ELECTRONIC JOURNAL OF
RESEARCH IN
EDUCATIONAL PSYCHOLOGY

No. 1 (2) 2003. ISSN: 1696-2095

Sensitivity of certain standardised tests to executive attention functioning in seven-year-old children

Luis J. Fuentes*, Carmen González**, Angeles F. Estévez*,
José A. Carranza**, Mayte Daza*, María D. Galián**,
and Dolores Álvarez*

* University of Almeria
** University of Murcia

Spain

lfuentes@ual.es

ABSTRACT

Introduction. In clinical and education environments, it is usual to assess attention skills through standardised tests. Because the concept of attention is rather vague, this study attempts to evaluate relationships between some of those tests and executive attention such as it is conceptualized in recent cognitive neuroscience approaches.

Method. Seven-year-old children were administered certain subtests of the WISC-R, the Trial Making Tests parts A and B, and the card version of the Stroop task. Then they carried out a computerized version of the Stroop task, which neuroimaging studies have associated with the neural circuitry involved in executive attention.

Results. Children scoring differently in only Picture Completion, Digits Backward, and TMT-B produced different patterns of Stroop interference effects. Also, scores in both versions of the Stroop task did not correlate.

Discussion. The results suggest that these tests are sensitive to executive attention functioning in schoolchildren. In addition, the lack of correlation between the two versions of the Stroop tests suggests that in the card version different attention mechanisms might predominate over those involved in the computerized version.

Keywords: Standardised tests, WISC-R, Trial Making Test, Stroop task, executive attention, counselling, neuropsychological testing.

Introduction

Certain standardised tests, such as the Weschler Intelligence Scales for Children (WISC), have become powerful instruments widely used in clinical and educational contexts to measure cognitive function in schoolchildren. In performing these tests, authors have attributed a relevant role to attention, along with other cognitive operations. However, the term attention today is rather vague, since recent neuro-cognitive approaches to cognition have thought of attention as a multi-component phenomenon that exerts control functions through the co-ordination of different networks. These networks would perform different and very specific attentional computations (Fuentes, in press). Some computations deal with orienting attention to particular locations in the visual field in anticipation of relevant targets. From a functional point of view, this orientation network of visual attention operates when attention is shifted to the location of a cue that signals where the target will come up. Disengage, movement and engage computations are involved in this kind of attention, depending on the posterior parietal lobe, the superior colliculus, and the pulvinar nucleus of the thalamus, respectively (for reviews, see Posner and Petersen, 1990; Posner and Raichle, 1994). This network is important because it helps the individuals search for potential relevant targets to foster further high-level processing of those targets. On the other hand, high-level processing related to cognitive functioning has been thought to be under the control of a higher form of attention known as the *executive attentional network*.

Executive attention involves activation of anterior areas of the brain, concretely portions of the dorsolateral prefrontal and midline frontal areas including the anterior cingulate cortex and the supplementary motor area. These areas activate in situations that involve planning, decision making, error correction, performing of responses not well-learned, or overcoming habitual (or automatic) responses (Posner and DiGirolamo, 1998), computations all that require executive control.

The purpose of this study was to assess whether some standardised scales that supposedly measure attention, as used in neuropsychological and counselling contexts, relate to the executive network of the attentional system in schoolchildren. The scales chosen from standardised tests were Picture Completion, Picture Arrangement, Digit Span, and the Mazes subtests of Weschler's Scales for Children Revised (the WISC-R; Weschler , 1974), and the

Trial Making Test (parts A and B). A computerized version of the Stroop task was chosen as a measure of executive attention.

Weschler's Scales for Children Revised (WISC-R)

Briefly, the WISC-R is conceived as a powerful instrument to assess general intelligence, and as a diagnostic tool to detect cognitive deficits in neuropsychological, counselling, and clinical environments. It consists of twelve subtests, 6 constitute the Verbal Scale, and 6 the Performance Scale. For present purposes we will briefly describe some of the subtests that have been thought to measure attention.

