Available online at www.sciencedirect.com

ScienceDirect

Intelligence 35 (2007) 369–380

INTELLIGENCE

Endophenotypes for intelligence in children and adolescents

Marieke van Leeuwen *, Stéphanie M. van den Berg,
Rosa A. Hoekstra, Dorret I. Boomsma

Department of Biological Psychology, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

Received 21 December 2005; received in revised form 29 September 2006; accepted 29 September 2006

Abstract

The aim of this study was to identify promising endophenotypes for intelligence in children and adolescents for future genetic studies in cognitive development. Based on the available set of endophenotypes for intelligence in adults, cognitive tasks were chosen covering the domains of working memory, processing speed, and selective attention. This set of tasks was assessed in a test–retest design in children and in adolescents. Working memory could be measured reliably using the *n*-back task and correlated with intelligence in both age groups. For processing speed, assessed with the Π-inspection time task and reaction time on the flanker task, test–retest reliability was good in both age groups, but processing speed only correlated significantly with intelligence in children. Selective attention, i.e., the effect of incongruent flankers on RT and accuracy, showed low reliability and neither correlated with intelligence in adolescents nor in children. Thus, working memory seems a promising endophenotype for intelligence in both children and adolescents. Inspection time and measures of selective attention based on the flanker task do not seem very promising endophenotypes for intelligence in these age groups.

Keywords: Processing speed; Selective attention; Test reliability; Working memory; Development

---

1. Introduction

Variance in children's IQ test performance is for 25 to 50% accounted for by genetic variation between individuals (Bartels, Rietveld, Van Baal, & Boomsma, 2002; Jacobs et al., 2001; Plomin, 2003; Rietveld, Dolan, Van Baal, & Boomsma, 2003; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003) and in adults for even more than 50% (Posthuma, De Geus, & Boomsma, 2001). In spite of this high heritability actual identification of genes is currently limited to mutations with rather severe neurological effects (De Geus, Wright, Martin, & Boomsma, 2001; Nokelainen & Flint, 2002). Genes that influence normal variation in cognitive ability in children have yet to be identified, although recently, several QTLs (quantitative trait loci, i.e. locations of genes that influence complex traits) have been suggested (Butcher, Meaburn, Dale et al., 2005; Butcher, Meaburn, Knight et al., 2005; Hewitt, 2004; Posthuma et al., 2005). One of the complexities of identifying genes affecting a complex trait such as intelligence is that it is influenced by many genes, and therefore each gene is likely to have a relatively small effect (Plomin, DeFries, McClearn, & McGuffin, 2001). The initial goal of QTL research is not to find *the* gene for intelligence, but rather those genes that contribute to different pathways that explain individual differences in intelligence (Plomin, DeFries, Craig, & McGuffin, 2002).

---

* Corresponding author. Tel.: +31 20 5988992; fax: +31 20 5988832.

E-mail address: M.van.leeuwen@psy.vu.nl (M. van Leeuwen).

0160-2896/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2006.09.008

Genetic influences on cognitive ability are likely to be mediated by a complex network of multiple subcortical and cortical brain structures each influenced in part by its own set of genes (De Geus et al., 2001). These sets of genes influencing intelligence may be localized and identified using the strategy of endophenotyping that is studying confined cognitive components or elements that relate to intelligence variability. These components can be suggested from neuroscience and may get closer to the actual biological systems involved in intelligence (De Geus & Boomsma, 2001; De Geus et al., 2001; Deary, 2001; Plomin & Spinath, 2002). It is thought that variation in confined components of intelligence may be influenced only by a subset of all the genes involved in general intelligence. The primary idea behind the endophenotypic approach is that by studying these components it may be easier to isolate and identify the effects of each of these subsets of genes. Although these genes may explain only a small part of general intelligence, they may explain a large part of the variance in the endophenotype itself, thereby improving the statistical power to detect genes for general intelligence (De Geus, 2002; De Geus et al., 2001).

The aim of this study is to identify promising endophenotypes for intelligence in childhood and adolescence that may play crucial roles in future genetic studies in cognitive development. For adults, a small set of endophenotypes for intelligence is already available, as will be outlined below. For children, however, much less is known about the suitability of these cognitive measures as endophenotypes for intelligence. Are children at all able to perform the tasks and are the measures reliable? Are the same constructs involved in children and adolescents as in adults?

A promising endophenotype for intelligence in children should be relatively stable, show reliable within-age individual differences, and should have an association with intelligence that is also theoretically meaningful. The endophenotype should also be heritable and have a strong genetic correlation with intelligence (De Geus et al., 2001). A genetic correlation is the extent to which genetic effects on one trait correlate with the genetic effects on another trait independent of the heritability of the two traits (Deary, Spinath, & Bates, 2006). Only when this correlation is strong, it does make sense to look for the genes that explain variability in the endophenotype. With a low genetic correlation, a gene variant found for the endophenotype is probably not involved in the variation of general intelligence. Based on suitable endophenotypes used in adults, promising endophenotypes for intelligence in children and adolescents may cover the following domains: working memory (particularly working memory capacity), processing speed and selective attention.

