Sapienza Università Editrice

Psychology Hub (2020)
XXXVII, 2, 7-16

Relationship between Accuracy and Speed in the Raven's Coloured Progressive Matrices test: normative data for Italian children aged 5-6 years

Teresa Gloria Scalisi[a], Eleonora Cannoni[a], Lena Traversari[a]

[a] *Department of Developmental and Social Psychology, Sapienza University of Rome, Rome, Italy*

Article info

Submitted: 25 February 2020
Accepted: 24 June 2020
DOI: 10.13133/2724-2943/17165

Abstract

The Raven's Coloured Matrices (CPM) is one of the most used tests for assessing fluid intelligence in children. Although the speed factor is considered important for the measurement of this construct, normative data for speed are not available for the CPM test, and the speed contribution on the test performance is not known in the literature. To help fill these gaps, we provide the CPM Accuracy and Speed norms, and data on the relationship between the two measures, concerning a sample of 468 Italian children aged 5-6 years. A negative correlation emerged between accuracy and speed for the Ab and B Sets of the test, which include the most complex problems. The association in Set B (the most difficult) between a decrease in accuracy and an increase in speed suggests the prevalence of random responses, advising the exclusion of this Set from the computation of the total score in the considered ages. Comparisons between three groups of children (inaccurate and fast; inaccurate and slow; accurate and slow) indicate that the poor accuracy of fast children may be due to an impulsive approach to the task. This pattern of results calls into question the practice of evaluating the test performance taking into account only the accuracy factor.

Keywords: preschool children; Raven's Coloured Matrices; accuracy; speed; normative data.

*Corresponding author.
Teresa Gloria Scalisi
Department of Developmental
and Social Psychology,
Sapienza University of Rome
Via dei Marsi, 78, 00185, Rome, Italy
Phone: +39 06 4991 7919
E-mail: gloria.scalisi@uniroma1.it
(T. G. Scalisi)

Introduction

The Raven's Coloured Progressive Matrices Test or CPM (Raven et al., 1998) measures the main components of general intelligence (g factor), in particular fluid intelligence, which uses reasoning to solve logical problems that cannot be resolved based on previous knowledge (Cattell, 1987). It is considered a cultural-free test because it evaluates the ability to understand relationships between abstract figural elements and is widely used internationally to assess the intellectual level in developmental age (Raven, 2008).

The CPM contains three Sets of items (A, Ab, and B), ordered by ascending level of difficulty. The complete test totals 36 items; each item displays a matrix consisting either of a figure with a piece missing or four figures (arranged on two rows and two columns), with the fourth one missing. A factor analysis (Muniz et al., 2016) identified one general factor and three factors respective to each scale. The easiest items are prevalent in Set A. Their components, such as colour, shape, size, and orientation, form a gestalt, and must be analysed to give the correct answer. The items of intermediate difficulty – which are present in Sets Ab and B – require considering rows and columns of the matrix at the same time. The most difficult items are the last five of the Set B; rows and columns must be considered together even for these items, but the three figures of the matrix are different from each other, thus the task requires a higher level of abstraction than all other items. The Italian standardization of the booklet form (Belacchi et al., 2016) presents normative data for children aged 3-11 years for individual administration and 7-11 years for the collective one. However, a later study of children aged 3-6 years (Giofrè & Belacchi, 2015) showed that children sometimes answer randomly until the age of 4, and many of them are unable to perform the test adequately; therefore, the reliability of the results of children aged 3-4 years are doubtful.

Although both the original and the Italian manual provide instructions for recording the time taken to perform the test, neither specifies how to use and interpret this measure, because normative data are not provided. Recent research shows that accuracy and speed are both critical for assessing cognitive skills (e.g., Ren et al., 2018), especially fluid intelligence (Sheppard & Vernon, 2008). As early as 1927, Thorndike (Thorndike et al., 1927) claimed that, all things being equal, the less time a person takes to provide a correct answer to an intelligence test, the better his ability is in that test. Therefore, accuracy should be positively correlated with speed in intellectual tests. However, Partchev and De Boeck (2012) and Goldhammer et al. (2014) showed that the positive accuracy-speed correlation emerges in adult samples when the task assesses acquired knowledge, but not when the task requires the understanding of new problematic situations and involves complex cognitive processes. The authors found that higher accuracy is associated with a lower speed in this type of tasks, as the careful examination of the characteristics of the problem and the strategic control of cognitive resources increase the probability of a correct answer, but require more time, thus producing a speed-accuracy trade-off. In fact, a study conducted on the Advanced form of Raven's Matrices (Goldhammer et al., 2015) showed that the time taken by adults to provide the correct answer increases as a function of the item difficulty.

In problem-solving tasks such as Raven's Matrices, therefore, impulsive responses have negative effects on accuracy and generally depend on the tendency to guess the answer without having carefully examined all elements of the problem. Among the possible causes of guessing, Wise (2017) suggests low test taker motivation, fatigue, and drop in sustained attention that occurs towards the end of long tasks or when other tests have previously been faced. The author underlines that rapid-guessing answers are uninformative about the construct that is tried to be measured, therefore they reduce the validity of the test and should not be considered.

Another variable that can influence the relationship between accuracy and speed in problem-solving tasks is the reflective/impulsive (R/I) cognitive style. Reflective people delay the response until they are sure that the answer is correct, therefore being very accurate but slow; on the contrary, impulsive people choose the first answer that comes to their mind, resulting faster but less accurate (Quiroga et al., 2011). The R/I index is usually calculated as the difference between the z scores of errors and response latencies (e.g., Quiroga et al., 2011) or between correct answers and speed measures (Nietfeld & Bosma, 2003); in the first case, the highest indices point greater impulsivity, and the lower indices point higher reflectivity, while in the second case the opposite occurs. The importance of the relationship between accuracy and speed in cognitive tests has brought some authors to promote the use of an efficiency index (independent of the R/I index) based either on the sum of the accuracy and speed z scores (e.g., Fernández-Martín & Hinojo-Lucena, 2006) or on the average number of correct answers provided in a unit of time (Ren, Wang, Sun, Deng & Schweizer, 2018). Based on this index, children who obtain high scores in both accuracy and speed would be more efficient (and vice versa for inefficient children).

Phillips and Rabbit (1995) found significant positive relations between impulsivity and the total score obtained in four intelligence tests, including Raven's Matrices, after partialling out the overall accuracy of the four tests. They showed that the R/I cognitive style was a stable characteristic in adults and was independent of the general intellectual level. Lozano, Hernández, and Santacreu (2015) found that university students who were slow and inaccurate (inefficient) in visual discrimination tasks, obtained low scores in reasoning tests, while impulsive students did not differ from reflective students in these tasks. The authors conclude that poor accuracy and slowness in cognitive tasks are probably associated with less intellectual capacity, while poor accuracy and high speed are associated with an impulsive cognitive style.