The *Picture Completion* subtest asks the participant to discover and name (or point to) the missing part of an incompletely drawn picture. The cognitive operations involved include visual scanning and visual search; recall of items in visual memory; comparison between the stimulus and the visually recalled image; and importantly, the ability of the examinee to select essential from non-essential details of the pictures.

The *Picture Arrangement* subtest asks the participant to place a series of pictures in the right sequence, so that when it is done correctly the pictures tell a little story. This measures the ability of the examinee to grasp the general idea of the story in order to be successful. The cognitive operations involved include planning ability, anticipation of the consequences of initial situations, visual organisation, and temporal sequencing.

The *Digit Span* subtest requires the child to repeat a series of digits given orally by the examiner. This test is a measure of the child's short-term auditory memory and attention. The subtest has two parts: *Digit Forward* and *Digit Backward*. Because these two parts seem to involve different components of working memory, most authors claim that they must be analysed separately (Lezak, 1995). In the Digit Forward part, children have to repeat series of digits ranging in length from two to nine, in the same order that was given by the examiner. This test involves primarily the auditory short-term storage of working memory. In the Digits Backward part, children have to repeat series of digits ranging in length from two to eight, in the reverse order to that given by the examiner. This test involves the active manipulation of stored information (i.e., the executive control system of working memory; Baddeley, 1986).

The *Mazes* subtest requires the child to solve paper and pencil mazes that differ in the level of difficulty. In each maze the examinee has to draw a line from the centre to the outside without crossing any of the lines that indicate walls. This test involves attention to the directions and appears to measure the child's planning ability and perceptual organisation ability.

The Trial Making Test (TMT)

The TMT was developed by the US Army in 1944. It consists of two parts. Subtest A (TMT-A) requires the examinee to connect randomly-located circles with numbers (1-25) in numerical order as quickly as possible. In subtest B (TMT-B) the task is to connect circles alternately with numbers (1-13) and with letters (A-L) in their respective sequence as quickly as possible. Previous research with neurological patients showed that deficient performance in TMT-B might be related to impaired inhibitory processing (Amieva et al., 1998). Inhibitory processing has been thought of as an important function of the executive attentional network (Fuentes, in press; Fuentes, Vivas, and Humphreys, 1999; Posner and Raichle, 1994).

The Stroop Task

The Stroop task (Stroop, 1935) has been amply used to measure selective attention in both neuropsychological and cognitive studies. In the task, participants are required to name the colour in which words are printed. In order to do that, participants have to attend to the relevant dimension of the word (the ink colour) while simultaneously ignoring its irrelevant but overriding dimension (the word meaning). When naming the colour of incongruent words (e.g., the word RED printed in blue colour), participants have difficulty ignoring the intrusive effects of the words, showing a deterioration in performance (longer reaction times, and /or more errors) compared with a neutral condition in which meaningless stimuli (e.g., a string of Xs) are used. Thus, the Stroop interference effect is computed by comparing performance in the incongruent condition with that in the neutral condition.

In the single-trial computerized version of the Stroop task, stimuli are individually presented in the centre of the computer screen. Studies of neuroimaging techniques (e.g., positron emission tomography, PET; functional magnetic resonance imaging, fMRI) have revealed that Stroop interference is associated with activation in the anterior cingulate cortex (Fan, Flombaum, McCandliss, Thomas, and Posner, 2003; Pardo, Pardo, Janer, and Raicle, 1990) or prefrontal cortex (Taylor, Kornblum, Lauber, Minoshima, and Koeppe, 1997), the brain areas that have been associated with executive attention (for a review, see Bush, Luu, and Posner, 2000). Stroop interference can be observed only from age 6-7 years, coinciding with the necessary reading skills that make semantic processing an automatic process, and with maturation of the frontal and prefrontal cortices involved in the executive control of information processing.