1.1. Working memory

Working memory in adults is related to intelligence. In a meta-analysis of 86 studies Ackerman, Beier, and Boyle (2005) found a correlation of .48 between working memory and intelligence. Colom, Rebollo, Palacios, Juan-Espinosa, and Kyllonen (2004) showed a large overlap in variance between working memory and general intelligence. In adults the correlation of gray and white matter volume with full scale IQ and the Working Memory dimension is completely mediated by common genetic factors (Posthuma, De Geus et al., 2002).

In a Japanese sample of young adults, using a spatial as well as a verbal working memory task, Ando, Ono, and Wright (2001) found that higher-order spatial and verbal cognitive abilities are mediated by a genetic factor they have in common with working memory. This common genetic factor explained 20–22% of variation in working memory and 64% and 26% of variation in spatial and verbal ability respectively. Thus, in adults working memory is a suitable endophenotype for intelligence.

In children, however, this is less clear. It is known that working memory and intelligence develop in concert in children. In their review on the relationships between processing speed, working memory, and fluid intelligence in children, Fry and Hale (2000) argue that much of the age related increase in intelligence in children could be attributed to developmental improvements in working memory. The greater the capacity of a child's working memory, the more information the child has available for solving problems.

In children, correlations between intelligence and working memory range from .38 to .67 (Alloway, Gathercole, Willis, & Adams, 2004; Fry & Hale, 1996; Swanson, 2004) and even .82 in children aged 4 to 6 (Swanson & Beebe-Frankenberger, 2004). In a study of De Ribaupierre and Lecerf (2006) working memory accounted for 54% of the variance in the Raven Standard Progressive Matrices (Raven, 1960) in a children and young adult sample.

Furthermore, working memory is associated with scholastic achievement. For instance, links have been found with reading ability (Cain, Oakhill, & Bryant, 2004; De Jonge & De Jonge, 1996; Gathercole, Alloway, Willis, & Adams, 2006) and solving mathematical problems (Adams & Hitch, 1997; Geary, Hoard, Byrd-Craven, & Catherine DeSoto, 2004; Swanson, 2004; Swanson & Beebe-Frankenberger, 2004; Lee, Ng, Ng, & Lim, 2004). Moreover, Van der Sluis, Van der Leij, and

De Jong (2005) found that when corrected for fluid intelligence most links between working memory and reading and arithmetic-related learning disabilities disappeared, suggesting that most of the relations between working memory and learning disabilities can be explained by IQ.

Using twins at the age of twelve years and their siblings Polderman et al. (2006) reported heritability estimates (the proportion of phenotypic differences among individuals that can be attributed to genetic differences in a particular population) for working memory capacity of 54% and 56%. Luciano, Wright et al. (2001) found in 16-year-olds that genetic differences explain 48% of the variance in working memory. The genetic correlation between intelligence and working memory is moderate in this age group; a phenotypic correlation (the correlation between observed characteristics) of .26 between intelligence and a delayed response task and a genetic correlation of .34 was found.

Working memory is related to activity in the dorsolateral prefrontal cortex (DL-PFC; Casey et al., 1995), a brain area still developing during childhood (Casey, Giedd, & Thomas, 2000), and to the anterior cingulate, a brain area of which the gray matter density is positively correlated with full scale IQ in adolescents (Frangou, Chitins, & Williams, 2004). This suggests working memory is still developing in children and is related to intelligence.

Summarizing, in children as well as in adults working memory and intelligence are related. In children variance in working memory is in part explained by genetic factors, and in adolescents and adults a genetic correlation between working memory and intelligence has been found.

1.2. Processing speed

Intelligence also co-develops with information processing speed. In children and adults inspection time (a measure for processing speed) and IQ are correlated and the correlations observed in children are similar to the ones found in adults (Fry & Hale, 2000). A meta-analysis conducted by Grudnik and Kranzler (2001) indicated that inspection time and IQ correlate around −.50. De Ribaupierre and Lecerf (2006) found in a sample of children and young adults that processing speed (as measured by a task in which subjects had to judge whether two patterns were identical) accounted for 61% of the total variance in the Raven's task. Taken together, processing speed and working memory explained 67% of variance in Raven's performance. Vickers and McDowell (1996) found in a sample with children aged 8 to 10 years a correlation between inspection time and full scale IQ of −.51. Fry and Hale (1996) found a correlation of −.44 between processing speed and the Raven.

