.org/10.1037/0894-4105.16.3.335

Borella, E., De Ribaupierre, A., Cornoldi, C., & Chicherio, C. (2013). Beyond interference control impairment in ADHD: Evidence from increased intraindividual variability in the color-stroop test. *Child Neuropsychology*, *19*(5), 495–515. https://doi.org/10.1080/09297049.2012.696603

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Brickenkamp, R. (1962). *Aufmerksamkeits-Belastungs-Test (Test d2)* [The d2 test of Attention]. 1st ed.). Hogrefe. [Spanish adaptation by Seisdedos Cubero, N. (2004); 2nd ed., revised and expanded. Madrid: TEA Ediciones]. Q15

Brickenkamp, R., & Zillmer, E. (1998). *The d2 Test of Attention* (U.S. 1st Ed.). Hogrefe & Huber Publishers. Q16

Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, *15*(1), 17–32. https://doi.org/10.1177/1094428110392383

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An H215O PET study of Stroop task performance. *NeuroImage*, *2*, 264–272. https://doi.org/10.1006/nimg.1995.1034

Cataldo, M. G., Nobile, M., Lorusso, M. L., Battaglia, M., & Molteni, M. (2005). Impulsivity in depressed children and adolescents: A comparison between behavioral and neuropsychological data. *Psychiatry Research*, *136*(2-3), 123–133. https://doi.org/10.1016/j.psychres.2004.12.012

Chevalier, H., Metz-Lutz, M. N., & Segalowitz, S. J. (2000). Impulsivity and control of inhibition in Benign Focal Childhood Epilepsy (BFCE). *Brain and Cognition*, *43*(1-3), 86–90.

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2010). Validity. In D. B. Hash (Ed.), *Psychological testing and assessment: An introduction to tests and measurement* (pp. 181–210). McGraw Hill.

Comalli, P. E., Jr., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood, and aging. *The Journal of Genetic Psychology*, *100*(1), 47–53. https://doi.org/10.1080/00221325.1962.10533572

Conners, C. K. (1990). *Conners' abbreviated symptom questionnaire: Parent version, teacher version*. Multi-Health Systems.

Conners, C. K. (2000). *Conners' Continuous Performance Test II*. Multi-Health Systems.

Cox, C. S., Chee, E., Chase, G. A., Baumgardner, T. L., Schuerholz, L. J., Reader, M. J., … Denckla, Q17 M. B. (1997). Reading proficiency affects the construct validity of the Stroop test interference score. *The Clinical Neuropsychologist*, *11*(2), 105–110. https://doi.org/10.1080/13854049708407039

Craig, F., Margari, F., Legrottaglie, A. R., Palumbi, R., De Giambattista, C., & Margari, L. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatric Disease and Treatment*, *12*, 1191–1202. https://doi.org/10.2147/NDT.S104620

Demakis, G. J. (2004). Frontal lobe damage and tests of executive processing: A meta-analysis of the category test, stroop test, and trail-making test. *Journal of Clinical and Experimental Neuropsychology*, *26*(3), 441–450. https://doi.org/10.1080/13803390490510149

Q18 Demontis, D., Walters, R. K., Martin, J., Mattheisen, Q19 M., Als, T. D., Agerbo, E., … Cerrato, F. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, *51*(1), 63–75. https://doi.org/10.1038/s41588-018-0269-7

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, *45*(3), 259–268. https://doi.org/10.1016/j.jrp.2011.02.004

Elosúa, M. R., Del Olmo, S., & Contreras, M. J. (2017). Differences in executive functioning in children with Attention Deficit and Hyperactivity Disorder (ADHD). *Frontiers in Psychology*, *8*, Article 976. https://doi.org/10.3389/fpsyg.2017.00976

Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test performance measures and ADHD behaviors. *Journal of Abnormal Child Psychology*, *31*(5), 543–554. https://doi.org/10.1023/A:1025405216339