There are other versions of the original task employed by Stroop in 1935. In the card version, all Stroop stimuli are simultaneously presented in a card. The first card contains a list of words printed in black ink. Participants have to read the words. Stroop interference is calculated from the other two cards in which participants have to name the stimulus colours. One card contains strings of Xs in different colours and serves as the neutral condition. The other card contains colour-incongruent words and serves as the incongruent condition. Given that the computerized version has proved to reflect the operations of executive attention, and the card version is the most commonly used in neuropsychological and counselling studies, a comparison between the two forms of Stroop interference is specially relevant for determining whether the widely-used card version of the Stroop task is actually sensitive to executive control. If the card version measures executive attention, we should find a correlation with scores in the computerized version, which has already been proved to measure executive attention.

Briefly, 7-year-old children were administered the subtests of the WISC-R, the parts A and B of the TMT, and the card version of the Stroop task. Then, all of them performed the computerized version of the Stroop task to measure their executive attention functioning. The goal of this study was to investigate whether children scoring differently in the above tests show different patterns of Stroop interference, and therefore in the functioning of executive attention operations. This will help us to select the particular subtests of the above tests connected with executive attention when used in clinical settings.

Method

Participants

Two hundred and four children from the Spanish schools Francisco de Goya (Almería), Lope de Vega (Almería), and San Buenaventura (Murcia) participated. All children were aged 7 years and were enrolled in the second grade of primary school. None re-

Electronic Journal of Research in Educational Psychology. No. 1 (2), 23-36. ISSN: 1696-2095

ported problems in identifying colours. At testing time all had normal or corrected-to-normal vision. None had evidenced reading or learning problems.

Instruments and apparatus

For the standardised tests, four subtests from the WISC-R were administered to all children: Picture Completion, Picture Arrangement, Digit Span (Forward and backward parts) and Mazes. Children also completed parts A and B of the Trial Making Test, and the card version of the Stroop task (Golden, 1978; test distributed by TEA ediciones).

For the executive attention measure we used a computerized version of the Stroop task. In the task, targets were the words ROJO (red), AZUL (blue), and VERDE (green) and a string of four Xs, displayed in red, blue, or green colour. Stroop stimuli were presented one by one on a colour screen (VGA) of a IBM-compatible computer. Naming responses were recorded trial by trial through a voice-key interfaced to the parallel port of the computer.

Procedure

For the computerized Stroop task, the computer showed a plus sign lasting 500ms that served as the fixation point. The target display was then presented until the participant responded. Participants had to name the colour of the target and then the experimenter entered a code for later coding of performance accuracy. After 2000ms, a new trial began with the presentation of the fixation point. Participants carried out one block of 72 trials. The first 12 were practice trials and were not included in the data analysis. The remaining 60 were experimental trials, 30 for the neutral condition (10 trials per colour), and 30 for the incongruent condition (10 trials per colour). Reaction times (RTs) and error percentage were the dependent variables. Thus, Stroop interference was measured by subtracting RTs and error percentage in the neutral condition from the incongruent condition.

All standardised tests (WISC, TMT and Stroop cards) were administered according to standard procedures used in counselling and clinical environments.

Data analysis

For the Picture Completion, Picture Arrangement, Digit Span, and Mazes subtests, children were split into two groups according to the median score. Children scoring below or above the median were categorised as *low* or *high* on each test, respectively (see González, Fuentes, Carranza, and Estévez, 2001, for a similar analysis method). For the TMT, scores were 0 if children were unable to perform the task, 1 if they completed the task but it took longer than the established time limit (2 minutes and 30 seconds), and 2 if they completed the task within the time limit. Children with a score of 0 were categorised as *low* and those with a score of 2 were categorised as *high* for analysis purposes.

We also conducted correlational analyses to assess the relationships between the two Stroop versions. Difference-score measures of Stroop interference (incongruent – neutral) were computed on data from the two tasks to calculate the correlations.