In the literature, a broad range of tasks is used to measure processing speed, ranging from simple inspection time tasks in which subjects have to distinguish the longest of two lines of different length (Brand & Deary, 1982; Luciano, Smith et al., 2001) to complex reaction time tasks in which subject have to memorize 5 digits and have to indicate whether a newly presented digit is one of the memorized digits (Neubauer, Spinath, Riemann, Borkenau, & Angleitner, 2000). This complicates the assessment of processing speed as a possible endophenotype, because the more complex an elementary cognitive task is, the higher the correlation with intelligence (Colom et al., 2004; Neubauer et al., 2000). More complex tasks are more meaningfully and more strongly related to intelligence, but on the other hand, a suitable endophenotype should not be too complex, since the more complex the task, the more genes are likely to be involved. Therefore a relatively simple task measuring processing speed should be preferred.

In adults and adolescents a high genetic correlation between inspection time and full scale IQ has been found (Luciano et al., 2005; Posthuma et al., 2001). In 13 to 15-year-olds, a genetic correlation of −.63 was reported. A common genetic factor accounted for 36% in the variance of inspection time, the remaining variance in inspection time was accounted for by a unique environmental factor (Luciano, Smith et al., 2001), suggesting that a considerable part of the variance in inspection time could consist of measurement error variance.

Posthuma et al. (2001) hypothesized that the genetic factor that influences intelligence as well as speed of processing is a factor that determines axonal myelination in the central nervous system. Fry and Hale (2000) concluded in their review that much of the age related improvement in children in intelligence is due to increases in speed and this seems to be mediated through the effect of speed on working memory, that is, the faster the brain, the more information can be retained in working memory.

Concluding, in both young adolescents and adults there is a relationship between processing speed and intelligence. Moreover, there is a genetic correlation between these two abilities in adolescents and adults. However, it is yet unknown to what extent processing speed is a suitable endophenotype for intelligence in children.

1.3. Selective attention

Concepts of selective attention are included in almost all theories of higher cognitive functioning. Dempster

(1991) claims that intelligence cannot be understood without reference to inhibitory processes. One of his arguments is that individuals who are more distractible, score generally lower on intelligence tests. Using the flanker task (Eriksen & Eriksen, 1974), Posthuma, Mulder, Boomsma, and De Geus (2002) showed a significant genetic correlation between IQ and incongruency effects (difference in performance between congruent and incongruent trials) on accuracy, varying between −.37 and −.68 depending on the cohort (old or young) and IQ-scale (verbal or performance IQ). Stins, Van Baal, Polderman, Verhulst, and Boomsma (2004) found little evidence for heritability of flanker performance in 12-year-old twins. Possible explanations for this discrepant finding can be that the task is not suitable for children and cannot reliably measure inhibitory processes, that twelve-year-olds perform the task in a more prudent way that leads to ceiling effects on accuracy, or that inhibitory processes have not yet fully developed at this age.

As far as we know no study has been done up till now relating incongruency effects to intelligence in children. The one study (Censabella & Noël, 2005) we could find reported on the relationship between incongruency effects and learning disabilities. This study could not show that children with learning disabilities exhibit significantly larger incongruency effects than children without learning disabilities.

Slowing down of reaction time and loss of accuracy in flanker task performance as a consequence of incongruencies in the stimuli may reflect an impairment in the top–down inhibitory control of the prefrontal cortex (Posthuma, Mulder et al., 2002). During performance on this task the left prefrontal cortex is activated (Fan, Flombaum, McCandliss, Thomas, & Posner, 2003), a brain area shown to be developing between the ages of 7 and 11 years (Sowell, Trauner, Gamst, & Jernigan, 2002) and implied in intelligence (Frangou et al., 2004).

A limitation of all studies relating working memory and processing speed to intelligence, is that none of these studies (Alloway et al., 2004; De Ribaupierre & Lecerf, 2006; Fry & Hale, 1996; Swanson, 2004; Swanson & Beebe-Frankenberger, 2004; Vickers & McDowell, 1996) corrected the observed relationships for the biasing effects of measurement error. When one is interested in the relationship between actual traits, rather than relationships between specific measures of traits it is important to make corrections for biases induced in research data by measurement error (Schmidt & Hunter, 1996).

Moreover, the few studies reporting short-term test–retest stability of tasks measuring processing speed, selective attention, and working memory, show that most test–retest reliabilities are rather low in children. Test–retest reliabilities for working memory in children are reported for various tasks, ranging from .52 to .76. However, most of these tasks are verbal in nature. Alloway et al. (2004) reported test–retest reliabilities for three working memory tasks in children aged 5 to 8 years. For the *backwards digit recall test* they report a reliability of .53, for the *counting recall test* (children need to count the number of dots in an array, and then recall the tallies of dots in the arrays that were presented) they report a reliability of .74, and for the *sentence completion and recall task* (the child listens to a series of short sentences with a missing word at the end, produces a word to complete the sentence, and recalls the word she or he produced for each sentence in a sequence) test–retest reliability was .52. In another study test–retest correlations were found of .54 using the *sentence completion and recall* and the *counting span task* (child needs to count yellow dots on a card with blue and yellow dots, after 2 to 5 of these cards, the child has to recall the total number of dots; Kuntsi, Stevenson, Oosterlaan, & Sonuga-Barke, 2001). Archibald and Kerns (1999) found a test–retest reliability of .76 for the *Self-Ordered pointing* in which children had to point to a different drawing in a booklet, whereby every time the location of drawings changed. Vickers and McDowell (1996) found a test–retest reliability of .30 for an inspection time task where children had to discriminate which of the two lines was the longest. For the standard Stroop task for interference a test–retest reliability of .81 over three sessions has been reported by Neyens and Aldenkamp (1996).