Epstein, J. N., & Loren, R. E. (2013). Changes in the definition of ADHD in DSM-5: Subtle but important. *Neuropsychiatry*, *3*(5), 455–458. https://doi.org/10.2217/npy.13.59

Evans, G. W. (2004). The environment of childhood poverty. *American Psychologist*, *59*(2), 77–92. https://psycnet.apa.org/doi/10.1037/0003-066X.59.2.77. https://doi.org/10.1037/0003-066X.59.2.77

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 31/2–7 years old on Stroop-like day-night test. *Cognition*, *53*, 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Golden, C. J. (1978). *The Stroop Color and Word Test*. Stoelting Company.

Golden, C. J. (2001). *Stroop, Test de Colores y Palabras*. TEA Ediciones.

Golden, Z. L., & Golden, C. J. (2002). Patterns of performance on the Stroop Color and Word Test in children with learning, attentional, and psychiatric disabilities. *Psychology in the Schools*, *39*(5), 489–495. https://doi.org/10.1002/pits.10047

Gregory, R. J. (2007). *Psychology testing: History, principles, and applications* (5th ed.). Ally & Bacon.

Gruber, S. A., Rogowska, J., Holcomb, P., Soraci, S., & Yurgelun-Todd, D. (2002). Stroop performance in normal control subjects: An fMRI study. *NeuroImage*, *16*(2), 349–360. https://doi.org/10.1006/nimg.2002.1089

Guo, G., & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, *37*(4), 431–447. https://doi.org/10.1353/dem.2000.0005

Friedman, R. (1971). The relationship between intelligence and performance on the Stroop color-word test in second-and fifth-grade children. *The Journal of Genetic Psychology*, *118*(1), 147–148. https://doi.org/10.1080/00221325.1971.10532602

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1997). *WCST, Test de clasificación de tarjetas de Wisconsin*. TEA Ediciones. Q20

Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology*, *19*(6), 725–743. https://doi.org/10.1016/j.acn.2003.09.003

Houghton, S., Douglas, G., West, J., Whiting, K., Wall, M., Langsford, S., … Carroll, A. (1999). Q21 Differential patterns of executive function in children with attention-deficit hyperactivity disorder according to gender and subtype. *Journal of Child Neurology*, *14*(12), 801–805. https://doi.org/10.1177/088307389901401206

Johnson, W., Bouchard, T. J. Jr., Segal, N. L., Keyes, M., & Samuels, J. (2003). The Stroop Color-Word Test: Genetic and environmental influences; Reading, mental ability, and personality correlates. *Journal of Educational Psychology*, *95*(1), 58–65. https://doi.org/10.1037/0022-0663.95.1.58

Kaufman, A. S., & Kaufman, N. L. (2000). *K-BIT*. [Test Breve de Inteligencia de Kaufman (K-BIT) (2ª edición)]. TEA Ediciones.

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*, 407–428. https://doi.org/10.1207/S15326942DN2001_6

Korkman, M., Kirk, U., & Kemp, S. L. (1998). *NEPSY. A developmental neuropsychological assessment.* Psychological Corporation.

Krain, A. L., & Castellanos, F. X. (2006). Brain development and ADHD. *Clinical Psychology Review*, *26*(4), 433–444. https://doi.org/10.1016/j.cpr.2006.01.005

Q22 Krikorian, R., & Bartok, J. A. (1998). Developmental data for the Porteus Maze Test. *The Clinical Neuropsychologist*, *12*, 305–310. https://doi.org/10.1076/clin.12.3.305.1984

Langleben, D. D., Monterosso, J., Elman, I., Ash, B., Krikorian, G., & Austin, G. (2006). Effect of methylphenidate on Stroop Color–Word task performance in children with attention deficit hyperactivity disorder. *Psychiatry Research*, *141*(3), 315–320. https://doi.org/10.1016/j.psychres.2005.09.007

