Archives Italiennes de Biologie, 144: 45-62, 2006.

STROOP EFFECTS FROM 3 TO 10 YEARS: THE CRITICAL ROLE OF READING ACQUISITION

A. PERU[1, 2], C. FACCIOLI[1], G. TASSINARI[1]

[1] *Department of Neurological and Visual Sciences, University of Verona, Italy;* [2] *"Casa del Sole" Rehabilitation Center, Curtatone, Italy*

INTRODUCTION

Stroop (27) described a series of color-word interferences. The most popular of his observations is that it takes much longer to name the color of incongruently-colored color names than to name the color of colored squares. In turn, reading a color name printed in a color that does not correspond with the name takes longer (although not much) than reading the same color name printed in black. The first of these phenomena is known as the classic Stroop effect (or Stroop effect *tout court*), the second as the reverse Stroop effect.

Since then, Stroop's evidence has been reproduced many times and in many variants, and different interpretations have been proposed (see 19 for a review; more recently, 4, 8, 14, 20, 21, 23, 28, 30, 31).

Given that the two main processes involved in the Stroop effects, i.e., reading and the inhibition of compelling but inappropriate responses, are both linked to maturational changes, much research into color-word interferences has been conducted in the context of developmental studies. Considering the classic Stroop effect, Comalli and co-workers (12) first included young participants in a very large sample ranging in age from 7 to 80 years. They found the greatest interference in the young children; the interference then decreased in adulthood and increased again with most advanced age. A few years later, testing students from Grade 1 to College, Schiller (26) demonstrated that the Stroop effect shows an inverted U-shaped pattern. In particular, interference was minimal in 1st graders, became maximal in 2nd and 3rd graders, and then declined gradually.

This pattern has been interpreted as the result of two concomitant but opposing trends (11): an increasing automation of the reading process, and an increasing maturation of the executive and attentional functions based on circuits in the prefrontal and anterior cingulate cortex (1, 10, 32). The interference increases for as long as the first process prevails over the second, after which, in spite of reading skills being fully acquired, it decreases: reading the incongruent name becomes an increasingly

Corresponding Author: Dott. G. Tassinari, Dipartimento di Scienze Neurologiche e della Visione, Sezione di Fisiologia Umana, Strada Le Grazie 8, I-37134 Verona, Italy - Tel. +39 045 8027153; Fax: +39 045 580881 - E-mail: giancarlo.tassinari@univr.it

inescapable step in the color-naming process, but at the same time the ability to inhibit the response corresponding to the written word improves.

Many other studies demonstrated that the magnitude of the classic Stroop effect is strictly related to reading proficiency. Among others, Cox et al. (13) studied parents of children with learning disabilities (therefore suspected of poor reading skills), and found that the Stroop effect correlated with other measures of response inhibition only in those subjects who satisfied a reading automaticity criterion. Armengol (5) examined Mexican children from both private and public school, and found that an increase in Stroop interference between the ages of 6 and 7 was present in public but not in private school children, the latter being more proficient in reading at the age of 6 and consequently showing a larger degree of interference from the outset. Similar results were obtained by Bonino and Ciairano (7) on an Italian sample. Their study shows that interference decreases comparing age 7 with age 9 and age 9 with age 11. In further agreement with the issue that Stroop interference is related to reading ability, there are indications that dyslexics show a larger Stroop interference than controls of the same chronological age, but roughly the same interference as reading age controls (15).

By assuming that the amount of interference can be predicted on the basis of reading proficiency, one should expect to find no interference at all in prereading children. However, even if they cannot grasp the meaning of written words, relatively old prereading children are usually able to distinguish alphabetical from non-alphabetical stimuli, thus one could expect to find some degree of interference also in preschool age. To the best of our knowledge, to present time, the few attempts made to present prereading children with Stroop-like situations did not use verbal stimuli. Santostefano (25) first reported a relatively strong interference effect among kindergarten children presented with a non-verbal extension of the Stroop test, named Fruit Distraction Test, which required children to name the color of drawings representing correctly and incorrectly colored fruits. More recently, Gerstadt et al. (17) developed another test to study inhibitory functions in prereading children, the Day-Night test. In this test, children are required to say "day" when faced with a nighttime picture, and "night" in response to a daytime picture. Although devised as a non-verbal extension of the Stroop test, the Day-Night test is rather different from the Stroop test in terms of cognitive demands, both because of the limited set size of the stimuli, and because children are explicitly required to inhibit the most compelling response, and to provide a counterintuitive response. Finally, Wright and co-workers (30) reported an interference effect among a large sample of children ranging in age from 3 to 16 years, by presenting them with chimeric drawings made by juxtaposing the head of an animal with the body of a different one. So far indeed the youngest children presented with a color-naming task using verbal stimuli were those attending the first grade of the primary school in the study of Schiller (26), who were able to read. Thus, the amount of the classic Stroop effect in prereading children remains to be investigated.

Analogously, none of the previous developmental studies appears to have investigated the reverse Stroop effect, i.e. the interference of color with reading (27), in the appropriate conditions. To cite MacLeod (19), "It is surprising how few studies investigating interference have been concerned at all with Stroop's Experiment 1: reading words in black versus incongruent colors". Instead, since color naming is usually acquired before reading, it is of particular interest to study whether at the beginning of reading acquisition color interferes with reading more or less than in skilled readers; and, on the other side, whether at this stage young children are facilitated in reading color names by the congruence between names and colors.

In the present study, we aimed at a careful characterization of the Stroop effects in relation with reading proficiency. We therefore presented a large sample of primary school children, ranging in age from 6 to 10 years, with a color-naming task using verbal stimuli; and extended our sample to study with the same verbal stimuli even prereading, nursery school children. Primary school children underwent also a word-reading task: meanwhile, we produced normative data for both classic and reverse Stroop effects across ages and grades. Moreover, we introduced a congruent condition (in which word meaning and color coincided), and a baseline color-naming condition in which pseudo-alphabetical strings, intended to match the visual appearance of letter strings, were employed in addition to the more commonly used geometrical shapes. In an attempt to normalize the presentation, the stimuli were displayed on a computer screen for a controlled time duration.

METHODS

Participants

Two hundred and forty-one children (113 males, and 128 females), ranging in age from 3 years and 6 months to 11 years and 3 months, participated in the study. Data collection was done during the second term of the scholastic year (March-June). According to the Italian laws, 3- and 5-year-old children were in their first and respectively last year of nursery school, 6-year-olds were in the first grade of primary school, and so on.