In the current study, test–retest reliability will be investigated for various tasks measuring working memory, processing speed, and selective attention to identify promising endophenotypes for intelligence in children and adolescents. For working memory the *n*-back task was used (Casey et al., 1995), information processing speed was assessed using the π-inspection time task (Luciano, Smith et al., 2001) and the flanker task (Eriksen & Eriksen, 1974). The flanker task was also used to measure selective attention. All tasks were specifically adapted for children. In order to assess the relationship between performance on these tasks and intelligence, all correlations were corrected for test reliability.

2. Materials and methods

2.1. Subjects

Three groups of subjects participated in this study. The first group consisted of 108 children who were recruited from the 5th grade of six primary schools located in different social economic areas in the Netherlands. 105 children returned two to three weeks later for retest, children were 8–11 years of age (M=8.7, SD=.6). Of these children, 55.4% were female. After completing the test protocol children received a present worth €5. The second group consisted of 98 children participating in an ongoing longitudinal study recruited via the Netherlands Twin Registry (NTR). For the current study one twin or sibling was randomly selected from a family (age: M=9.7, SD=1.1, 52% female). After participation they received a present of €10. The third group consisted of 30 adolescents in the age range of 14 to 20 (M=18.4, SD=1.6), from which 29 returned for retest. In this group 70% was female. Adolescents received a token of €25. When children were under 14, their parents signed an informed consent form. If not, participants signed an informed consent themselves.

2.2. Testing procedures

Children in the first group were administered the *n*-back task, Eriksen flanker task, and π-inspection time task, as part of a larger neuropsychological test battery. Because of practical reasons not all children were administered the complete battery. Administration of the complete battery required approximately 50 min. Children in these group were individually tested during school hours in a quiet room at school. Children in the second group and adolescents were individually tested at the Vrije Universiteit. They were administered the π-inspection time task, Eriksen flanker task and *n*-back task at the end of a larger test battery – including the Wechsler Adult Intelligence Scale-III (WAIS-III) or Wechsler Intelligence Scale for Intelligence-III (WISC-III) – of which administration required approximately 4.5 h. Retest for children in the first group as well as adolescents took place two to three weeks after initial testing. There was no retest for the WAIS-III and no retest session for children in the second group.

2.3. n-back task

Subjects performed a spatial variant of the *n*-back task, designed after Gevins and Cutillo (1993) and Jansma, Ramsey, Coppola, and Kahn (2000), with increasing levels of difficulty. The task was adapted to make it more attractive for children. Subjects had to look at an apple presented on a screen. The apple had four holes in which a caterpillar could appear. The participants were told to catch the caterpillar to prevent it from eating the apple. They were instructed to respond to the caterpillar by pushing one of four buttons with thumb and index finger of both hands. The layout of the four buttons corresponded spatially to the four holes in which the caterpillar could appear. Subjects had to indicate where the caterpillar was one move back (1-back), two moves back (2-back) or three moves back (3-back). Adolescents received also a session with a delay of 4 moves (4-back). The caterpillar appeared in a hole for 1 s; after its disappearance there was a warning sound. Subjects had to respond after this warning. Between two caterpillar moves, the apple was empty for 1 s.

Sessions were given in blocks of 20 trials. After each block participants received feedback on the number of apples they had saved from the caterpillar (correct button presses) and how many had been eaten (incorrect button presses). The 1-back condition consisted of a practice block only. The 2-back, 3-back and 4-back conditions contained one practice block, and three blocks in which performance was measured. Practice blocks were added if the subject did not understand the task. Children were motivated during the task by counting the moves of the caterpillar. In the 2-back version the test administrator counted continuously to three and in the 3-back version the administrator counted to four. The task requires that subjects have to respond to all stimuli and continuously have to monitor and update all movements of the caterpillar. Performance on the task was scored by using the total number of correct responses. Maximum score per condition was 60.

2.4. π-inspection task

The π-inspection task was designed after Luciano, Smith et al. (2001). For this task subjects had to identify the longer of two lines which were presented by the test administrator as worms that the subjects needed to catch. This task was complicated by the fact that the worms burrowed quickly into the ground (i.e., disappeared quickly from the screen). If subjects caught five worms they had enough worms to catch a fish, which would appear at the lower left-hand side of the screen. It was stressed that it was important to be accurate and that it did not matter how long it took them to catch the worms.