The R/I cognitive style has also been studied in children and adolescents, using the Matching Familiar Figures Test-20 (MFFT-20). It comprises 20 items, each consisting of a model drawing (target) and six highly similar drawings, one of which is identical to the target. The task is to identify the exact copy of the target. The response latency for the first choice and the number of errors for each item are recorded. The total number of errors and the mean response latency are calculated as total scores (see Mazzocchi et al., 2010; Riaño-Hernández et al., 2015). Arán-Filippetti and Richaud de Minzi (2012) found negative and significant correlations between cognitive impulsivity, attentional control, and some Executive Functions

– such as working memory, planning, and cognitive flexibility – in children aged 8-12 years. Furthermore, as also observed in previous works (e.g., Zelniker & Jeffrey, 1976), reflective children tended to approach the MFFT-20 task through an analytical procedure, while impulsive children tended to approach the task through a global and holistic procedure. The authors therefore exclude that impulsive children perform poorly on cognitive tasks due to inadequate skills; instead, they hypothesize that impulsive children process stimuli in a global fashion even when the accurate analysis of the details is critical to providing the correct answer. For this reason, the authors conclude that cognitive impulsivity is not necessarily associated with cognitive deficits or neurological developmental disorders. Salkind and Nelson (1980) showed that cognitive impulsivity gradually decreases in typically developing children aged from 5 to 10 years, and can be corrected through school interventions (e.g., Gargallo, 1993). Hence, a greater cognitive impulsivity can be expected in 5-year-old children, and the decreasing trend observed with increasing age is likely due to the progressive maturation of attentional processes (see Fisher & Kloos, 2016).

To our knowledge, no study of children has examined the relationship between accuracy and speed in logical reasoning tasks such as CPM. Investigating this relationship could allow us to understand whether impulsivity/reflectivity is also evident in problem-solving tasks. However, some studies show that impulsive cognitive style, measured with an external criterion, can negatively influence children's performance on intelligence tests. The errors made in the MFFT-20 test were indeed negatively correlated with the total score on the WISC-R scale (Mollick & Messer, 1978), with an intelligence score based on a reasoning test (Buela-Casal et al., 2002) and with a measure of the G factor (Solís Cámara Reséndiz & Servera Barceló, 2003). Moreover, impulsive children, identified through the MFFT-20 test, performed lower on WISC-R (Brannigan & Ash, 1977), on CPM (Bush & Dweck, 1975), on the Standard form of Raven's Matrices (Lawry et al., 1983) and on different tests of visuospatial reasoning (McKinney, 1973; 1975) compared to reflective children. McKinney (1973) observed that children classified as reflective based on their MFFT-20 scores attempt to regard several alternative hypotheses in reasoning tests, using a strategy that assesses the relevance of conceptual categories. On the contrary, impulsive children often approach reasoning tests using a random trial-and-error procedure and are less likely to form abstract hypotheses.

Research goals and hypotheses

Our first goal is to verify if the time taken by preschool children to perform the CPM test allows obtaining more information about the individual differences in approaching such a task, in particular as regards the possibility that poor accuracy may be the result of an impulsive approach. Our second goal is to provide normative data also for speed. We have chosen to consider preschool children because the tendency to respond impulsively is more widespread among younger children (Salkind & Nelson, 1980); we excluded children below the age of 4.5 years to avoid unreliable results (see Giofrè & Belacchi, 2015). We expect to observe in our sample a negative correlation between accuracy and speed, according to the results obtained with adults (Goldhammer et al., 2015; Phillips & Rabbit, 1995); this correlation would indicate that the children who resolve the highest number of problems are also the slowest (and vice versa). We also expect the speed-accuracy trade-off to be larger for the Sets containing more difficult items, i.e. Ab and B.

According to Facon and Nuchadee (2010), if two groups of children with the same total CPM accuracy score show a different trend across the items as a function of the item difficulty, this suggests that the performance of the two groups is influenced by different factors. Hence, our third goal is to compare the performance on the three CPM Sets of slow and fast children obtaining the same low accuracy total score. If faster children are less disadvantaged in the Set A, whose problems favour a more global processing (Muniz et al., 2016), this could suggest that their poor accuracy is influenced by a tendency to respond impulsively (see Arán-Filippetti & Richaud de Minzi, 2012). Moreover, we expect that the performance of faster children will accelerate in the presence of the more difficult Set B since impulsive children often adopt a fast guessing strategy in solving difficult cognitive problems (McKinney, 1973), but we expect the slower group to be just as slow in Set B as in Set Ab. We also expect to find a difference in Set A between less-accurate-faster and more-accurate-slower children. In fact, if the global procedure is more suitable for the children of the first group, they should be more efficient for this Set than the children of the second group, and thus be able to solve a higher number of problems in one unit of time.

Method

Participants

Participants were 468 typically developing preschool children, 221 were 5-year-olds (mean age 5 years and 1 month; SD = 3.1 months; MIN = 4 years and 6 months, MAX = 5 years and 5 months; 105 females) and 247 were 6-year-olds (mean age 6 years; SD = 3 months; MIN = 5 years and 6 months, MAX = 6 years and 6 months; 122 females). Children lived in Central Italy (i.e. Roma, Terni, Latina, Napoli, and their respective provinces) and were examined in the years 2017-2019. Informed consent was obtained from parents before testing, and children provided verbal assent at the time of testing. The study was approved by the Ethics Committee of the authors' Department.

Materials and procedure

Participants were individually administered the Raven Coloured Progressive Matrix (CPM) test (Raven et al., 1998; Italian standardization of Belacchi et al., 2016) in the booklet version. The children leafed through the booklet independently. The first three items of Set A were used for training: if an error occurred, the problem was presented again until obtaining the correct answer. Responses to these three items were always recorded as correct. No assistance was provided for problem-solving or feedback on performance from item 4 of Set A to the end of the task, except repeating instructions in case of need. For Set A only, the recording of the execution time (carried out by the examiner by means of a stopwatch) was started from item 4 instead of item 1.

Accuracy, speed and efficiency scores computation

According to the test manual (Belacchi et al., 2016), the number of correct responses (CR) was calculated separately for each Set (A, Ab and B). The total accuracy score was then obtained by summing the CR for the three Sets. In order to obtain a "positive" measure of speed, the mean number of responses (correct + incorrect) per minute (RM) was calculated for each Set, including only the items 4 to 12 for the Set A. The total speed score was calculated by averaging the speed scores across the three Sets. A joint accuracy and speed score (*efficiency index*) was also calculated based on the mean number of correct responses per minute (CM), both overall and for each Set.