Since we aimed at a clear-cut distinction between readers and prereaders, the reading ability of nursery school and first grade children was assessed prior to the experimental session. This was done by asking them to read the four color names used as stimuli, as well as 8 high frequency Italian words (see Appendix) taken from a list used to assess the reading proficiency of first grade children at the intermediate level (22). All the words were printed in black block capitals on a white background.

Primary school children underwent both the *Color-Naming* and the *Word-Reading* tasks (see below), whereas children attending nursery school, who, with a few exception, were unable to read, were scheduled to perform only the *Color-Naming* task. Although the experiment was run on all the above children, only data from those meeting the following inclusion criteria were analyzed: a) to have been exposed to Italian language since birth; b) not to suffer or have suffered from neurological or psychiatric problems; c) to have normal or corrected-to-normal visual acuity, and normal color perception; d) (only for children attending the primary school) to have achieved a satisfactory proficiency in reading; e) (only for children attending the nursery school) to prove to be unable to read. Thus, we discarded data from a) seven non-native Italian speakers, b) two children suffering from neurological impairments and six children undergoing speech therapy for language disorders, and c) one pupil with impaired color perception. Finally, data from the few children escaping the opposite inclusion criteria adopted for readers and prereaders (six 6-year-olds attending the first-grade but still unable to read at the time of the experiment, as well as seven 5-year-old children who proved to be able to read) were removed from the main analysis, but will be

Table I. - *Demographic characteristics of the sample.*

	n	M	F	Age (years and months)			
Nursery School (in brackets: Grade)				Mean	*(SD)*	Min	Max
3-year-olds (I)	13	3	10	3y 11m	*(3.6 m)*	3y 6m	4y 5m
5-year-olds (III)	36	12	24	5y 10m	*(2.9 m)*	5y 5m	6y 5m
Subtotal	*49*						
Primary School (in brackets: Grade)							
6-year-olds (I)	23	15	8	6y 8m	*(4.2 m)*	6y 3m	7y 3m
7-year-olds (II)	38	17	21	7y 10m	*(3.7 m)*	7y 3m	8y 2m
8-year-olds (III)	38	16	22	8y 8m	*(3.5 m)*	8y 3m	9y 3m
9-year-olds (IV)	26	13	13	9y 9m	*(3.4 m)*	9y 3m	10y 3m
10-year-olds (V)	38	15	23	10y 8m	*(3.4 m)*	10y 3m	11y 2m
Subtotal	*163*						
TOTAL	212						

referred to in the Discussion. The main analyses were therefore performed on 212 children for the *Color-Naming* task, and 163 children for the *Word-Reading* task. Demographic characteristics of the analyzed sample are given in Table I.

Apparatus

A computer-assisted version of the Stroop paradigm, implemented by means of a commercial software program (E-Prime, Psychology Software Tools, Inc.), was used. All the experimental sessions were conducted on an IBM compatible PC. Stimuli were displayed on a 15-inch monitor. The acoustic warning signals were delivered from an acoustic box placed behind the monitor. A microphone connected to the computer recorded voice onset times, while the computer keyboard was used to enter the participants' answers and to run the experiment. Finally, a head and chin rest was used to ensure that a fixed distance was maintained between the participants' eyes and the center of the monitor.

Stimuli and Task

Four different categories of experimental stimuli were used: a) the Italian names of 4 colors [i.e., BLU (blue), ROSSO (red), VERDE (green), and GIALLO (yellow)], printed either in the corresponding color [e.g., the word ROSSO (red) printed in red] or in one of the three conflicting colors [e.g., the word ROSSO (red) printed in blue, green or yellow]; b) the Italian names of 4 colors [i.e., BLU (blue), ROSSO (red), VERDE (green), and GIALLO (yellow)] printed in black; c) rectangular shapes (2.5 x 0.6 cm in size) fully colored in blue, red, green, and yellow; d) pseudo-alphabetic strings (pseudowords): printed in blue, printed in red, printed in green, and printed in yellow. The color names, both black and colored, were written in block capitals (font: SWISS 721 HvBT, size 25). They were 0.6 cm tall, and ranged in length from 1.6 cm for the word BLU, to 3.3 cm for the word GIALLO. Similarly, the pseudowords were written in 0.6 cm tall characters and, on the computer screen, ranged in length from 1.6 cm (the stimulus printed in blue), to 3.3 cm (the stimulus printed in yellow). All the stimuli were displayed singly in the center of the screen, against a white background.

In the *Color-Naming* task, the participants were required to name the color of the stimulus displayed on the screen, regardless of the category to which it belonged. Forty-eight stimuli were presented, divided into 3 blocks of trials. Each block involved the presentation of stimuli belonging to one of the above categories: in the first block, each of the 4 colored rectangles was displayed twice for a total of 8 trials; in the second block, each of the 4 colored pseudowords was displayed twice for a total of 8 trials. In the last block, the 8 congruent, and 24 incongruent colored words, as described above, were displayed in a fixed, predetermined, pseudo-random sequence. Thus, to sum up, in this task four different stimulus types were presented: rectangles, pseudowords, congruent, and incongruent words.

In the *Word-Reading* task, the participants were required to read aloud the word presented on the screen, regardless of the color of the stimulus. Forty stimuli were presented, divided into 2 blocks of trials. In the first block, the 4 color names printed in black were each presented twice for a total of 8 trials; in the second block, 32 colored words were presented in another fixed, predetermined, pseudo-random sequence: in 8 cases, word meaning and ink color coincided (congruent trials), while in the remaining 24 cases, the meaning of the word conflicted with the color of the ink (incongruent trials). Congruent and incongruent trials were intermingled. Thus, to sum up, in this task, participants were presented with three different stimulus types: black, congruent, and incongruent words.

Procedure

In a sound- and light-attenuated room, the subject sat in front of a computer monitor with his/her head positioned in a head- and chin-rest so that the distance between the eyes and the center of the monitor was always 57 cm, and centimeters corresponded to degrees of visual angle.