Mixed 2 x 2 x 2 ANOVAs, with gender (boys and girls) and test score (low vs. high) as the between-subjects factors, and Stroop condition (neutral vs. incongruent) as the within-subjects factor, were conducted for the computerized Stroop task.

Results

General Stroop effect

Table 1 shows the mean of median RTs and percentage of errors for the computerized Stroop task. The main effect of condition was significant for both RTs, $F(1, 202) = 300.49$, $p < 0.0001$; and errors, $F(1, 202) = 68.73$; $p < 0.0001$. Incongruent stimuli produced longer RTs and more errors than neutral stimuli. That is, we observed the standard Stroop interference effect. Neither gender nor the Gender x Condition interaction proved significant. Pearson product-moment correlational analyses showed a poor relationship between the two Stroop effects ($r = 0.18$, for RTs; and $r = 0.012$, for errors).

Table 1. Mean of Median Reaction Times and Percentage of Errors for the Computerized Stroop Task, as a Function of Gender and Condition.

	Boys		Girls	
Condition	RT/NI	PE	RT/NI	PE
Neutral	944	3.8	962	3.0
Incongruent	1123	7.7	1173	6.9

Note. RT = reaction time; PE = percentage of errors.

Interactions between standardised tests and Stroop interference effect

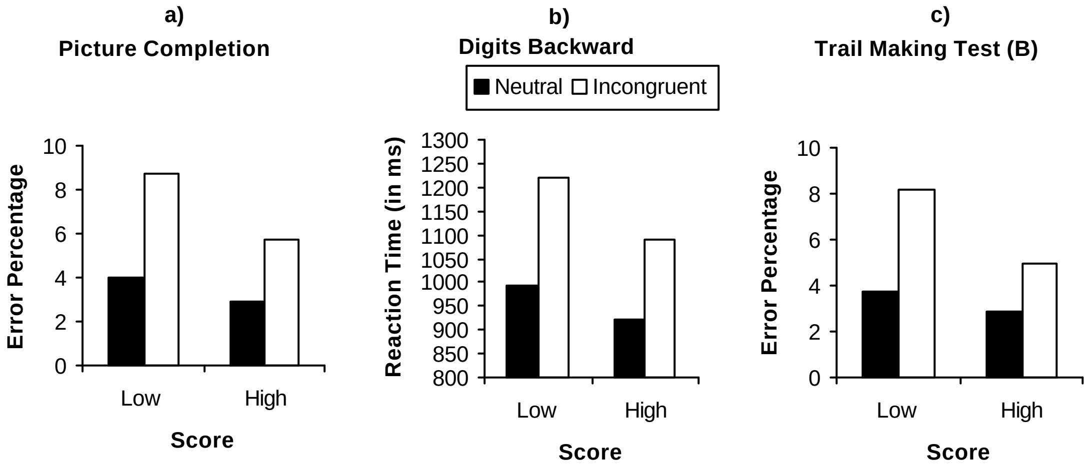
Reaction times and percentage of errors from the computerized version were taken as dependent variables for the following analyses. Given that in these analyses gender did not interact with any other factor, data from this variable were collapsed. For the sake of simplicity, data are presented only for children scoring low or high in each one of the subtests in which a significant Stroop effect x Test interaction was found (see Figure 1).

For the Picture Completion test, the error percentage analysis showed significant main effects of score and condition, $F(1, 184) = 9.99$, $p < 0.01$; and $F(1, 184) = 62.84$, $p < 0.001$, respectively. Children scoring low produced more errors than children scoring high (6.4 vs. 4.3%), and the incongruent condition produced more errors than the neutral condition (7.2 vs. 3.4%). Importantly, the Score x Condition interaction was also significant, $F(1, 184) = 3.78$, $p = 0.05$. The analysis of the interaction showed that children scoring low produced a greater Stroop interference than children scoring high (see Figure 1a).