Results and Discussion

Descriptive statistics and comparisons between the two age groups

Table 1 shows the descriptive statistics of accuracy (CR), speed (RM) and efficiency (CM) scores, for both the three Sets and the overall task. The skewness and kurtosis values of the variables suggest that the assumption of normality was not violated, thus making parametric statistics applicable.

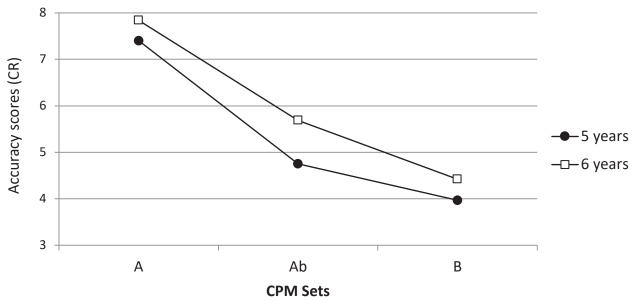
Tab. 1. Descriptive statistics of Accuracy (CR), Speed (RM) and Efficiency (CM) scores both total and separate for the three CPM Sets.

CPM Scores	Means	Min	Max	SD	Skewness	Kurtosis
A CR	7.63	1.00	12.00	1.47	-0.211	0.781
Ab CR	5.25	1.00	12.00	2.04	0.555	0.547
B CR	4.21	1.00	12.00	1.70	0.697	1.196
A RM	10.61	3.18	20.59	3.01	0.513	0.402
Ab RM	8.30	3.40	18.20	2.35	0.788	1.213
B RM	8.68	2.63	19.86	2.74	0.871	1.331
A CM	6.72	1.03	15.56	2.24	0.626	0.780
Ab CM	3.54	0.00	9.66	1.53	0.937	1.477
B CM	2.93	0.00	7.56	1.23	0.489	0.239
TOT CR	17.09	6.00	33.00	4.12	0.491	0.611
TOT RM	9.20	3.42	18.42	2.41	0.691	0.969
TOT CM	4.40	1.29	9.18	1.33	0.558	0.597

Note. CR = Correct Responses; RM = Responses (correct + incorrect) per Minute; CM = Correct responses per Minute.

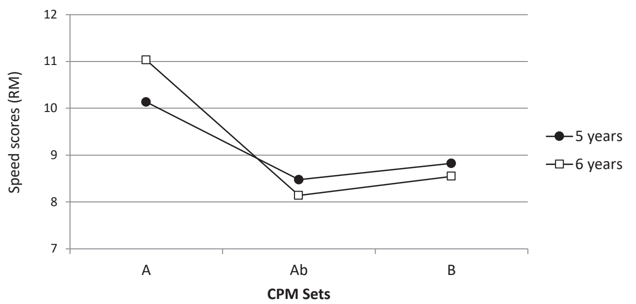
A factorial ANOVA, with two between-subject factors (Age and Gender) and one within-subject factor (Sets A, Ab and B), was performed separately for accuracy and speed in order to compare the two age groups of 5- and 6-year-old children. The gender variable will not be discussed as it did not produce any significant results. The interaction between Age and Sets (when significant) was examined in more detail through planned comparisons aimed at verifying whether the differences between the two age groups were significant for all three Sets. The ANOVA on the accuracy scores (CR) revealed a significant main effect for Age ($F(1, 464) = 23.9$; partial $\eta^2 = 0.05$; $p<0.001$; mean CR scores: 5.4 at age five and 6.0 at age six). The main effect for Sets was also significant ($F(2, 928) = 810.3$; partial $\eta^2 = 0.64$; $p<0.001$; mean CR scores: A=7.6; Ab=5.2; B=4.2), with both A-Ab and Ab-B significant differences ($F(1, 464) = 714.7$ and 140.3 respectively; $p<0.001$). Even the interaction between Age and Sets, displayed in Figure 1, was significant ($F(2, 928) = 5.3$; partial $\eta^2 = 0.01$; $p<0.01$). The differences between the age groups were significant for all three Sets A, Ab and B; $F(1, 464)$ values were 10.5 ($p=0.001$), 25.6 ($p<0.001$) and 8.36 ($p<0.01$) respectively. Although the accuracy of both groups decreased as the task difficulty increased across the three Sets, the interaction in Figure 1 shows a greater decrease in performance from A to Ab by 5-year-olds compared to 6-year-olds, who exhibited a more linear decreasing trend.

Fig. 1. Accuracy scores (Correct Responses - CR) of children aged 5 and 6 years for the three CPM Sets



The ANOVA on speed scores (RM) revealed a significant main effect for Sets ($F(2, 928) = 309.3$; partial $\eta^2 = 0.40$; $p<0.001$; mean RM scores: A=10.6; Ab=8.3; B=8.7). The interaction between Age and Sets, shown in Figure 2, was also significant ($F(2, 928) = 25.2$; partial $\eta^2 = 0.05$; $p<0.001$) and indicated that the difference between the two age groups was significant only for Set A ($F(1, 464) = 10.5$; $p=0.001$), with 6-year-olds faster than 5-year-olds. However, Figure 2 shows that the performance of both groups slowed markedly in the transition from Set A to Ab ($F(1, 464) = 494.3$; $p<0.001$) and accelerated in a more moderate but significant way from Set Ab to B ($F(1, 464) = 24.2$; $p<0.001$).

Fig. 2. Speed scores (Responses per Minute - RM) of children aged 5 and 6 years for the three CPM Sets



The comparison of Figures 1 and 2 shows that, regardless of age, children significantly slowed down their performance in the impact with Set Ab (more challenging than A) probably to pay more attention to the characteristics of the stimuli, despite committing a higher number of errors. The errors still increased significantly in the transition from Ab to B, but a significant speed increase accompanied the accuracy decrease. Results concerning Set B suggest that children mainly responded too quickly and at random, probably because of fatigue and or loss of attention. It is also possible that the difficulty level of the problems in Set B exceeds the ability of many preschool children, and thus they disengage from the test-taking process. Therefore, in line with Wise's (2017) opinion on guessing, the responses to Set B should not be counted in the total score for children aged 5-6 years, and a reduced version of the test, including only Sets A and Ab, could be used for this age-band, according to Giofrè and Belacchi (2015) suggestions.

The relationship between accuracy and speed: is there a trade-off?