Each trial proceeded as follows: as the child fixated a black cross at the center of the screen (fixation point), the examiner initiated the trial by pressing the space bar on the computer keyboard; an acoustic warning signal prompted the child to look at the fixation point, and to maintain fixation until after the presentation of the stimulus, that started 200-500 ms after the sound. The stimulus duration was 150 ms, an exposure too brief to allow eye movements to occur during stimulus presentation. Nevertheless, the examiner monitored the maintenance of fixation, and trials with failure to fixate were aborted. Children were required to verbally report the color of the stimulus displayed in the *Color-Naming* task, and to read aloud the word presented in the *Word-Reading* task. Both speed of response and accuracy were encouraged. Voice onset latencies were measured from stimulus offset by means of a microphone connected to the computer, and reaction times (RTs) calculated by adding stimulus duration. In addition, the examiner, using the keyboard, entered the children's answers, so that errors, inappropriate utterances, and extraneous noises, which triggered the microphone erroneously, were recorded in order to be subsequently filtered from the data. As the examiner keyed in the children's answer, the cycle of events was repeated. Before starting the experiment, a few practice trials with a stimulus duration of 3000 ms were performed to ensure that the child had understood the procedure. Each child underwent a single test session in which both tasks, *Color-Naming* and *Word-Reading*, were performed; the order of the two tasks was systematically counterbalanced across participants.

In each task (*Color-Naming* and *Word-Reading*), both accuracy and speed of response were analyzed.

For each child, the number of errors was taken as a measure of accuracy on both tasks. Since the different types of stimuli were not equiprobable, percentages of errors rather than raw scores were considered. Inappropriate utterances, and extraneous noises that triggered the microphone erroneously (see above), as well as anticipations and delays (see below) were considered as spurious errors and filtered from the data. As a consequence, only errors due to interference [e.g. reporting red in response to the word (red) printed in blue in the *Color-Naming* or, respectively, reporting blue in response to the word (red) printed in blue in the *Word-Reading*] were considered as a genuine measure of accuracy, and entered into the analyses.

Analogously, for each child, the median correct RT was taken as a measure of the speed of response to the different stimulus types. RTs shorter than 250 ms and longer than 2500 ms were regarded as anticipations and delays, respectively, and discarded.

COLOR-NAMING TASK

As mentioned above, data from 212 children (49 attending nursery school, 163 attending primary school) were considered.

RESULTS

Accuracy. All the subjects perfear ceiling when requested to name the color of three out of four types of stimuli, namely: rectangles, pseudowords, and congruent words. Indeed, the rate of errors on these types of stimuli was absolutely negligible, ranging from 0.8% for rectangles and congruent words to 1.1% for pseudowords, thus accuracy in naming the color of these three types of stimulus will not be further considered. In contrast, the overall performance of the seven groups of children in naming the color of incongruent words was far from flawless (11.2% of errors). To investigate whether accuracy varied across the different groups of children, for each subject, the percentage of errors in response to the incongruent words types was entered in an univariate ANOVA with Age as the between subjects factor (seven levels: from 3 up to 10 years).

The factor Age was significant [$F(6,211) = 10.312$; $p < .001$], with prereaders 3- (5.1% of errors) and 5-year-old children (2.1% of errors) giving the most accurate performance. In details, *post-hoc* comparisons demonstrated that the percentage of errors in the incongruent condition was significantly smaller in 3- than in 6-, 7- and 9-year-old children ($p < .05$ in each case), as well as in 5-year-old pupils than in all the groups of older children (mean accuracy = 14.2%, $p < .02$ in each case), which, in turn, did not differ from each other.

Speed of response. For each subject, the median correct RT to the different types of stimulus was entered in a repeated measures ANOVA with Age (seven levels: from 3 to 10 years) as the between subjects factor, and Type of Stimulus (four levels: rectangles, pseudowords, congruent, and incongruent words) as the within subjects factor. Bonferroni correction was used for multiple comparisons. Table II summarizes speed data for each group of participants.

This analysis confirmed the existence of an Age effect [$F(6,205) = 24.454$; $p < .001$]: a *post-hoc* comparison demonstrated that the overall speed of response increased with age, with the sole exception of 5-year-old nursery school pupils who performed significantly faster than 6-year-old children. However, comparison did not reach significance in all cases. In details, younger were significantly slower than older children in these cases: 3- *vs.* 5-, 7-, 8-, 9- and 10-year-old children; 5- *vs.* 9- and 10-year-old children; 6- *vs.* 8-, 9- and 10-year-old children; 7- and 8- *vs.* 9- and 10-year-old children, $p < .05$ in each case.

The factor Type of Stimulus was also significant [$F(3,615) = 381.181$; $p < .001$]. RT was significantly slower in response to incongruent than congruent words; in turn, RT in response to the presentation of congruent words was longer than RT in response to pseudowords, which was longer than RT to rectangles, $p < .001$ in each case.

However, the significance of the interaction Type of Stimulus by Age [$F(18,615) = 16.788$; $p < .001$] indicated that the difference between RTs to the different stimulus types varied across the seven groups of pupils. The different components of this interaction are illustrated in Figure 1.

Indeed, although subsequent *t*-tests demonstrated a clear Stroop effect within each group of participants (namely, within each group RT in response to incongruent

Table II. - *Color-Naming Task: Speed of response to the different stimulus types (RT and SD in ms).*

	Age group (years)							
	3 Mean *(SD)*	5 Mean *(SD)*	6 Mean *(SD)*	7 Mean *(SD)*	8 Mean *(SD)*	9 Mean *(SD)*	10 Mean *(SD)*	Mean
RT *Rectangles*	864.4 *(133.0)*	689.7 *(102.5)*	649.1 *(85.3)*	590.9 *(74.1)*	614.2 *(119.2)*	533.9 *(70.6)*	511.4 *(68.9)*	961.7
RT *Pseudowords*	953.2 *(145.2)*	739.2 *(124.1)*	738.2 *(81.7)*	673.3 *(107.7)*	661.2 *(96.4)*	568.4 *(86.1)*	578.6 *(74.5)*	836.1
RT *Congruent Words*	1013.6 *(159.2)*	820.0 *(145.9)*	913.8 *(212.7)*	856.6 *(205.0)*	832.4 *(173.8)*	708.2 *(128.5)*	708.0 *(113.6)*	701.8
RT *Incongruent Words*	971.0 *(99.2)*	807.5 *(135.5)*	1156.6 *(216.8)*	1122.5 *(195.9)*	1002.8 *(183.8)*	875.4 *(112.3)*	796.3 *(102.1)*	636.2
Mean	950.6	764.1	864.4	810.8	777.7	671.5	648.6	