The vertical lines measured 22 and 27 mm in length, were 9 mm apart, and joined at the top to a horizontal line 12 mm long. The probability of the longer line appearing on the left or right was equal. The stimulus duration ranged between 14.2 and 2000 ms. A dynamic mask, consisting of two vertical lines (37 mm) shaped as lightning bolts, immediately followed the stimulus and was presented for 300 ms to limit further stimulus processing. On each trial, a fixation cross appeared in the center of the screen for 1 s (an alerting beep was sounded for 100 ms at the onset of the dot), followed by a blank screen for 100 ms. The π figure was then presented, and the participant's first response (left or right) was noted. The screen was blanked for 750 ms before the next trial was presented. This produced an effective response-stimulus interval of approximately 2 s. Stimulus duration depended on the response of the subject. The initial duration was 210 ms. For every four correct consecutive responses the stimulus duration was decreased, and for every incorrect response the stimulus duration was increased by a step size depending on previous performance (for details of the actual algorithm, see Luciano, Smith et al., 2001). The algorithm decided when minimal stimulus duration was achieved and stopped the program. If the minimum was not reached within 96 trials the program also stopped. This way, for each subject the minimal stimulus duration time could be assessed. On each trial the stimulus duration and whether or not the correct key was pressed, was stored. Data of the subjects were excluded from the analyses when they responded more often than 10 times before the stimulus had disappeared. The stimulus duration at the last correct trial was used as the measure for inspection time.

2.5. Eriksen flanker task

In the Eriksen flanker task (Eriksen & Schultz, 1979) subjects were presented with a horizontal array of five arrows. Before each trial a fixation cross was presented for 1000 ms. The stimulus was then presented for 200 ms. Between two trials there was a random interval of either 2000, 2250, 2500, 2750 or 3000 ms. Subjects were instructed to pay attention to the direction of the center arrow and ignore the four flanking ones. They were told that the cross in the center of the screen between two trials could help them to focus on the middle arrow. Subjects had to press the left key to a left facing central arrow, and the right key to a right facing central arrow. The flanking arrows could either all point in the same direction as the target arrow (<<<<< or >>>>>; congruent), or they all pointed in the opposite direction (<<><< or >><>>; incongruent). Children received 40 congruent and 40 incongruent trials in random order after an eight trial practice session. After each ten correct responses a smiley was presented in the center of the screen. On each trial the reaction time (RT) and whether or not the correct key was pressed, was stored. Maximum score was 40 correct congruent and 40 correct incongruent trials for the children. The adolescents performed a shorter version of the task; this was 20 congruent and 20 incongruent trials. Trials in which reaction time was below 300 ms or exceeded 1500 ms were excluded from analysis. If, as a result of this rule, more than 25% of a subject's trials were excluded, all data from this subject's session were excluded from analysis. Average RT was calculated over the accurate trials. Average RTs on incongruent trials were subtracted from average RTs on congruent trials and served as a measure for selective attention. The same applies to the error rates that were similarly subtracted as an additional measure for selective attention.

2.6. Data analysis

Test–retest reliability and correlations between the endophenotypes and IQ were calculated with Pearson correlation coefficients. Correlations between IQ and the endophenotypes were subsequently corrected for test–retest reliability using the disattenuation formula $\boldsymbol{r}_{xtyt}=\boldsymbol{r}_{xy}/(\boldsymbol{r}_{xx}\boldsymbol{r}_{yy})^{1/2}$, where $\boldsymbol{r}_{xtyt}$ is the correlation between the true scores of the measures x and y, $\boldsymbol{r}_{xy}$ is the observed correlation, and $\boldsymbol{r}_{xx}$ and $\boldsymbol{r}_{yy}$ are the reliabilities of x and y, respectively (Schmidt & Hunter, 1996). For the reliability of the Dutch WISC-III a Cronbach's α of .93 was used (Wechsler et al., 2002) and for the reliability of the Dutch WAIS-III a test–retest correlation was used of .94 (Kessels & Wingbermühle, 2001). Reaction time and inspection time data were log transformed prior to analysis since their distributions were positively skewed.

A partial correlation analysis was conducted to determine the independent contributions of the endophenotypes to IQ. Only variables were included that showed

Table 1
Problems encountered during testing: (I) children of the first group, (II) children of the second group, (A) adolescents

	Test	Retest
2-back	I: 2 subjects not recorded	
3-back	I: 2 subjects not able	I: 3 subjects not able/ did not want to participate
	II: 3 subjects not able	
4-back	A: 1 subject not able	
π-task	I: 8 subjects excluded	I: 2 subjects excluded
	II: 8 subjects excluded	
		A: 1 subject excluded
flanker	I: 1 subject excluded	I: 2 subjects excluded
	II: 2 subjects excluded	
		A: 1 subject excluded

Note: For the π-task, subjects were excluded when they pressed the keys too early more often than 10 times. For the flanker task, subjects were excluded when more than 25% of their RTs was either <300 or >1500 ms.