The relationship between accuracy and speed was first examined globally, considering the overall sample and the total test scores, and then separately for the two age groups and the three Sets. Since 5-year-olds were less accurate than 6-year-olds, the total accuracy and speed (CR and RM) scores were converted to z scores separately for the two age groups in order to control for the effects of age. The correlation between CR and RM z scores in the total sample was -0.27 ($p < 0.001$). Consistent with our hypothesis, a significant trade-off emerged, showing that accuracy decreased as speed increased (and vice versa). The scatterplot of Figure 3 displays the relationship between the two variables. Considering the extreme scores of the CR and RM distributions, it is apparent that no very fast-accurate or very slow-inaccurate children emerged, but fast-inaccurate and slow-accurate children could be identified.

Fig. 3. Scatterplot of the z scores concerning the total number of Correct Responses (CR) and Responses per Minute (RM) of the overall sample (N = 468; r = -0.27; p <0.001).

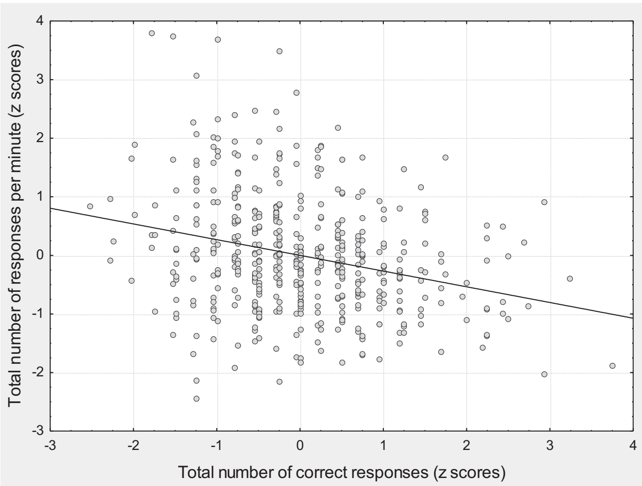


Table 2 shows the CR-RM correlations calculated on raw data separately for the two age groups and the three Sets. The Set A correlation was not significant for both groups (r = -0.13 for 5-year-olds and -0.07 for 6-year-olds), but a significant trade-off was observed for Sets Ab and B. The difference between Ab and B correlation coefficients (Lee & Preacher, 2013) was not significant for 5-year-old children (r = -0.28 and -0.27 for Ab and B Sets respectively; p>0.05) while it was for 6-year-olds (-0.18 and -0.30 for Ab and B respectively; p <0.05). The two groups thus exhibited a different trend in the trade-off, with a significant increase only between Sets A and Ab for the younger group and a quite linear increase across Sets for the older one. Interestingly, Figure 1 shows that the accuracy of younger children mostly decreased when moving from Set A to Set Ab, while older children accuracy decreased almost linearly across Sets. In both groups, therefore, the slowest children were the ones who provided the highest number of correct answers as the problems became most challenging. This result is in agreement with Goldhammer et al.'s (2015) findings concerning the performance of adults on problem-solving tasks.

Tab. 2. Intercorrelations among total CPM scores and Set scores. The values below the diagonal refer to 5-year-olds (N = 221), and those above the diagonal refer to 6-year-olds (N = 247). Values >0.16 are significant at the 0.01 level.

	A CR	Ab CR	B CR	A RM	Ab RM	B RM	TOT CR	TOT RM
A CR		**0.42**	**0.33**	-0.07	-0.16	**-0.19**	**0.69**	-0.15
Ab RC	**0.40**		**0.46**	-0.07	**-0.18**	**-0.22**	**0.85**	**-0.17**
B CR	**0.31**	**0.54**		-0.12	**-0.22**	**-0.30**	**0.77**	**-0.23**
A RM	-0.13	-0.12	-0.09		**0.74**	**0.66**	-0.11	**0.90**
Ab RM	**-0.25**	**-0.28**	**-0.25**	**0.62**		**0.81**	**-0.24**	**0.93**
B RM	**-0.27**	**-0.28**	**-0.27**	**0.64**	**0.79**		**-0.30**	**0.90**
TOT CR	**0.71**	**0.85**	**0.79**	-0.14	**-0.34**	**-0.35**		**-0.23**
TOT RM	**-0.24**	**-0.25**	**-0.23**	**0.86**	**0.89**	**0.91**	**-0.31**	

Note. CR = Correct Responses; RM = Responses (correct + incorrect) per Minute

Selection of children with different relationships between accuracy and speed

An accuracy-speed *discrepancy* criterion was used, based on the difference between the total CR and RM z scores. Six children were excluded from the selection as their CR z scores were <-2. The following two groups were identified, which coincidentally were equal in size (N = 42): *inaccurate-fast* (ACC-SPEED+), with zCR<0, zRM>0 and a difference zCR-zRM<-2; *accurate-slow* (ACC+SPEED-) with zCR>0, zRM<0 and zCR-zRM>2. We chose a discrepancy criterion based on a difference greater than |2| z scores because z = ± 1.96 is the critical value for a = 0.05. We rounded the cut-off to |2| to obtain an integer and more restricted criterion. A third group of 42 children was then selected (ACC-SPEED-) to obtain a group of children paired for accuracy with ACC-SPEED+ children, but showing longer execution times. To avoid including in this group children scoring close to the other two groups, the discrepancy criterion was a difference zCR-zRM ranging from -1.5 to 1.5.

Table 3 shows the characteristics of the three groups and the z scores concerning the selection criteria. The three groups were not significantly different for mean age ($F(2, 123) = 0.8$; $p = 0.45$), gender composition (Chi-square = 1.8; df = 2; $p = 0.41$) and age group composition (Chi-square = 1.03; df = 2; $p = 0.60$). Consistent with the selection criteria, the two groups ACC-SPEED+ and ACC-SPEED- did not differ for accuracy ($F(1, 82) = 0.002$; $p = 0.96$) but only for speed ($F(1, 82) = 272.3$; $p < 0.001$). It should be noted that the ACC-SPEED- group was defined as slow as the ACC+SPEED- one, but the first group was faster than the second, due to the negative accuracy-speed correlation found in the overall sample (see Table 3). The abbreviation ACC-SPEED- therefore indicates a "group of children equally inaccurate but slower than the ACC-SPEED+ ones". Appendix A shows the descriptive statistics of the three groups separately for the three Sets.

Tab. 3. Composition by gender and age of the three groups that differ in accuracy and speed; means and standard deviations (in brackets) of the scores used for selecting the children of the three groups

	GROUPS		
	ACC-SPEED+	ACC-SPEED-	ACC+SPEED-
N	42	42	42
N of females	22	18	16
N of 5-years-old	16	20	20
Age in months	68.12 (4.8)	67.33 (5.1)	66.75 (5.0)
CR TOT (z scores)	-1.08 (0.5)	-1.08 (0.4)	1.67 (0.8)
RM TOT (z scores)	1.81 (0.9)	-0.83 (0.6)	-1.15 (0.4)
Differences zCR-zRM	-2.89 (0.8)	-0.25 (0.7)	2.81 (0.8)

Note. CR = Correct Responses; RM = Responses (correct + incorrect) per Minute

Comparison between inaccurate children who differ in speed

One aim of the study was to find out whether different factors could influence the CPM performance of fast- and slow-inaccurate children. According to this hypothesis, they would show a different trend across the Sets as a function of the item difficulty (see Facon & Nuchadee, 2010). In particular, we expected the faster group to provide a higher number of correct answers on Set A compared to the slower group, and to show a different trend in speed when moving across the three Sets (see the Research Goal and Hypotheses section).