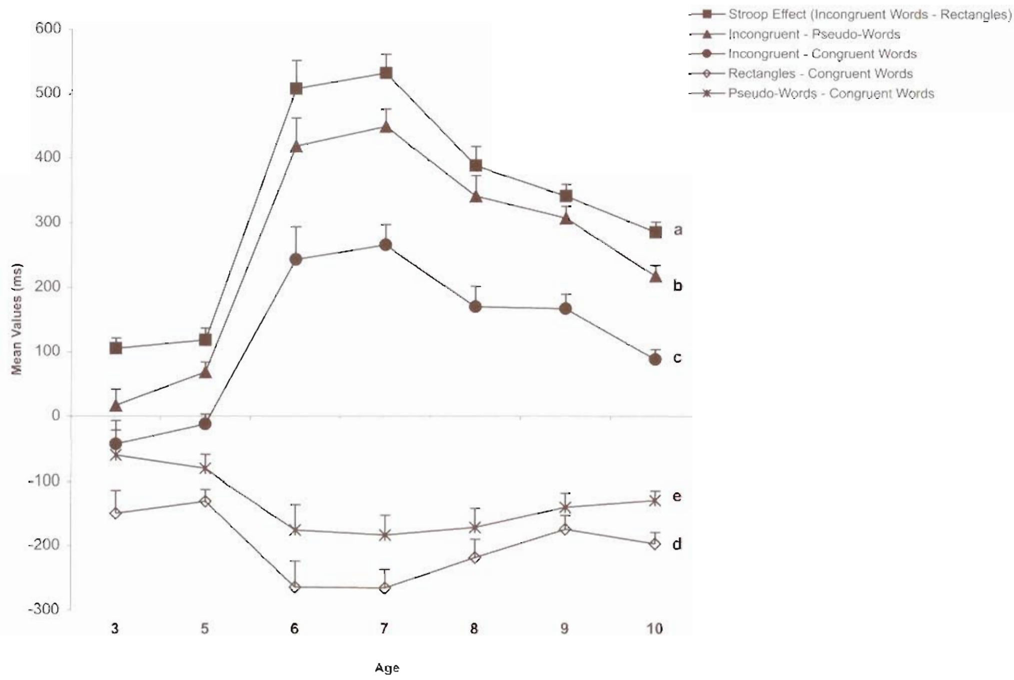

Fig. 1. - *Color-Naming*.
The different components of the interaction Type of Stimulus by Age. The lines a), b), and c) represent the three measures of interference: mean differences in ms between incongruent words and rectangles (the Stroop effect), a); between incongruent words and pseudowords, b); and between incongruent and congruent words, c), in the seven levels of education. The lines d) and e) represent the mean differences between rectangles, and respectively pseudowords, and congruent words (lack of facilitation). Error bars correspond to standard errors.

words was longer than RT to rectangles), the magnitude of this effect varied across the groups (Tab. II and Fig. 1, a): it was minimal among 3- (105.7 ms) and 5-year-old children attending nursery school (118.2 ms), very large among 6-year-old pupils (507.5 ms); it reached a maximal level among 7-year-old pupils (531.6 ms), and then progressively decreased (388.6 ms; 341.6 ms; 284.9 ms for 8-, 9-, and 10-year-old children, respectively). On an univariate ANOVA of the differences incongruent words-rectangles, this pattern of the Stroop effect across ages proved to be significant [F (6,205) = 36.143; $p < .001$], and to correspond to a significant quadratic trend; *post-hoc* comparisons were significant for 3- and 5-year-old children *vs.* all the primary school groups; for 6- *vs.* 9- and 10-year-old children, and for 7- *vs.* 8-, 9-, and 10-year-old children ($p < .01$ in each case).

A very similar pattern of results was obtained when entering the differences between RTs to incongruent words and to pseudowords (Tab. II and Fig. 1, b) into an univariate ANOVA [F (6,205) = 33.533; $p < .001$]; *post-hoc* comparisons were significant for 3- and 5-year-old pupils *vs.* all other children; for 6- *vs.* 10-year-old children, for 7- *vs.* 8-, 9-, and 10-year-old children, as well as for 8- *vs.* 10-year-old children ($p < .05$ in each case). A *t*-test within each group of participants showed that all the groups of primary school children and 5-year-old pupils attending nursery school were significantly slower at responding to incongruent words than to pseudowords ($p < .001$ for each group), whereas the sole 3-year-old pupils group failed to show any difference between these two types of stimuli.

Also the difference between RTs to incongruent and congruent words across ages (Tab. II and Fig. 1, c) proved to be significant on an univariate ANOVA [$F(6,205) = 14.943$; $p < .001$]. As with the Stroop effect and the difference between incongruent words and pseudowords, this congruence effect corresponded to a significant quadratic trend, being smaller among 3- and 5-year-old children than in the pupils belonging to all the other age groups ($p < .05$ in each case), except 10-year-old children; the difference was also significantly smaller among 10-year-old than 6- and 7-year-old children ($p < .006$ in each case). A t-test within each group of participants demonstrated that all the groups of primary school children were significantly slower at responding to incongruent than congruent words ($p < .001$ for each group), whereas 3- and 5-year-old pupils attending nursery school failed to show any congruence effect.

In addition, no advantage (facilitation) was induced by congruent words *vs.* each of the two baseline conditions, but responses to congruent words were found to be slower than responses to both rectangles and pseudowords in all age groups (Tab. II and Fig. 1, d and e), the differences being minimal among 3- and 5-year-old children attending nursery school and maximal among 7-year-old children. This pattern of differences across ages was significant for both rectangles [$F(6,205) = 3.540$; $p = .002$] and pseudowords [$F(6,205) = 2.637$; $p = .017$] on further univariate analyses. Within-group t-tests showed that both differences were significant in all the groups of children ($p < .003$ in each case), with the exception of 3-year-old nursery school pupils whose responses to congruent words did not differ from those to pseudowords.

Finally, the difference between the two baseline conditions, rectangles and pseudowords, did not vary significantly across ages, and within-groups t-tests showed that if was significant for each group of children ($p < .02$ in each case). Therefore this difference did not contribute to the interaction (and does not appear in Figure 1).

DISCUSSION

With respect to the primary school children, our results further substantiate findings from classic studies (12, 26), as well as a more recent study on an age-discontinuous Italian sample (7): the Stroop effect (i.e. the interference due to the incongruence between names and colors) is maximal in children attending the first grades of primary school, and then gradually decreases.

The main, and in some ways unexpected, result of the present study, however, is that the Stroop effect was present not only in 6- to 10-year-old children attending primary school, but already in 5- and even 3-year-old, nursery school children (Fig. 1, a), who should be, and in fact proved to be, unable to read (apart from a few exceptions, excluded from previous analyses). Since even prereaders are frequently exposed to alphabetical material, this interference could derive from an at least rough awareness of the alphabetical nature of the stimuli, which is independent from the ability to catch the meaning of written words. Experimental data provided strong support to this interpretation. According to their status of prereaders, the percentage of errors made by 3- and 5-year-old children in naming the color of incongruent words was negligible when compared with the performance of older children. Moreover, both 3- and 5-year-old children showed a Stroop effect in absence of any effect of congruence (no difference in RTs to congruent *vs.* incongruent words). Therefore, the Stroop effect shown by prereaders must represent a conflict between the color and the verbal form of the stimuli, instead of a conflict between the color and the verbal meaning of the stimuli like in readers. We will come back to this point in the General Discussion.