For the Digit Backward test, the analysis of the RTs showed significant main effects of both score and condition, $F(1, 202) = 13.20$, $p < 0.01$; and $F(1, 202) = 310.60$, $p < 0.0001$, respectively. Children scoring low produced longer RTs than children scoring high (1108 vs. 1005 msec), and the incongruent condition produced longer RTs than the neutral condition (1156 vs. 958 msec). The Score x Condition interaction was also significant, $F(1, 202) = 5.89$, $p < 0.025$. Again, children scoring low in this test showed greater Stroop interference than children scoring high (see Figure 1b).

For the TMT-B test, the error percentage analysis produced significant main effects of both score and condition, $F(1, 142) = 7.21$, $p < 0.01$; and $F(1, 142) = 33.08$, $p < 0.001$, respectively. Children scoring low produced more errors than children scoring high (6.0 vs. 3.9%), and the incongruent condition produced twice as many errors as the neutral condition (6.6 vs. 3.3%). As before, the Score x Condition interaction proved significant, $F(1, 142) = 4.12$, $p < 0.05$. Children scoring low in the TMT-B showed greater Stroop interference than children scoring high (see Figure 1c).

Figure 1. Stroop effect for the computerized version as a function of score in the Figure Completion (a), Digit Backward (b), and Trial Making Test B (c) tests.



No other tests produced differential Stroop effects as a function of score, neither with RTs nor with errors.

Discussion

The present work was aimed at assessing the relationships between some standardised tests that are thought to measure attention, and executive attention as measured by the computerized version of the Stroop task. The reason for studying such relationships is twofold. First, tests such as the WISC-R, the TMT and the card Stroop task, are amply used in education and clinical neuropsychology contexts and they provide a measure of attention capabilities. Second, the concept of attention has evolved so as to be considered as a system formed by different attentional networks, each one performing different functions (Fuentes, in press; Posner and Raichle, 1994). Therefore, there seems to be a kind of dissociation between how attention is conceived in educational/clinical contexts, and how it is defined in neurocognitive studies. In the present study we asked whether some tests that are thought to measure attention in schoolchildren are related to executive attention.

A preliminary analysis showed that both boys and girls are susceptible to interference from colour words when performing the computerized Stroop task. These results confirm previous findings that children aged 7 years show Stroop interference effects (Bonino and Ciairano, 1997; González et al., 2001; Schiller, 1966). However, when we looked at the in-

 Electronic Journal of Research in Educational Psychology. No. 1 (2), 23-36. ISSN: 1696-2095

teractions between the standardised tests and Stroop interference, the results showed that only 3 tests produced significant differences in the Stroop effect.

When using a Stroop task that neuroimaging studies have clearly associated with the executive network, only Picture Completion (errors), Digit Backward (RTs) and TMT-B (errors) appeared to be sensitive to differential Stroop interference. Importantly, scores in both versions of the Stroop task did not correlate.

From the present study, therefore, we can draw two important implications. First, only two subtests of the WISC-R scales that supposedly measure attention (Picture Completion and Digits Backward), and part B of the Trial Making Test were related to executive attention in seven-year-old children when interference effects are measured by the computerized Stroop task. Poor performance on these tests might be due to dysfunction of the neural circuitry associated with the executive attentional network.

Second, given the above results, the Stroop card version may not be an adequate index of executive attention, at least at this age. This is further supported, although rather indirectly, by the poor correlation between the two versions of the Stroop task. Dissociations between these two versions have been previously reported when attention is assessed in schizophrenic patients (Boucart, Mobarek, Cuervo, and Danion, 1999). Boucart et al. (1999) pointed out that there are important methodological differences between the two versions. Whereas stimuli are presented one by one on the computer screen, the card version displays all stimuli simultaneously, having the subjects select the current target among a great number of distractors. Also, conditions are mixed in the computerized version but blocked in the card version. All these differences might suggest that different attention mechanisms are stressed in each version of the task. Whereas resolution of conflict, associated with the executive attentional network, might be predominant in the computerized version (Botvinick, Nystrom, Fissell, Carter, and Cohen, 1999), location of current target and filtering out distractor information, a role in which the pulvinar nucleus of the thalamus plays a role (LaBerge and Buchsbaum, 1990), might predominate in the card version (Boucart et al., 1999), involving the orienting attentional network.