Two separate ANOVAs were performed on the raw accuracy and speed (CR and RM) scores, considering only the interaction between Group (ACC-SPEED+ and ACC-SPEED-) and Sets, followed by planned comparisons conducted to test our hypotheses. Figures 4 and 5 display the mean CR and RM scores of the two groups and the mean scores of the overall sample. In the ANOVA on CR scores the Group x Sets interaction only approached significance ($F(2, 164) = 2.7$; partial $\eta^2 = 0.03$; $p=0.07$). However, consistently with our hypothesis, the ACC-SPEED+ group was significantly more accurate on Set A compared to the other group ($F(1, 82) = 4.8$; $p < 0.05$). The Group x Sets interaction was significant in the ANOVA on RM scores ($F(2, 164) = 3.2$; partial $\eta^2 = 0.04$; $p < 0.05$). The differences between the A-Ab and Ab-B Sets (not reported for brevity) were all significant ($p < 0.001$) for both groups, except for the Ab-B difference in the ACC-SPEED- group ($F(1, 82) = 0.2$; $p = 0.64$).

Fig. 4. Accuracy scores (Correct Responses - CR) of inaccurate-fast children (ACC-SPEED+), inaccurate-slow children (ACC-SPEED-), and the overall sample

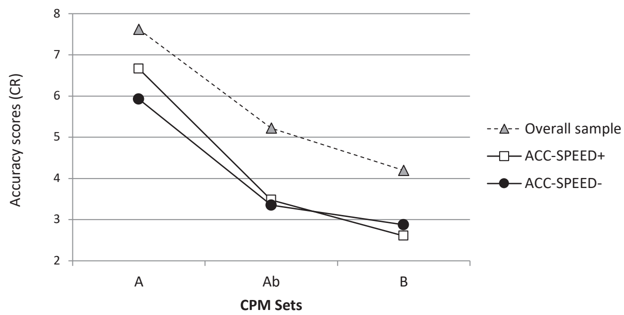
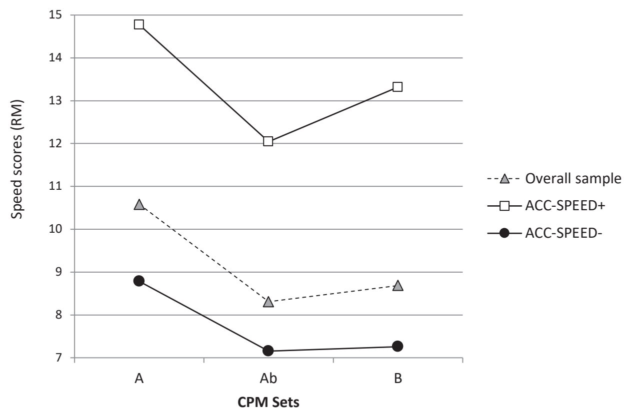


Fig. 5. Speed scores (Responses per Minute - RM) of inaccurate-fast children (ACC-SPEED+), inaccurate- slow children (ACC-SPEED-), and the overall sample



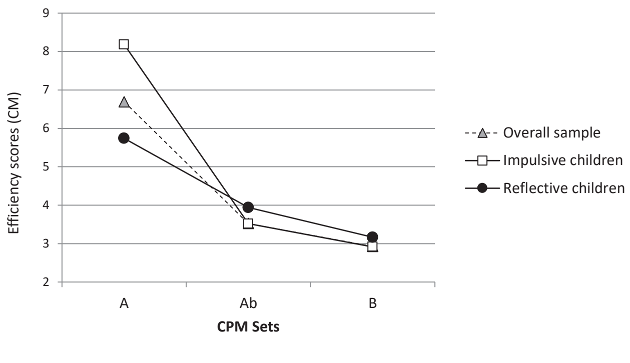
Overall accuracy being equal between the two groups, the different trend shown across the Sets indicates that different factors affected their poor performances. It is apparent from Figures 4 and 5 that the gap between ACC-SPEED- children and the overall sample was nearly the same for all three Sets, meaning that the impairment in problem-solving of ACC-SPEED- children was similar for both simple and complex items. On the contrary, ACC-SPEED+ children displayed a smaller accuracy gap on Set A and increased their speed more than the overall sample while solving the Set B problems. Some elements contribute to support the hypothesis that a tendency to respond impulsively was responsible (at least in part) for their poor accuracy. Data on Appendix A show that a mean difference of -2 z scores between accuracy and speed was already present in this group from Set A, thus suggesting that faster answers were not due to inability to solve the most challenging problems. Furthermore, the association between higher accuracy on Set A and higher acceleration on Set B (in comparison to the other group) suggests the prevalence in the ACC-SPEED+ group of a global and holistic way of processing (more appropriate to the properties of Set A items) combined with lower attention on the details of more complex stimuli, which were mainly answered quickly and at random. Research on cognitive style showed that all these characteristics are typical of impulsive children (e.g. McKinney, 1973; Arán-Filippetti & Richaud de Minzi, 2012).

Are impulsive children less efficient than reflective ones?

In this analysis, we compared the group hypothesized as *impulsive* (ACC-VEL+) with the group of accurate-slow children (ACC+SPEED-) whose approach to solving the CPM problems was indicative of a higher degree of reflectivity. We wanted to check whether impulsive children were not only less accurate but also less efficient than reflective children. A Group x Sets ANOVA was conducted on the efficiency score calculated as the mean number of correct answers provided per minute (CM). To contain type I error increase, resulting from including ACC-SPEED+ children again in this analysis, the α level was set at 0.01.

The ANOVA revealed significant main effects for Group (F (1, 82) = 8.4; partial η^2 = 0.09; p <0.01) and Sets (F (2, 164) = 233.2; partial η^2 = 0.74; p <0.001). The interaction between the two variables was also significant (F (2, 164) = 39.7; partial h^2 = 0.33; p <0.001) and is represented in Figure 6, in which CM scores of the total sample are also reported. As shown in Figure 6, the performance was more efficient on Set A compared to Sets Ab and B for all the groups examined. Consistent with our hypothesis, the differences between impulsive and reflective children only concerned Set A (F (1, 82) = 38.7; p<0.001), with impulsives more efficient than reflectives. Examining the mean CR scores and the mean execution times of the two groups for Set A, it appears that impulsive children provided seven correct answers in 50 seconds, while reflective children took almost twice as long (96 seconds) to provide nine correct answers. As a consequence, reflective children could be more disadvantaged than impulsives in tests administered under time pressure, which are now used even in school settings.