That the amount of interference can be predicted more on the basis of reading proficiency than on the basis of age alone (5, 15) is indeed supported by data from the 5-year-old children who were able to read, and the 6-year-old children who were unable to read (seven and six cases, respectively, all excluded from previous analyses): the former showed a Stroop effect larger than the average amount recorded in their age group: 252.7 *vs.* 118.2 ms, t (41) = 2.68, p = .01, while the contrary was true in the latter: 323.3 *vs.* 507.5 ms, t (27) = 1.71, p < .1. In addition, unlike other prereaders, the seven 5-year-old children able to read showed a clear congruence effect (i.e. the difference between RTs to incongruent and congruent words), even larger than that observed in the six 6-year-old children unable to read (mean value = 177.9 *vs.* 96.5 ms). The latter was far smaller than the mean of same age group readers, 242.8 msec.

WORD-READING TASK

As mentioned in the Methods section, we analyzed data from 163, 6- to 10-year-old children, attending the five grades of primary school.

RESULTS

Accuracy. All these children read extremely accurately all the different stimulus types displayed. Indeed, in all the conditions (black, congruent, and incongruent words) the percentage of errors was negligible (mean < 4%), thus accuracy scores will not be further considered.

Speed of response. Each subject's median correct RT to the different stimulus types was entered in a repeated measures ANOVA with Age (five levels: from 6 to 10 years) as the between subjects factor, and Type of Stimulus (three levels: black, congruent, and incongruent words) as the within subjects factor. Bonferroni correction was used for multiple comparisons. Table III summarizes speed data for each group of participants.

This analysis showed a clear Age effect [F (4,158) = 44.702, p < . 001]: a *post-hoc* comparison by means of two-sample t-test demonstrated that the overall performance of 6-year-old pupils was significantly worse than that of pupils in the other

Table III. - *Word-Reading Task: Speed of response to the different stimulus types (RT and SD in ms).*

	Age group (years)					
	6 Mean *(SD)*	7 Mean *(SD)*	8 Mean *(SD)*	9 Mean *(SD)*	10 Mean *(SD)*	Mean
RT *Black Words*	804.4 *(151.1)*	547.4 *(83.0)*	532.9 *(69.6)*	452.4 *(42.5)*	496.6 *(133.1)*	566.8
RT *Congruent Words*	752.6 *(102.4)*	560.4 *(80.2)*	567.3 *(72.2)*	465.4 *(56.9)*	504.5 *(127.0)*	570.1
RT *Incongruent Words*	794.7 *(141.7)*	566.0 *(78.8)*	564.9 *(80.2)*	476.7 *(47.1)*	509.5 *(144.4)*	582.4
Mean	783.9	548.0	555.1	464.9	503.5	

four age groups ($p < .001$ in each case). In addition, the overall performance of 9-year-old pupils was significantly better than the performance of 7- ($p = .001$) and 8-year-old pupils ($p = .002$).

The factor Type of Stimulus was also significant [F (2,316) = 4.692, $p = .013$], with children across the age groups found to be significantly slower at reading incongruent than black (the reverse Stroop effect, $p = .029$), and congruent words ($p = .013$) which, in turn, did not differ each other.

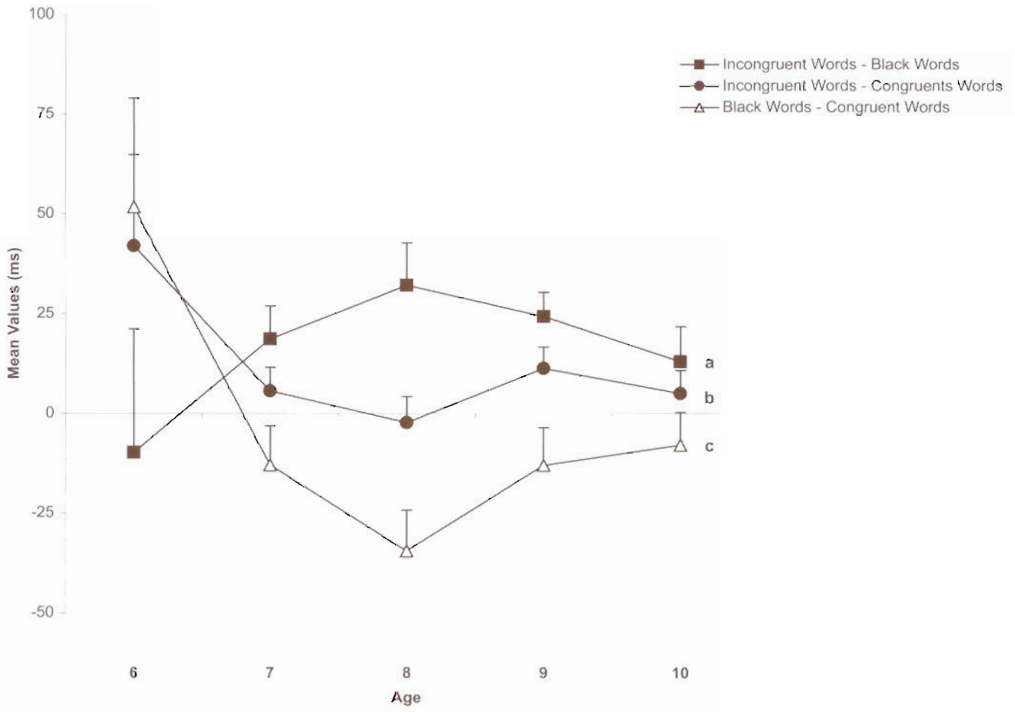

Fig. 2. - *Word-Reading.*
The lines a) and b) represent the two measures of interference: mean differences in ms between incongruent and black words (reverse Stroop effect), a), and between incongruent and congruent words, b), in the five groups of primary school children. A facilitation due to congruent words (positive difference between black and congruent words, c) is present only in 6-year-old children. Error bars correspond to standard errors.

However, as for the *Color-Naming* task, the significance of the interaction Type of Stimulus by Age [F (8,316) = 3.138, p = .003] indicated that the speed of reading the three different types of stimulus varied across the five groups of pupils.