Table 2
Descriptive statistics and test–retest correlations for *n*-back, flanker and π-task in children of group I (accuracy is reported in proportion correct, reaction time in ms)

	N	Mean (SD)-test	*N* retest	Mean (SD)-retest	*N*	r (95% confidence intervals)
2-back	59	.51 (.24)	59	.66 (.26)	58	.65 (.51–1.0)
3-back	58	.40 (.15)	56	.50 (.19)	56	.70 (.60–1.0)
flanker RT congruent	76	566 (98)	74	563 (91)	74	.66 (.56–1.0)
flanker RT incongruent	76	668 (137)	74	638 (106)	74	.62 (.49–.96)
flanker incongruency effect (RT)	76	102 (86)	74	75 (57)	74	.48 (.29–.76)
flanker Acc. congruent	76	.95 (.07)	74	.96 (.08)	74	.06 (−.17–.29)
flanker Acc. incongruent	76	.85 (.19)	74	.90 (.15)	74	.46 (.26–.73)
flanker incongruency effect (Acc.)	76	.10 (.08)	74	.06 (.06)	74	.29 (.07–.53)
π-task	68	164 ms (61)	70	128 ms (50)	63	.65 (.52–1.00)

a significant correlation with IQ in one of the two age groups in the previous correlation analyses. To determine how much variance in IQ scores was explained by the endophenotypes, a multiple regression analysis was conducted, in which all variables were entered simultaneously.

3. Results

Mean IQ score of the children in the second group was 101.6 (SD = 14.3). For the adolescents the mean IQ was 108.4 (SD = 12.2). Table 1 presents a description of the problems encountered during testing. In general the following problems were encountered: during administration of the *n*-back, it was observed that some children were not able to push the button while at the same time paying attention to where to caterpillar went. In the π-inspection task sessions were excluded, because children and adolescents pushed the button before the π figure disappeared. From the Eriksen flanker task data from some subjects were excluded from analysis, because more than 25% of data were excluded (reaction times were below 300 ms or exceeded 1500 ms).

Results of testing and test–retest reliabilities in children and adolescents are presented in Tables 2 and 3. Because specific abilities in children are tested, reliabilities of .7 or higher are considered satisfactory, whereas reliabilities of .5 and .6 are considered modest (Kuntsi et al., 2001). As shown in Table 2, all test–retest correlations in children exceeded .60, except for accuracy and stimulus congruency effects of the flanker task. The low accuracy test–retest correlation is very likely due to ceiling effects. For the adolescents the same holds true, with an exception of the 2-back and the π-inspection task. The low test–retest correlation on the 2-back can be explained by ceiling effects at the second time of testing.

Table 4 presents observed correlations and corrected correlations between IQ and the endophenotypes. In children and adolescents *n*-back performance was significantly related to IQ. Better performance on the *n*-back task was related to higher IQ-scores. No correction for test–retest reliability is reported for the 2-back task in adolescents, since this reliability is influenced by ceiling effects. Reaction time on the congruent and incongruent trials of the flanker was significantly related to IQ for children only; the longer the reaction time, the

Table 3
Descriptive statistics and test–retest correlations for *n*-back, flanker, and π-task in adolescents (accuracy is reported in proportion correct, reaction time in ms)

	N	Mean (SD)	*N* retest	Mean retest (SD)	*N*	r (95% confidence intervals)
2-back	30	.89 (.15)	29	.96 (.09)	29	.16 (−.22–.55)
3-back	30	.72 (.17)	29	.84 (.14)	29	.70 (.48–1.0)
4-back	29	.61 (.15)	29	.69 (.17)	28	.66 (.40–1.0)
flanker RT congruent	30	434 (72)	28	419 (56)	28	.66 (.40–1.0)
flanker RT incongruent	30	495 (72)	28	475 (61)	28	.65 (.38–1.0)
flanker incongruency effect (RT)	30	61 (29)	28	56 (28)	28	.48 (.13–.91)
flanker Acc. congruent	30	.97 (.05)	28	.96 (.08)	28	.42 (.06–.84)
flanker Acc. incongruent	30	.96 (.08)	28	.94 (.08)	28	.35 (−.03–.76)
flanker incongruency effect (Acc.)	30	.01 (.06)	28	.03 (.07)	28	.14 (−.25–.53)
π-task	28	94 ms(35)	26	70 ms (17)	25	.58 (.24–1.0)

Table 4
Observed correlations ($\boldsymbol{r_{xy}}$) and correlations corrected for test–retest ($\boldsymbol{r_{xtyt}}$) of *n*-back, flanker, and π-task with IQ in children and adolescents