Lavoie, M. E., & Charlebois, P. (1994). The discriminant validity of the stroop color and word test: Toward a cost-effective strategy to distinguish subgroups of disruptive preadolescents. *Psychology in the Schools*, *31*(2), 98–107. https://doi.org/10.1002/1520-6807(199404)31:2<98::AID-PITS2310310203>3.0.CO;2-F

Lazarus, P. J., Ludwig, R. P., & Aberson, B. (1984). Stroop Color-Word Test: A screening measure of the selective attention to differentiate LD from non LD children. *Psychology in the Schools*, *21*, 53–60. https://doi.org/10.1002/1520-6807(198401)21:1<53::AID-PITS2310210110>3.0.CO;2-N

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

León-Carrion, J., García-Orza, J., & Pérez-Santamaría, F. J. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience*, *114*(10), 1291–1311. https://doi.org/10.1080/00207450490476066

Leverett, J. P., Lassiter, K. S., & Buchanan, G. M. (2002). Correlations for the Stroop Color and Word Test with measures of reading and language achievement. *Perceptual and Motor Skills*, *94*(2), 459–466. https://doi.org/10.2466/pms.2002.94.2.459

López-Villalobos, J. A., Serrano-Pintado, I., Andrés-De Llano, J. M., Delgado Sánchez-Mateos, J., Alberola-López, S., & Sánchez-Azón, M. I. (2010). Utilidad del test de Stroop en el trastorno por déficit de atención/hiperactividad. *Revista de Neurología*, *50*, 333–340. https://doi.org/10.33588/rn.5006.2009418

Lufi, D., Cohen, A., & Parish-Plass, J. (1990). Identifying attention deficit hyperactive disorder with the WISC-R and the Stroop color and word test. *Psychology in the Schools*, *27*(1), 28–34. https://doi.org/10.1002/1520-6807(199001)27:1<28::AID-PITS2310270105>3.0.CO;2-L

MacLeod, C. M. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. https://psycnet.apa.org/doi/10.1037/0033-2909.109.2.163. https://doi.org/10.1037/0033-2909.109.2.163 Q23

Mainville, M., Brisson, J., Nougarou, F., Stipanicic, A., & Sirois, S. (2015). Inhibition development: Comparison of neuropsychological and eye tracking measures. *Revista Argentina de Ciencias del Comportamiento*, *7*, 17–25.

Martín, R., Hernández, S., Rodríguez, C., García, E., Díaz, A., & Jiménez, J. E. (2012). Datos normativos para el Test de Stroop: Patrón de desarrollo de la inhibición y formas alternativas para su evaluación. *European Journal of Education and Psychology*, *5*(1), 39–51. https://doi.org/10.30552/ejep.v5i1.76

Mattson, S. N., Crocker, N., & Nguyen, T. T. (2011). Fetal alcohol spectrum disorders: Neuropsychological and behavioral features. *Neuropsychology Review*, *21*(2), 81–101. https://doi.org/10.1007/s11065-011-9167-9

Matute, E., Rosselli, M., Ardila, A., & Ostrosky-Solís, F. (2007). *Evaluación neuropsicológica infantil (ENI).* Manual Moderno.

Matute, E., Sanz, A., Gumá, E., Rosselli, M., & Ardila, A. (2009). Influencia del nivel educativo de los padres, el tipo de escuela y el sexo en el desarrollo de la atención y la memoria. *Revista Latinoamericana de Psicología*, *41*, 257–276. Q24

Méndez-Castellano, H., & de Méndez, M. C. (1994). *Sociedad y estratificación. Método Graffar—Mendez Castellano* [Society and stratification. Graffar–Mendez Castellano Method]. Fundacredesa.