First, the reverse Stroop effect (i.e. the advantage for black *vs.* incongruent words, see Table III and Figure 2, a) was present in all the pupils but the 6-year-olds (mean difference incongruent-black words = -9.8 ms, not significant). Such an effect reached significance in 7-, 8-, and 9-year-old children (mean difference = 18.6, 32.1 and 24.3 ms, respectively, $p < .05$ in each case), but not in the 10-year-old children (mean difference = 12.9 ms). Moreover, all the pupils but the 8-year-olds were faster at reading congruent than incongruent words (mean difference = 42.0, 5.6, -2.4, 11.3, 5.0 ms for the five age groups, see Table III and Figure 2, b); *post-hoc* comparisons reached significance among the 9-year-old pupils (p = .044). Finally, analogously to the pattern observed when comparing groups performance on black *vs.* incongruent words, we found that all the pupils ranging from 7 to 10 years were also slower at reading congruent than black words (mean difference black-congruent words = -12.9, -34.5, -13.0 and -7.9 ms, respectively, the difference being significant only among 8-year-old children, p = .002). On the opposite (Tab. III and Fig. 2, c), 6-year-old pupils were 51.8 ms faster at reading congruent than black words, albeit this difference only approached significance (p = .070).

DISCUSSION

As mentioned above – to the best of our knowledge – none of the previous developmental studies appears to have investigated the reverse Stroop effect in the appropriate conditions, i.e., the results of reading incongruently colored names minus the results of reading names printed in black (19). Therefore, our experiment represents the first attempt to present young children with a word-reading task properly devised to investigate whether the incongruence between names and colors interferes with reading, as well as, on the other side, whether the congruence between names and colors may facilitate reading. Results were straightforward. Consistently with the pattern originally observed in adults (27), pupils from 7 to 10 years, attending from the second to the fifth grade of primary school, respectively, were slower at reading incongruent than black words, clearly demonstrating the existence of an interference of color with reading (i.e. the reverse Stroop effect). Although a mild effect in quantitative terms, such an interference reached significance in all the children but the 10-year-olds, likely because of the more efficient inhibitory processes at work in older subjects. In contrast, 6-year-old children attending the first grade of primary school failed to show any reverse Stroop effect. In addition, they were the sole group displaying facilitation when reading congruent *vs.* black words. Although the beneficial effect of reading color names printed in a color that corresponded with the name only approached significance, this result suggests that congruence between word meaning and ink color was effective in improving the performance of children at the early stages of reading acquisition. We will come back to this point in the General Discussion.

An additional point concerning the reverse Stroop effect deserves some consideration. Recent studies in adult subjects (2, 3) showed that in a Stroop task there is an asymmetric cost of task switching: practice in color naming induces a reverse Stroop effect in a consecutive word reading, but not vice versa. Considering this possible order effect, we analyzed separately the data of the primary school children performing *Word-Reading* before and respectively after *Color-Naming*. In fact, the interference was smaller in the first than in the second group [5.0 *vs*. 29.2 ms, $t(161) = 2.09$, $p = .04$], indicating that, also in our developmental sample, practice in the more demanding task (i.e. inhibition of reading in color naming) is necessary for the interference to become evident in the successive word reading.

GENERAL DISCUSSION

Stroop interference has been estimated in different ways, and this may contribute to the differences across studies, critically reviewed by Salo et al. (24). In the original Stroop (27) study, the total time of the baseline condition (reading of 100 color names printed in black in the word-reading task, and naming of 100 colored squares in the color-naming task) was subtracted from the total time of the relative experimental condition (reading, or naming the color of, 100 color names, all printed in ink of a color incongruent with the name). To cite only a few examples, Comalli et al. (12), Bonino and Ciairano (7), and Armengol (5) adopted the same procedure, while Schiller (26) and Everatt et al. (15) used the total experimental time: baseline time quotient in order to reduce the variability, and other, recent studies used even more indirect scores (1, 6, 18).

In our study, as in previous ones (19, 24), the amount of interference was calculated on the basis of single responses to each stimulus, subtracting the appropriate baseline conditions. In addition, stimuli were presented one at a time for a short, fixed duration and in central vision following a warning signal, so that each subject viewed them for the same amount of time, with as little distraction as possible and no need of scanning movements. Moreover, in this study a condition of congruence was introduced in both tasks, thus allowing us to estimate possible facilitatory effects related to the correspondence between color and word. Finally, pseudowords were used in addition to rectangles as the baseline for the *Color-Naming* task.

The new data we found was that a Stroop interference is already present in prereading children, as a result of the interference between the color and the verbal form more than the verbal meaning, since we found no difference between their responses to congruent and incongruent words. Six-year-old children attending the first grade of primary school display a full Stroop effect but no reverse Stroop effect, are slower in reading than in naming (see below) and benefit from congruence in reading. The Stroop effect continues to increase from 6- to 7-year-old children, then diminishes from 7- to 10-year-old children, in agreement with the developmental changes of reading abilities and executive functions during childhood. In the following, we will discuss in some detail these points by age groups and across tasks, then conclude with perspectives for reading disables.

Three- and 5-year-olds (prereaders)

Formally, the delay in naming the color of incongruent words with respect to geometrical shapes in prereaders can be taken as a true Stroop effect. However, as expected, and already discussed, 3- and 5-year-old, unable to read children, did not show any difference between incongruent and congruent stimuli in the *Color-Naming*. Therefore, the delay in naming the color of incongruent words with respect to geometrical shapes in prereaders should be better considered an apparent Stroop effect, in the sense that it is independent from the meaning of the verbal stimuli. On the other side, color-naming ability is so well acquired before learning to read that 5-year-old, nursery school children were practically as quick and accurate as 6-year-old children in the baseline conditions of the *Color-Naming* task (Tab. III). It would remain to establish whether, even when color naming only was required, prereaders processed letter strings less efficiently than the baseline stimuli because of some awareness of their verbal nature, or because of purely visual reasons. Our data support a specific less efficient processing of letter strings, at least in the group of relatively older prereaders. We base this conclusion on the performance of prereaders on pseudo-alphabetical strings, that were purposefully drawn to match the visual appearance of letter strings: in 3-year-old children RTs to alphabetical and pseudo-alphabetical strings were statistically indistinguishable, while in 5-year-olds, like in primary school children, RT difference was larger and significant. This result seems indeed to suggest some verbal awareness at age 5 but not at age 3. On the other hand, like older children, both groups of prereaders responded to pseudowords more slowly than to rectangles, suggesting that this difference was based on purely visual features.