	Children			Adolescents		
	N	r_{xy}	r_{xtyt}	*N*	r_{xy}	r_{xtyt}
2-back	94	.41**	0.53	30	.66**	–
3-back	95	.44**	0.55	30	.55**	0.68
4-back	–	–		29	.40*	0.51
flanker RT congruent	96	−.35**	−0.45	30	−.07	−0.09
flanker RT incongruent	96	−.35**	−0.46	30	−.06	−0.08
flanker incongruency effect (RT)	96	−.04	−0.06	30	.14	0.21
flanker Acc. congruent	96	−.10	−0.42	30	.10	0.16
flanker Acc. incongruent	96	−.06	−0.09	30	.02	0.03
flanker incongruency effect (Acc.)	96	−.02	−0.04	30	−.06	−0.17
π-task	88	−.28**	−0.36	28	−.33	−0.45

Note: **=$p<.01$, *=$p<.05$.

lower the IQ. Incongruency effects on reaction time, accuracy on the congruent and incongruent trials, as well as incongruency effects on accuracy were not related to IQ in children or in adolescents. Inspection time was related to IQ in children, the shorter the inspection time the higher the IQ, but was not significantly related to IQ in adolescents.

In Table 5 the results of the partial regression analyses are presented. It is important to note that in this table only the subjects are included for whom data are available on all tasks. This leads to a lower correlation between IQ and 3-back performance and flanker reaction time in children due to a biased sample of relatively smarter subjects: children who performed poorly had a higher probability of not being able to perform all of the cognitive tasks sufficiently well. As can be seen in Table 5, none of the tasks contributed completely independently to the variance in intelligence in children. The π-task and the flanker task did not contribute any significant independent part of the variance in intelligence. The 3-back showed significant covariance with intelligence even after the flanker and π-task had been controlled for. Also in adolescents, the 3-back task contributed to the variance in intelligence independently of the other two tasks, while the contribution of performance on the π-task could partly be explained by performance on the 3-back task.

Regression analyses revealed that 2-back, 3-back, flanker reaction time on the congruent and incongruent trials and inspection time could explain a total of 17% (adjusted R^2) of the variance in IQ in children ($R=.47$, $F(5, 78)=4.49$, $p<.001$). Since it was clear from the other analyses that flanker reaction time did not contribute to the variance in intelligence, this variable was not included in the regression analysis in adolescents. In this analysis 2-back, 3-back, 4-back and inspection time could explain a total of 45% (adjusted R^2) in the variance of intelligence ($R=.73$, $F(4, 22)=6.28$, $p<.01$).

4. Discussion

The goal of this study was to determine whether endophenotypes for intelligence previously used in adults and sometimes in adolescents are promising endophenotypes for intelligence in children and adolescents. A good endophenotype for intelligence must meet the following criteria (De Geus et al., 2001): it must be a reliable trait, it must show evidence of genetic influence, it must be associated with intelligence, the association between endophenotype and intelligence must derive partly from the same genetic source (i.e., there should be a genetic correlation) and the association between endophenotype and intelligence must be theoretically meaningful.

In this paper we examined the reliability and the relation to intelligence of three candidate endophenotypes: working memory, processing speed, and selective attention. The choice for these three endophenotypes was based on prior research, which was mainly conducted in adults.

Table 5
Correlations of 3-back, flanker RT congruent and π-task with IQ controlling for respectively 3-back accuracy, flanker reaction time in congruent trials and π-inspection time

R controlled for:	–	3-back	flanker	π-task	3-back & flanker	3-back & π-task	flanker & π-task
Children (N=84)							
3-back	.42**	–	.36**	.35**	–	–	.32**
flanker RT congruent	−.31**	−.22*	–	−.24*	–	−.19	–
π-task	−.28**	−.15	−.20	–	−.10	–	–
Adolescents (N=27)							
3-back	.57**	–	.57**	.54***	–	–	.54***
flanker RT congruent	−.03	−.12	–	−.01	–	−.09	–
π-task	−.31	−.23	−.31	–	−.22	–	–

4.1. Working memory

A spatial version of the *n*-back task was used, specifically adapted to measure working memory in children and adolescents. In children 2-back as well as 3-back performance could be measured reliably and in adolescents the *n*-back 3 and 4 could be measured reliably. This result is comparable to the test–retest reliability reported by Hockey and Geffen (2004) who found a test–rest correlation in students on the 3-back of .73. It is also comparable to the results from a spatial working memory task reported by Archibald and Kerns (1999).

In both children and adolescents performance on the *n*-back 2 and 3 was correlated with IQ. In the adolescent group performance on the 4-back task was also correlated with intelligence. The observed correlations in this study are comparable to correlation reported in previous studies (Ackerman et al., 2005; Alloway et al., 2004; Fry & Hale, 1996; Swanson, 2004).