Fig. 6. Efficiency scores (Correct responses per Minute - CM) of inaccurate-fast children (ACC-SPEED+), accurate- slow children (ACC+SPEED-), and the overall sample



Conclusions

Our work does not concern important critical debates related to the intelligence construct (Cornoldi, 2009) such as the validity of the g factor concept (see Kovacs & Conway, 2016), the specific diagnostic categories based on the scores obtained for the intelligence tests (see Vianello & Cornoldi, 2018) or the real independence of the CPM test from cultural factors (see Fox & Mitchum, 2013). It only questions the classical notion that ability in reasoning tasks such as CPM can be assessed taking into account only accuracy scores, without considering speed and efficiency factors.

Our hypotheses on the overall sample have been confirmed, revealing that even in preschool children aged 5-6 years there is a negative correlation between accuracy and speed in problem-solving tasks, in line with results concerning adult participants (Goldhammer et al., 2015; Phillips & Rabbit, 1995). This correlation, as expected, was not observed in the easier Set A, while it was significant in the more difficult Sets Ab and B. Furthermore, the performance on Set B was characterized by the association between the greatest number of errors and the highest speed, thus suggesting that the responses to the most challenging problems were often based on a guessing strategy. For this reason, the inclusion of Set B in the CPM total score of preschool children is not recommended for diagnostic purposes, as also suggested by Giofrè and Belacchi (2015) based on the analysis of the correct answers only.

Studies of adults (Phillips & Rabbit, 1995; Lozano et al., 2015) show that cognitive impulsivity is independent of reasoning skills, whereas slow-inaccurate performance on cognitive tasks is an indicator of low abstract reasoning ability. However, many studies have shown that impulsive children are disadvantaged in intelligent tests (e.g. Bush & Dweck, 1975; Brannigan & Ash, 1977; Lawry et al., 1983). This discrepancy can be explained by considering that impulsivity is more common among children than adults (Salkind & Nelson, 1980), thus the tendency to respond very quickly can impair the performance on cognitive tasks that require careful analysis of details.

According to this hypothesis, we found that inaccurate-fast children showed on Set A better problem-solving abilities than the equally inaccurate but slower children, and greater performance efficiency than the slower and more accurate (reflective) children. Moreover, they accelerated on Set B more than the other group and the overall sample. Thus they were less impaired on items requiring more global processing procedures and mainly adopted a fast guessing strategy in solving more difficult cognitive problems. Both these characteristics are considered typical of impulsive children (e.g. McKinney, 1973; Arán-Filippetti & Richaud de Minzi, 2012), hence our results cast doubts on the interpretation of the poor overall accuracy of inaccurate-fast children in terms of low abstract reasoning ability, suggesting that it was influenced (at least in part) by the tendency to respond more impulsively, giving poor attention to details.

Inaccurate-slow children showed a different trend compared to inaccurate-fast ones, as the gap between this group and the overall sample in the level of accuracy and speed was almost the same across the three CPM Sets (see Figures 4 and 5). Thus their impairment in accuracy and speed was as severe for the easiest items as it was for the most difficult ones. This result, in line with Lozano et al.'s (2015) finding on adults, supports an explanation in terms of poor ability to solve abstract reasoning problems, despite the long times taken to find the correct answers.

The investigation on the relationship between accuracy and speed in CPM also showed that very accurate but slow preschool children could be disadvantaged in tests administered with a time limit. This needs to be taken into account when screenings of basic cognitive skills or subsequent learning assessments are carried out.

Our results also suggest issues for future research. Many studies examined the reflexivity/impulsivity dimension using the MFFT-20 test and found significant correlations with accuracy in intelligence and problem-solving tasks (e.g. Bush & Dweck, 1975; Brannigan & Ash, 1977; Lawry et al., 1983). It would be informative to check whether children classified as impulsive or reflective by the MFFT-20 test maintain these characteristics even in abstract reasoning tests such as CPM, by calculating the R/I index in both tests and checking if a significant correlation emerges.

The main weakness of this study is that we did not assess the relationship between cognitive style in CPM and tasks other than problem-solving, such as MFFT-20. Moreover, no tests were administered to evaluate the attentional processes of faster children. Despite these limitations, our results offer some insights for further deepening into whether CPM accuracy-speed scores are related to performance on other cognitive tasks. Our findings also suggest that it is important to evaluate speed, in addition to accuracy, in the diagnostic use of CPM in preschool age. Assessing some other skills (e.g. attention and executive functions) of inaccurate-fast children may help understand whether their poor accuracy is really due to low reasoning ability or is affected by other factors. The evidence of an impulsive approach to CPM could also be a warning for other problems, since cognitive impulsivity has negative effects on learning (e.g. Barret, 1977), but can be corrected with targeted interventions, improving also academic performance (Gargallo, 1993).

We display in Appendix B the normative data of Accuracy (CR), Speed (RM) and Efficiency (CM) for the two age groups, separately for the three Sets of the test; however, as already mentioned, taking into account the Set B score for diagnostic purposes is not advisable. The measures used in this research approximated the normal distribution, therefore the transformation in z scores is allowed. Unlike the percentile ranks, which cannot be added or subtracted, the difference between the z scores of accuracy and speed can be calculated to compute the R/I index, in particular for Set A, which was crucial in discriminating between groups. The CPM test is already widely used in both diagnostics and research, thus it could be helpful to obtain some other information on children just by recording even the time taken to perform each Set. Obviously, the individual R/I index obtained through CPM could be referred only to this test and the child classification as impulsive or reflective cannot be generalized to other tasks or behaviours.

Lastly, we recommend not transforming raw CPM scores of preschool children into IQ scores: the tendency to impulsivity, which has negative effects on accuracy, progressively decreases with age (Salkind & Nelson, 1980), thus in some cases the IQ score might be significantly lower than it would be some years later.

Author Contributions

The authors contributed equally to this manuscript.

Compliance with Ethical Standards

Conflict of interest

The authors declare that they have no competing interests.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent

Each participant dealt with the process of informed consent.

References

Arán-Filippetti, V., & Richaud de Minzi, M. C. (2012). Análisis de la relación entre reflexividad-impulsividad y funciones ejecutivas en niños escolarizados mediante un modelo de ecuaciones estructurales. *International Journal of Psychology and Psychological Therapy*, *3*, 427–440.