Six-year-olds

According to the fact that the 6-year-old children were in the early stages of reading acquisition, their performance is very peculiar. In particular, they were the only group which was faster in performing the "pure" naming task (rectangles) than the "pure" reading task (black words), compare Table II and Table III and see Schiller (26), Figure 1, for a similar result. Coherently, the 6–year-old children were the only group displaying no reverse Stroop effect, but some facilitation (i.e. a benefit from reading congruent *vs.* black words) in the Word Reading task. Their poor reading skills may account for both the lack of any reverse Stroop effect and the trend for facilitation. Given that their reading performance was near to floor level, no interference effect can be expected by presenting them with incongruently colored color names. On the other hand, when requested to read congruently colored color names, 6-year-old children may solve the task by simply telling the color of the words, that is easier and faster for them. However, the large Stroop effect found in the *Color-Naming* task demonstrates that, albeit poor, reading is an inescapable process also for this group of children. In fact, the delay in naming the color of incongruent words with respect to geometrical shapes, that is already present in 3- and 5-year-olds, further increases from 5- to 6-year-olds, finally becoming a true Stroop effect in this age group.

Seven- to 10-year-olds

Seven-year-olds are at the top of the inverted U-shaped curve depicting the Stroop effect in our overall sample. The interference increases from 6- to 7-year-olds, then decreases. As we wrote in the Introduction, this pattern is best interpreted as resulting from two developmental changes displaying a different time-course: increasing automation in reading, and increasing maturation of executive-attentional functions (11). Therefore, Stroop interference increases until reading automation prevails over the controlled processes, then decreases.

In the case of the reverse Stroop Effect, we cannot compare our results to the existing developmental literature, since none of the previous developmental studies appear to have investigated the reverse Stroop effect in the appropriate conditions, i.e., the results of reading incongruently colored names minus the results of reading names printed in black (19). As originally reported in adults (27), we found that the reverse Stroop effect, i.e. the interference of color with reading, was far smaller than the classic effect (compare the performances of 7- to 10-year-old children in Figure 1, a, and Figure 2, a). Moreover, the two effects proved not to be statistically correlated across the participants (Pearson test, $p = .147$, among all the primary school children). In fact, while the classic Stroop effect was significant in all age groups, the reverse effect did not reach significance in 10- and, as already underlined, was absent in 6-year-old children. These results are in keeping both with the fact that in skilled readers reading is generally faster and more automatic than naming, and with the more efficient inhibitory processes at work in older children.

Due to the fact that responses to congruent stimuli were generally slower than responses to baseline stimuli (see below, lack of facilitation), the difference between incongruent and congruent stimuli was generally smaller than that found between responses to incongruent and baseline stimuli, that are the classic Stroop effect in the *Color-Naming* and the reverse Stroop effect in the *Word-Reading*, respectively (Fig. 1, a and c, and Fig. 2, a and b).

Concerning facilitation, by introducing congruent stimuli (corresponding name and color) in the two tasks, we might expect that the congruence would facilitate the response, therefore producing general benefits somewhat symmetrical to the costs produced by the incongruent *vs.* the baseline conditions. Instead, the benefit displayed by 6-year-olds in reading congruently colored color names is the only instance of facilitation, and is likely to be a spurious effect, depending on simply telling the color of the words, rather than a genuine facilitatory effect based on a sort of synergy between reading and color processing. All the five primary school age groups in the *Color-Naming*, and the present four groups in the *Word-Reading* task, see Figure 1, c, and Figure 2, d and e, failed to show any facilitation attributable to the color-name congruence, but responded more slowly to congruent than to baseline stimuli. Facilitation in the color-naming task has been reported in adult studies (19, 20, 21, 28), and its absence here might seem, at a first glance, to be a specific feature of the age groups we studied. In fact, we are not aware of any report of facilitation for this age range, the only partial overlap being with the 9-12 year olds studied by Carter et al. (9). On a closer investigation, however, facilitation appears to have been both reported and denied in the adult literature. To cite once more the review of MacLeod (19), "Facilitation is not a necessary concomitant of interference ... Congruence between the irrelevant word and the to-be-named ink color often produces facilitation. However, this facilitation is much less than the corresponding interference in the incongruent condition, and the choice of control condition may be crucial". In fact, among those finding the strongest facilitatory effects, both Tzelgov et al. (28) in adults, and Carter et al. (9) in children, used non-color words instead of color patches as a baseline, and this may be the reason why responses to congruent colored words were relatively facilitated. In conclusion, we can only speculate that the lack of facilitation in the *Color-Naming* task in our study was specifically age-related, being possibly due to the inescapable process of reading both congruent and incongruent words, that is however less efficient in children than in adults. We already discussed the special case of the *Word-Reading* task in the 6-year-olds.

Perspectives

We have shown how Stroop effects increase together with reading abilities, then decrease during the maturation of executive functions, in the normal population. Reading abilities and executive functions are both impaired in developmental dyslexia, therefore the present results may be used as normative data for studying developmental reading impairments across the crucial ages of reading acquisition. Although the effects of poor reading and poor control would seem to annihilate each other, our preliminary data indicate that Stroop interference is larger in developmental dyslexics than in normals (16).

SUMMARY

We studied 241 nursery and primary school 3- to 10-year old children with a discrete-trial version of the Stroop Test, including both a *Color-Naming* and (for the 6- to 10-year olds only) a *Word-Reading* task. The classic Stroop effect was present across all the ages, with an inverted U-shaped pattern: increasing from 3- to 7-year olds, then decreasing. Preschool children who were able to read showed a Stroop effect larger than same-age, unable to read children. The reverse Stroop effect was present across all the studied ages but the 6-year-olds, who instead displayed some facilitation in reading congruent *vs.* black words. Since the acquisition of reading skills turned out to be crucial for the Stroop effects, the present research may be useful to study developmental reading impairments by providing normative data.

APPENDIX

The 8 high frequency Italian words taken from a list used to assess the reading proficiency of first grade children at the intermediate level (Martini, 1995). All the words were written in block capitals, printed in black and presented in the center of the computer screen.

LUNA (MOON)
MARE (SEA)
ROSA (ROSE)
FILO (THREAD)
PORTA (DOOR)
STUFA (STOVE)
SAPONE (SOAP)
FAVOLA (FABLE)

Acknowledgments. - We thank children and teachers of the nursery school in Villaggio Eremo, Curtatone (Mantova), the nursery school "Angeli custodi", Palazzina (Verona) and the primary school "Berto Barbarani", San Martino Buon Albergo (Verona); Barbara Marocchi and Elena Rubini for their help in collecting data, Gianni Finizia for technical assistance and Marco Veronese for graphic help. This research has been supported by grants from the "Ministero dell'Istruzione, dell'Università e della Ricerca".