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. https://doi.org/10.1006/cogp.1999.0734

Nichols, S. L., & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry and Human Development*, *34*, 297–315. https://doi.org/10.1023/B:CHUD.0000020681.06865.97

Nigg, J. T., Blaskey, L. G., Huang-Pollock, M. A., & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD subtypes. *Journal of the American Academy of Child &*

Adolescent Psychiatry, *41*, 59–66. https://doi.org/10.1097/00004583-200201000-00012

Noble, K. G., Farah, M. J., & McCandliss, B. D. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development*, *21*(3), 349–368. https://doi.org/10.1016/j.cogdev.2006.01.007

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, *87*, 256–259. https://doi.org/10.1073/pnas.87.1.256

Perugini, E. M., Harvey, E. A., Lovejoy, D. W., Sandstrom, K., & Webb, A. H. (2000). The predictive power of combined neuropsychological measures for attention-deficit/hyperactivity disorder in children. *Child Neuropsychology*, *6*, 101–114. https://doi.org/10.1076/chin.6.2.101.7059

Piccolo, L. D. R., Arteche, A. X., Fonseca, R. P., Grassi-Oliveira, R., & Salles, J. F. (2016). Influence of family socioeconomic status on IQ, language, memory and executive functions of Brazilian children. *Psicologia: Reflexão e Crítica*, *29*, Article 23. https://doi.org/10.1186/s41155-016-0016-x

Porteus, S. D. (2006). *Laberintos de Porteus (4ª edición revisada)*. TEA Ediciones.

Pratt, M. L., Leonard, H. C., Adeyinka, H., & Hill, E. L. (2014). The effect of motor load on planning and inhibition in developmental coordination disorder. *Research in Developmental Disabilities*, *35*, 1579–1587. https://doi.org/10.1016/j.ridd.2014.04.008

Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology*, *54*, 251–282. https://doi.org/10.1016/j.cogpsych.2006.07.003

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills*, *55*, 839–844. https://doi.org/10.2466/pms.1982.55.3.839

Reitan, R. M., & Wolfson, D. (1992). *Neuropsychological evaluation of older children*. Neuropsychology Press.

Riccio, C. A., Reynolds, C. R., & Lowe, P. A. (2001). *Clinical applications of continuous performance tests: Measuring attention and impulsive responding in children and adults*. John Wiley & Sons.

Rivera, D., Morlett-Paredes, A., Peñalver Guia, A. I., Irías Escher, M. J., Soto-Añari, M., Aguayo Arelis, Q25 A., … García de la Cadena, C. (2017). Stroop color-word interference test: Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation*, *41*(3), 605–616. https://doi.org/10.3233/NRE-172246

Rodríguez Barreto, L., Pineda Roa, C. A., & Pulido, Q26 N. D. C. (2016). Propiedades psicométricas del Stroop, test de colores y palabras en población colombiana no patológica. *Universitas Psychologica*, *15*(2), 255–272. https://doi.org/10.11144/Javeriana.upsy15-2.ppst

Rosselli-Cock, M., Matute-Villaseñor, E., Ardila-Ardila, A., & Botero-Gómez, V. E. (2004). Evaluación Q27 Neuropsicológica Infantil (ENI): Una batería para la evaluación de niños entre 5 y 16 años de edad. Estudio normativo colombiano. *Revista de Neurología*, *38*, 720–731. https://doi.org/10.33588/rn.3808.2003400

Roy, A., Kefi, M. Z., Bellaj, T., Fournet, N., Le Gall, D., & Roulin, J. L. (2018). The Stroop test: A developmental study in a French children sample aged 7 to 12 years. *Psychologie Française*, *63*(2), 129–143. https://doi.org/10.1016/j.psfr.2016.08.001

Rucklidge, J. J., & Tannock, R. (2002). Neuropsy- Q28 chological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*(8), 988–1003. https://doi.org/10.1111/1469-7610.00227

Santos, D. N., Assis, A. M. O., Bastos, A. C. S., Santos, L. M., Santos, C., Strina, A., … Barreto, M. L. (2008). Q29 Determinants of cognitive function in childhood: A cohort study in a middle income context. *BMC Public Health*, *8*(1), Article 202. https://doi.org/10.1186/1471-2458-8-202

Sameroff, A. J. (1998). Environmental risk factors in infancy. *Pediatrics*, *102*(5), 1287–1292.