Barret, D. E. (1977). Reflection-Impulsivity as a Predictor of Children's Academic Achievement. *Child Development*, *48*, 1443–1447.

Belacchi, C., Scalisi, T. G., Cannoni, E., & Cornoldi, C. (2016). *Manuale CPM. Coloured Progressive Matrices. Standardizzazione italiana*. Firenze: Giunti O.S. Organizzazioni Speciali.

Brannigan, G. G., & Ash, T. (1977). Cognitive tempo and WISC-R performance. *Journal of Clinical Psychology*, *33*(1S), 212–212. doi: 10.1002/1097-4679(197701)33:1+<212::AID-JCLP2270330147>3.0.CO;2-S

Buela-Casal, G., De Los Santos-Roig, M., & Carretero-Dios, H. (2002). Reflexividad-Impulsividad e Inteligencia: Un Análisis de la Validez Teórica del Estilo Cognitivo y de la Validez Divergente del MFFT 20. *Revista Mexicana de Psicologia*, *19*(2), 155–166.

Bush, E. S., & Dweck, C. S. (1975). Reflections on conceptual tempo: Relationship between cognitive style and performance as a function of task characteristics. *Developmental Psychology*, *11*(5), 567–574. doi: 10.1037//0012-1649.11.5.567

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: North-Holland.

Cornoldi, C. (2009). *L'intelligenza.* Bologna: Il Mulino.

Facon, B., & Nuchadee, M. L. (2010). An item analysis of Raven's Colored Progressive Matrices among participants with Down syndrome. *Research in Developmental Disabilities*, *31*(1), 243–249. doi: 10.1016/j.ridd.2009.09.011

Fernández-Martín, F., & Hinojo-Lucena, F. (2006). El estilo cognitivo Reflexividad-Impulsividad (R-I) en el segundo ciclo de Educación Primaria: diferencias entre los sistemas de clasificación e implicaciones educativas. *Enseñanza & Teaching: Revista Interuniversitaria de Didáctica*, *24*, 117–130.

Fisher, A., & Kloos, H. (2016). *Development of selective sustained attention: the role of executive functions.* In J. A. Griffin, P. McCardle, & L. S. Freund (Eds) *Executive Function in Preschool-Age Children: Integrating Measurement, Neurodevelopment, and Translational Research*, American Psychological Association.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge-based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General*, *142*(3), 979–1000. doi: 10.1037/a0030155

Gargallo, B. (1993). Basic variables in reflection-impulsivity: A training programme to increase reflectivity. *European Journal of Psychology of Education*, *8*(2), 151–167. doi: 10.1007/BF03173159

Giofrè, D., & Belacchi, C. (2015). Versione ridotta delle matrici colorate di Raven: uno strumento rapido per l'assessment in bambini dai tre ai cinque anni e undici mesi. *Psicologia Clinica Dello Sviluppo*, *19*(1), 145–154. doi: 10.1449/79743

Goldhammer, F., Naumann, J., & Greiff, S. (2015). More is not Always Better: The Relation between Item Response and Item Response Time in Raven's Matrices. *Journal of Intelligence*, *3*(1), 21–40. doi: 10.3390/jintelligence3010021

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, *106*(3), 608–626. doi: 10.1037/a0034716

Kovacs, K., & Conway, A. R. A. (2016). Process Overlap Theory: A Unified Account of the General Factor of Intelligence. *Psychological Inquiry*, *27*(3), 151–177. doi: 10.1080/1047840X.2016.1153946

Lawry, J. A., Welsh, M. C., & Jeffrey, W. E. (1983). Cognitive Tempo and Complex Problem Solving. *Child Development*, *54*(4), 912–920.

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with no variable in common [Computer software]. Available from http://quantpsy.org.

Lozano, J. H., Hernández, J. M., & Santacreu, J. (2015). Impulsivity and speed-accuracy strategies in discrimination learning. *Journal of Cognitive Psychology*, *27*(1), 69–79. doi: 10.1080/20445911.2014.977793

Marzocchi, G. M., Re, A. M., & Cornoldi, C. (2010). *BIA. Batteria italiana per l'ADHD per la valutazione dei bambini con deficit di attenzione-iperattività*. Trento: Erickson.

McKinney, J. D. (1973). Problem-solving strategies in impulsive and reflective second graders. *Developmental Psychology*, *8*(1), 145. doi: 10.1037/h0033783

McKinney, J. D. (1975). Problem-solving strategies in reflective and impulsive children. *Journal of Educational Psychology, 67*, 807–820. doi:10.1037/0022-0663.67.6.807

Mollick, L. R., & Messer, S. B. (1978). The relation of reflection-impulsivity to intelligence tests. *The Journal of Genetic Psychology*, *132*(1), 157–158.

Muniz, M., Gomes, C. M. A., & Pasian, S. R. (2016). Factor structure of Raven's Coloured Progressive Matrices. *Psico-USF*, *21*(2), 259–272. doi: 10.1590/1413-82712016210204

Nietfeld, J., & Bosma, A. (2003). Examining the self-regulation of impulsive and reflective response styles on academic tasks. *Journal of Research in Personality*, *37*(3), 118–140. doi: 10.1016/S0092-6566(02)00564-0

Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated?. *Intelligence*, *40*(1), 23–32. doi: 10.1016/j.intell.2011.11.002

Phillips, L. H., & Rabbit, P. M. A. (1995). Impulsivity and speed-accuracy strategies in intelligence test performance. *Intelligence*, *21*(1), 13–29. doi: 10.1016/0160-2896(95)90036-5

Quiroga, M. Á., Martínez-Molina, A., Lozano, J. H., & Santacreu, J. (2011). Reflection-Impulsivity Assessed Through Performance Differences in a Computerized Spatial Task. *Journal of Individual Differences*, *32*(2), 85–93. doi: 10.1027/1614-0001/a000038

Raven, J. (2008). *Uses and Abuses of Intelligence. Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*, Royal Fireworks Press.

Raven, J., Court, J. H., & Raven, J. C. (1998). *Raven Manual, Section 1 (General overview) and Section 2 (Coloured Progressive Matrices)*. Oxford, UK: Oxford Psychologist Press.