REFERENCES

1. Adleman, N.E., Menon, V., Blasey, C.M., White, C.D., Warsofsky, I.S., Glover, G.H. and Reiss, A.L. A developmental fMRI study of the Stroop Color-Word Task. *NeuroImage*, **16:** 61-75, 2002.
2. Allport, A., Styles, E.A. and Hsieh, S. Shifting intentional set: Exploring the dynamic control of tasks. Pp. 421-452. In: Umiltà, C.A. and Moscovitch, M. (Eds.) *Attention and Performance XV*. Cambridge MA, MIT Press, 1994.
3. Allport, A. and Wylie, G. Task switching, stimulus-response bindings, and negative priming. Pp. 35-70. In: Monsell, S. and Driver, J. (Eds.) *Attention and Performance XVIII*. Cambridge MA, MIT Press, 2000.
4. Archibald, S.J. and Kerns, K.A. Identification and description of new tests of executive functioning in children. *Child Neuropsychol.*, **5:** 115-129, 1999.
5. Armengol, C.G. Stroop test in Spanish: Children's norms. *Clin. Neuropsychol.*, **16:** 67-80, 2002.
6. Barbarotto, R., Laiacona, M., Frosio, R., Vecchio, M., Farinato, A. and Capitani, E. A normative study on visual reaction times and two Stroop color-word tests. *Ital. J. Neurol. Sci.*, **19:** 161-170, 1998.
7. Bonino, S. and Ciairano, S. Effetto Stroop e capacità di inibizione: una ricerca in campo evolutivo. *Giornale Italiano di Psicologia*, **3:** 587-603, 1997.
8. Brown, T.L., Roos-Gilbert, L. and Carr, T.H. Automaticity and word perception: Evidence from Stroop and Stroop Dilution Effect. *J. Exp. Psychol. Learn. Mem. Cogn.*, **6:** 1395-1411, 1995.
9. Carter, C.S., Krener, P., Chaderjian, M., Northcutt, C. and Wolfe, V. Abnormal processing of irrelevant information in attention deficit hyperactivity disorder. *Psychiatry Res.*, **56:** 59-70, 1995.
10. Carter, C.S., Mintun, M. and Cohen, J.D. Interference and facilitation effects during selective attention: an $H_2^{15}O$ PET study of Stroop task performance. *NeuroImage*, **2:** 264-272, 1995.

11. COHEN, J.D., DUNBAR, K. AND MCCLELLAND, J.L. On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychol. Rev.*, **97:** 332-361, 1990.
12. COMALLI, P.E.JR., WAPNER, S. AND WERNER, H. Interference effects of Stroop color-word test in childhood, adulthood and aging. *J. Genet. Psychol.*, **100:** 47-53, 1962.
13. COX, C.S., CHEE, E., CHASE, G.A., BAUMGARDNER, T.L., SCHUERHOLZ, L.J., READER, M.J., MOHR, J. AND DENCKLA, M.B. Reading proficiency affects the construct validity of the Stroop test interference score. *Clin. Neuropsychol.*, **11:** 105-110, 1997.
14. DURGIN, F.H. The reverse Stroop effect. *Psychon. Bull. Rev.*, **7:** 121-125, 2000.
15. EVERATT, J., WARNER, J., MILES, T.R. AND THOMSON, M.E. The incidence of Stroop interference in dyslexia. *Dyslexia*, **3:** 222-228, 1997.
16. FACCIOLI, C., PERU, A. AND TASSINARI, G. Stroop effects from preschoolers to fifth graders in normals and dyslexics. Poster presented at the First Congress of the European Neuropsychological Societies, Modena, Italy, 2004.
17. GERSTADT, C.L, HONG, Y. AND DIAMOND, A. The relationship between cognition and action: Performance of children 3.5 to 7 years old on a Stroop-like day-night test. *Cognition*, **53:** 129-153, 1994.
18. HELLAND, T. AND ASBJØRNSEN, A. Executive functions in dyslexia. *Child Neuropsychol.*, **6:** 37-48, 2000.
19. MACLEOD, C.M. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.*, **109:** 163-203, 1991.
20. MACLEOD, C.M. Training on the integrated versus separated Stroop tasks: the progression of interference and facilitation. *Mem. Cognit.*, **26:** 201-211, 1998.
21. MACLEOD, C.M. AND MACDONALD, P.A. Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends Cogn. Sci.*, **4:** 383-391, 2000.
22. MARTINI, A. *Le difficoltà di apprendimento della lingua scritta: criteri di diagnosi ed indirizzi di trattamento*. Pisa: Edizioni del Cerro, 1995.
23. MONSELL, S., TAYLOR, T.J. AND MURPHY, K. Naming the color of a word: is it responses or task sets that compete? *Mem. Cognit.*, **29:** 137-151, 2001.
24. SALO, R., HENIK, A. AND ROBERTSON, L.C. Interpreting Stroop interference: an analysis of differences between task versions. *Neuropsychology*, **15:** 462-471, 2001.
25. SANTOSTEFANO, S. Cognitive controls and exceptional states in children. *J. Clin. Psychol.*, **20:** 213-218, 1964.
26. SCHILLER, P.H. Developmental study of color-word interference. *J. Exp. Psychol.*, **72:** 105-108, 1966.
27. STROOP, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.*, **18:** 643-662, 1935.
28. TZELGOV, J., HENIK, A. AND BERGER, J. Controlling Stroop effects by manipulating expectations for color words. *Mem. Cognit.*, **20:** 727-735, 1992.
29. WRIGHT, I., WATERMAN, M., PRESCOTT, H. AND MURDOCH-EATON, D. A new Stroop-like measure of inhibitory function development: typical developmental trends. *J. Child Psychol. Psychiatry*, **44:** 561-575, 2003.
30. WRIGHT, B.C. AND WANLEY, A. Adult's versus children's performance on the Stroop task: interference and facilitation. *Br. J. Psychol.*, **94:** 475-485, 2003.
31. YEE, P.L. AND HUNT, E. Individual differences in Stroop dilution: Tests of attention-capture hypothesis. *J. Exp. Psychol. Hum. Percept. Perform.*, **17:** 715-725, 1991.
32. ZYSSET, S., MÜLLER, K., LOHMANN, G. AND VON CRAMON, D.Y. Color-Word matching Stroop task: separating interference and response conflict. *NeuroImage*, **13:** 29-36, 2001.