Future research will determine whether the kinds of relationships found in the present study can be established between other standardised tests and different forms of attention, and as a function of development.

Authors' note

This research was supported by Grants BSO2003-00930 and PB98-0379 from the *Ministerio de Ciencia y Tecnología*. The authors wish to thank the children for participating in this study and the staff of Lope de Vega (Almería), Francisco de Goya (Almería) and San Buenaventura (Murcia) schools for their invaluable help.

Luis J. Fuentes

References

Amieva, H., Lafont, S., Auriacombe, S., Rainville, C., Orgogozo, J.M., Dartigues, J.F., & Fabrigoule, C. (1998). Analysis of error types in the Trial Making Test evidences an inhibitory deficit in dementia of the Alzheimer Type. *Journal of Clinical and Experimental Neuropsychology, 20*, 280-285.

Baddeley, A. (1986). *Working Memory*. Oxford, England: Oxford University Press.

Bonino, S., & Ciairano, S. (1997). Effetto Stroop e capacità di inhibizione: Una ricerca in campo evolutivo. [The Stroop Effect and inhibition capacity: Research in the developmental field.] *Giornale Italiano di Psicologia, 24,* 587-603.

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., & Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179-181.

Boucart, M., Mobarek, N., Cuervo, C., & Danion, J.M. (1999). What is the nature of increased interference in schizophrenia? *Acta Psychologica, 101*, 3-25.

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*, 215-222.

Fan, J., Flombaum, J.I., McCandliss, B.D., Thomas, K.M., & Posner, M.I. (2003). Cognitive and brain mechanisms of conflict. *NeuroImage, 18,* 42-57.

Fuentes, L.J. (in press). Inhibitory processing in the attentional networks. In M.I. Posner (Ed.), *Attention*. New York: Guilford Press.

Fuentes, L.J., Vivas, A.B. & Humphreys, G.W. (1999). Inhibitory mechanisms of attentional networks: Spatial and semantic inhibitory processing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1114-1126.

Golden, C.J. (1978). *Stroop color and word test*. Chicago: Stoelting.

González, C., Fuentes, L.J., Carranza, J.A., & Estévez, A.F. (2001). Temperament and attention in the self-regulation of seven-year-old children. *Personality and Individual Differences, 30,* 931-946.

LaBerge, D., & Buchsbaum, M.S. (1990). Positron emission tomographic measurements of pulvinar activity during an attention task. *Journal of Neuroscience, 10*, 613-619.

Lezak, M.D. (1995). *Neuropsychological assessment*. New York: Oxford University Press.

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 Science*, *87*, 256-259.

Posner, M.I., & DiGirolamo, G. J. (1998). Conflict, target detection and cognitive control. In R. Parasuraman (Ed.), *The attentive brain* (pp. 401-423). Cambridge, M. A.: MIT Press.

Posner, M.I. & Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Posner, M.I., & Raichle, M E. (1994). *Images of mind*. New York: Scientific American Library.

Schiller, P.H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology, 72*, 105-108.

Stroop, J R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology, 18*, 643-662.

Taylor, S.F., Kornblum, S., Lauber, E.J., Minoshima, S., & Koeppe, R.A. (1997). Isolation of specific interference processing in the Stroop task: PET activation studies. *NeuroImage, 6*, 81-92.

Weschler, D. (1974). *Manual for the Weschler Intelligence Scale for Children-Revised*. San Antonio, TX: The Psychological Corporation.