Ren, X., Wang, T., Sun, S., Deng, M., & Schweizer, K. (2018). Speeded testing in the assessment of intelligence gives rise to a speed factor. *Intelligence*, *66*, 64–71. doi: 10.1016/j.intell.2017.11.004

Riaño-Hernández, D., Riquelme, A. G., & Buela-Casal, G. (2015). Conceptualization and assessment of impulsivity in adolescents: A systematic review. *Universitas Psychologica*, *14*(3), 1077–1090. doi: 10.11144/Javeriana.upsy14-3.ceia

Salkind, N. J., & Nelson, F. (1980). A Note on the developmental nature of reflection-impulsivity. *Developmental Psychology*, *16*(3), 237–238. doi: 10.1525/aa.1889.2.4.02a00080

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*(3), 535–551. doi: 10.1016/j.paid.2007.09.015

Solís Cámara Reséndiz, P., & Servera Barceló, M. (2003). Los efectos del modelo probabilístico sobre el estilo cognitivo reflexividad-impulsividad. *Psicothema*, *15*(4), 545–549.

Thorndike, E. L., Bregmamn, E. O., Cobb, M. V., & Woodyard, E. (1927). *The measurement of intelligence*. New York: Teachers College, Columbia University.

Vianello, R., & Cornoldi, C. (2018). Commenti a «Riflessioni sul funzionamento intellettivo limite o borderline: proposta per una discussione». *Psicologia Clinica Dello Sviluppo*, *22*(1), 143–168. doi: 10.1449/87537

Wise, S. L. (2017). Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications. *Educational Measurement: Issues and Practice*, *36*(4), 52–61. doi: 10.1111/emip.12165

Zelniker T., & Jeffrey, W. E. (1976). Reflective and impulsive children: Strategies of information processing underlying differences in problem solving. *Monographs of the Society for Research in Child Development*, *41*, 1-59. doi: 10.2307/1166020

Appendix A

Descriptive statistics of Accuracy (CR), Speed (RM) and Reflectivity/Impulsivity (R/I) scores of the subgroups selected, calculated separately for the three CPM Sets.

		Subgroups					
		ACC-SPEED+		ACC-SPEED-		ACC+SPEED-	
		Mean	DS	Mean	DS	Mean	DS
Accuracy scores (CR)	Set A	6.76	1.27	6.21	1.02	9.02	1.32
	Set Ab	3.52	1.09	3.67	1.18	8.10	1.76
	Set B	2.64	1.01	2.86	1.14	6.67	1.76
Speed scores (RM)	Set A	14.94	2.92	8.30	2.20	7.59	1.54
	Set Ab	12.22	2.21	6.58	1.42	5.83	1.15
	Set B	13.54	2.70	6.71	1.43	5.86	1.40
Efficiency scores (CM)	Set A	8.48	2.48	4.27	1.27	5.72	1.45
	Set Ab	3.59	1.32	2.03	0.88	3.97	1.29
	Set B	3.01	1.36	1.59	0.71	3.21	0.98
R/I scores (zCR-zRM)	Set A	-2.06	1.15	-0.22	1.12	2.01	1.02
	Set Ab	-2.59	1.03	-0.06	0.79	2.48	0.87
	Set B	-2.74	1.07	-0.08	0.86	2.49	1.26

Note. ACC-SPEED+ = inaccurate-fast children; ACC-SPEED- = children equally inaccurate than the first group but slower; ACC+SPEED- = accurate-slow children; CR = Correct Responses; RM = mean number of Responses per Minute; CM = mean number of Correct responses per Minute; R/I = Reflectivity/Impulsivity scores

Appendix B

The normative data of the children of the two age groups are shown below, separately for the three Sets of the CPM test.

		Age-band (years; months) from 4;6 to 5;5 (N = 221)				
		Mean	SD	Median	Percentile 25	Percentile 75
Accuracy scores (CR)	Set A	7.41	1.59	7.00	6.00	8.00
	Set Ab	4.76	1.90	5.00	3.00	6.00
	Set B	3.97	1.65	4.00	3.00	5.00
Speed scores (RM)	Set A	10.14	2.95	10.05	8.00	11.57
	Set Ab	8.48	2.41	8.23	6.68	9.67
	Set B	8.83	2.91	8.31	6.72	10.63
Efficiency scores (CM)	Set A	6.21	2.11	6.01	4.77	7.41
	Set Ab	3.26	1.39	3.15	2.30	3.92
	Set B	2.81	1.25	2.73	1.94	3.44

		Age-band (years; months) from 5;6 to 6;6 (N = 247)				
		Mean	SD	Median	Percentile 25	Percentile 75
Accuracy scores (CR)	Set A	7.84	1.32	8.00	7.00	9.00
	Set Ab	5.69	2.08	5.00	4.00	7.00
	Set B	4.42	1.72	4.00	3.00	5.00
Speed scores (RM)	Set A	11.03	3.00	10.90	9.16	12.52
	Set Ab	8.14	2.28	8.01	6.55	9.44
	Set B	8.54	2.58	8.21	6.89	10.17
Efficiency scores (CM)	Set A	7.18	2.26	7.02	5.58	8.42
	Set Ab	3.79	1.60	3.44	2.71	4.73
	Set B	3.04	1.21	2.97	2.07	3.83

Note. SD = Standard Deviation; CR = Correct Responses; RM = mean number of Responses per Minute; CM = mean number of Correct responses per Minute.

The following procedures must be adopted to obtain Accuracy (CR), Speed (RM), and Efficiency (CM) scores for a child assessed in either clinical or research context.
The time taken in minutes (MINs), seconds (SECs), and hundredths of a second (Centiseconds - CSs) must be recorded for each Set (starting from item 4 for Set A) and transformed into total Centiseconds with the following formula:
TOTAL CSs = (MINs x 600)+(SECs x 100)+CSs
The total number of minutes must be calculated with the following formula:
TOTAL MINs = TOTAL CSs / 6000
Speed (RM) scores of the three Sets must be calculated with the following formulas:
RM Set A = 9/MINs Set A
RM Set Ab = 12/MINs Set Ab
RM Set B = 12/MINs Set B
Efficiency (CM) scores of the three Sets must be calculated with the following formulas:
CM Set A = Correct Responses Set A (excluding the first 3 items)/ MINs Set A
CM Set Ab = Correct Responses Set Ab/MINs Set Ab
CM Set B = Correct Responses Set B/MINs Set B
Accuracy (CR), Speed (RM) and Efficiency (CM) scores must be converted to z scores with the following formula:
z = (children score – normative sample mean) / normative sample SD
Z scores > |2| can be considered outside of the normal range.
The R/I score can be calculated for Set A only, by computing the difference between zCR and zRM scores. Positive z scores > 2 indicate a tendency to apply a reflective strategy to the task, whereas negative z scores < -2 indicate a tendency to respond impulsively.
The scores transformed by the formulas described above are reliable only if the administration and scoring procedures adopted in the assessment faithfully follow the procedures described in the method of this work.
In the event that clinicians or researchers encounter difficulties in calculating the different scores, they can send an email to traversarilena@gmail.com to receive an excel file that automatically calculates them by entering the child's raw data.