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., & Boyce, W. T. (2011). Family socioeconomic status and child executive functions: The roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*, *17*(1), 120–132. https://doi.org/10.1017/S1355617710001335

Scarpina, F., & Tagini, S. (2017). The stroop color and word test. *Frontiers in Psychology*, *8*, Article 557. https://doi.org/10.3389/fpsyg.2017.00557

Semrud-Clikeman, M., Steingard, R. J., Filipek, P., Biederman, J., Bekken, K., & Renshaw, P. F. (2000). Using MRI to examine brain–behavior relationships in males with attention deficit disorder with hyperactivity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 477–484. https://doi.org/10.1097/00004583-200004000-00017

Schiller, P. H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology*, *72*, 105–108. https://psycnet.apa.org/doi/10.1037/h0023358. https://doi.org/10.1037/h0023358

Schroeter, M. L., Zysset, S., Wahl, M., & von Cramon, D. Y. (2004). Prefrontal activation due to Stroop interference increases during development—an event-related fNIRS study. *NeuroImage*, *23*(4), 1317–1325. https://doi.org/10.1016/j.neuroimage.2004.08.001

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological test: Administration, norms* *and commentary* (2nd ed.). Oxford University Press.

Swerdlow, N. R., Filion, D., Geyer, M. A., & Braff, D. L. (1995). Normal personality correlates of sensorimotor, cognitive and visuospatial gating. *Biological Psychiatry*, *37*, 286–299. https://doi.org/10.1016/0006-3223(94)00138-S

Q30

Temple, C. M., Carney, R. A., & Mullarkey, S. (1996). Frontal lobe function and executive skills in children with Turner's syndrome. *Developmental Neuropsychology*, *12*, 343–363. https://doi.org/10.1080/87565649609540657

Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K. W. (2012). The five-point test: Reliability, validity and normative data for children and adults. *PLOS ONE*, *7*(9), Article e46080. https://doi.org/10.1371/journal.pone.0046080

van Mourik, R., Oosterlaan, J., & Sergeant, J. A. (2005). The Stroop revisited: A meta-analysis of interference control in AD/HD. *Journal of Child Psychology and Psychiatry*, *46*(2), 150–165. https://doi.org/10.1111/j.1469-7610.2004.00345.x

Q31

van Veen, C., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage*, *27*, 497–504. https://doi.org/10.1016/j.neuroimage.2005.04.042

Wechsler, D. (2010). *WISC IV, Escala de Inteligencia para niños de Wechsler—IV*. Adaptación Argentina. Paidos.

Q32

Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics*, *9*(3), 490–499. https://doi.org/10.1007/s13311-012-0135-8

Q33

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*, 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006

Wu, K. K., Chan, S. K., Leung, P. W., Liu, W. S., Leung, F. L., & Ng, R. (2011). Components and developmental differences of executive functioning for school-aged children. *Developmental Neuropsychology*, *36*(3), 319–337. https://doi.org/10.1080/87565641.2010.549979

Received April 29, 2020
Revision received September 22, 2020
Accepted September 28, 2020 ■

Queries

Q1. Please provide the expansion for the abbreviations “PET”, “fMRI”, “VF”, “ENI” and “IS” used in this article.

Q2. The reference “Fiske (1971)” is cited in text but not provided in the reference list. Please provide complete details of the reference or delete the text citation.

Q3. Please check that the edits made to the sentence that begins with "In turn, when examining . . . " are OK and that the meaning has not been changed.

Q4. Please update "XXXX", as applicable.

Q5. “Golder (1975)” is cited in text but not provided in the reference list. Please provide complete details of the reference or delete the text citation.

Q6. Please provide expansion for the abbreviation “NEPSY.”

Q7. The reference “Brickenkamp (2004)” is cited in text but not provided in the reference list. Please provide complete details of the reference or delete the text citation.

Q8. The reference “Heaton et al. (1993)” is cited in text but not provided in the reference list. Please provide complete details of the reference or delete the text citation.

Q9. Please provide expansion for the abbreviation “FAS.”

Q10. The same abbreviation “VIF” has been used for two different explanations, “variation inflation factor” and “variance inflation factor.” Please provide correct information.

Q11. The reference “Martín et al. (1993)” is cited in text but not provided in the reference list. Please provide complete details of the reference or delete the text citation.

Q12. Please provide expansion for the abbreviation “DSM-IV” and “WCST.”

Q13. Please check that the edits made to the sentence that begins with "Furthermore, our findings . . . " are OK and that the meaning has not been changed.

Q14. Kindly check and confirm the page ranges “1-15” has been changed to “Article E101” for the reference “Arán Filippetti (2013)”.

Q15. The reference “Brickenkamp (1962)” is not cited in text but provided in the reference list. Please cite in text or delete from the list.

Q16. Please provide the chapter title and page ranges for the reference “Brickenkamp and Zillmer (1998)”.

Q17. Please provide the complete author details for the reference “Cox et al. (1997)”.

Q18. CrossMark reports an erratum (or similar issue). The CrossMark type is "isPreprintOf". Additional information can be found at https://doi.org/10.1101/145581. (Ref. "Demontis, Walters, Martin, Mattheisen, Als, Agerbo, Cerrato, . . ., 2019")

Q19. Please provide the complete author details for the reference “Demontis et al. (2019)”.

Q20. The reference “Heaton et al. (1997)” is not cited in text but provided in the reference list. Please cite in text or delete from the list.

Q21. Please provide the complete author details for the reference “Houghton et al. (1999)”

Q22. The reference “Krikorian and Bartok (1998)” is not cited in text but provided in the reference list. Please cite in text or delete from the list.

Q23. The reference “McLeod (1991)” is not cited in text but provided in the reference list. Please cite in text or delete from the list. And also note that the author name “McLeod” has been changed to “MacLeod”, kindly check.

Q24. Please provide the complete author details because of author names and initials mismatch in the reference “Matute et al. (2009) ”.

Q25. Please provide the complete author details for the reference “Rivera et al. (2017)”.

Q26. Please provide the complete author details because of author names and initials mismatch in the reference “Rodríguez Barreto et al. (2016)”.

Q27. Please provide the complete author details because of author names and initials mismatch in the reference “Rosselli-Cock et al. (2004)”.

Q28. Please note that the journal title “*Journal of Child Psychology and Psychiatry, and Allied Disciplines*” has been changed to “*Journal of Child Psychology and Psychiatry*” for the reference “Rucklidge and Tannock (2002)”.

Q29. Please provide the complete author details for the reference “Santos et al. (2008)”

Q30. The reference “Swerdlow et al. (1995)” is not cited in text but provided in the reference list. Please cite in text or delete from the list.

Q31. Please note that the journal title “*Journal of Child Psychology and Psychiatry, and Allied Disciplines*” has been changed to “*Journal of Child Psychology and Psychiatry*” for the reference “van Mourik et al. (2005)”.

Q32. Please provide the complete details for the reference “Wechsler (2010)”.

Q33. Please note that the journal title “*Neurotherapeutics; the Journal of the American Society for Experimental NeuroTherapeutics*” has been changed to “*Neurotherapeutics*” for the reference “Willcutt (2012)”.

Q34. Please check and approve the updated running head. Else, kindly provide a running head not exceeding 50 characters including punctuations and spaces.

Q35. Please confirm if author names are correctly rendered in author group, in ORCID ID and in correspondence.