In this study we found a lower correlation between *n*-back performance and intelligence in children as compared to adolescents. This may be due to the lower performance levels of the children: some children were not able to push a button and meanwhile attend where the caterpillar was going. As a consequence in children the task may measure short-term memory, rather than working memory. This interpretation is in line with that of De Jonge and De Jonge (1996) who reported that no distinction could be made between tasks measuring short-term and working memory in children aged 10 to 12 years. This is not surprising since the prefrontal cortex, and particularly the DL-PFC, a brain region involved in working memory, appears to be the last brain region to mature (Casey et al., 2000).

Based on the findings from previous studies as well as from our own, it can be concluded that working memory as measured by the *n*-back task is a suitable endophenotype for intelligence in adolescents and to a somewhat lesser extent in children. Future research should establish whether there is a genetic correlation between performance on this task and intelligence in children and adolescents. Whether in children the task actually measures working memory rather than short-term memory is still a matter of discussion.

4.2. Processing speed

The π-inspection time task was used to measure processing speed. In children as well as adolescents this task showed good test–retest correlations. Test–retest reliability was substantially higher than the one reported in the study of Vickers and McDowell (1996). This discrepancy can possibly be explained by the reward incorporated in our task, which keeps children motivated during the task. The correlation between intelligence and processing speed was lower than what has been found in previous studies (Grudnik & Kranzler, 2001). One explanation for this finding is that in studies in which higher correlations with intelligence have been reported, more complex measures for processing speed were used, like for instance Sternberg's memory scanning task. As stated by Neubauer et al. (2000) and Colom et al. (2004) the more complex an elementary cognitive task is, the higher the correlation with intelligence. At first glance it may seem that inspection time is a suitable endophenotype for children. However, it must be noted that in children the average inspection time was quite long and showed large variation. This suggests that in children whose inspection times are long, it may not be the speed of processing that was measured. Many children seemed somehow unable to deal with this task, showing inspection times of over 500 ms. It is unclear what the task actually measured and therefore the test may be unsuitable as an endophenotype for intelligence in children of this age. An endophenotype must be simple to interpret, since its goal is to facilitate the search for genes. One of the prerequisites of an endophenotype is that the relationship with the phenotype of interest must be theoretically meaningful.

In adolescents inspection time was not significantly related to IQ. This result may be caused by a lack of statistical power, since the effect size is similar to the one reported by Luciano, Smith et al. (2001) and Posthuma et al. (2001). Nevertheless, the amount of variation in intelligence it might explain is limited, particularly after correction for working memory performance (cf. Fry & Hale, 2000). This finding suggests that inspection time task is of limited added value as an endophenotype in a test battery including a working memory task.

4.3. Selective attention

The flanker task did not measure accuracy and incongruency effects on RT reliably. The task measured reaction time on congruent and incongruent trials reliably, but it can be argued that reaction time is a measure of processing speed rather than selective attention (Fry & Hale, 1996). In adolescents and children we found no evidence for a relationship between intelligence and incongruency effects. Therefore, it can be concluded that selective attention as measured by flanker incongruency effects is not a suitable endophenotype for intelligence in children and adolescents.

The partial correlation analysis showed that in children working memory as measured by the 3-back task contributed a significant, though not completely independent, part to the variance in intelligence. Processing speed as measured by the flanker task also contributed a small part to the variance of intelligence, though not significantly or independently. In adolescents, working memory as measured by the 3-back contributed a significant part to intelligence that could not be explained by performance on the flanker or π-task. When controlled for working memory and processing speed as measured by the flanker task, no significant contribution of inspection time was left in children or in adolescents.

When exploring the variance contributed by the different tasks to intelligence it becomes clear that the same tasks explain more variance in adolescents than in children. A possible reason for this is that performance on these tasks in children is influenced by different and unknown processes which do not play a role in adolescents. This finding illustrates that the search for endophenotypes in children may be more complex than in adolescents and adults. Another explanation for the lower contribution of variance by the different cognitive tasks to the variation in intelligence in young children, is that in children variation in intelligence is less influenced by genes than in adolescents and in adults. It is possible that the association of IQ and working memory is mainly due to genetic covariation which will become more pronounced with increasing age.

To conclude, working memory capacity seems a good endophenotype for intelligence in children and adolescents: it can be reliably assessed using our version of the *n*-back and it correlates with intelligence. Processing speed is not an optimal endophenotype for intelligence in children (as measured by reaction time on the flanker task) and adolescents (as measured by the π-task). Once corrected for working memory, it contributes only a very small part to the variance of intelligence. Selective attention, at least when measured as the flanker incongruency effect on RT and accuracy, is not a suitable endophenotype for neither age groups. Future studies will be directed at investigating whether in children and adolescents, working memory is sufficiently heritable and genetically correlated with intelligence to be of use in QTL research.

Acknowledgements

We wish to thank the children and the staff of the following schools for the time and assistance they contributed to this study: Flevoparkschool in Amsterdam, St Michael in De Bilt, Cammingha RK in Bunnik, ABC RK basis school locatie de Bijenkorf in Gouda, and De Paltrok and De Wissel in Uitgeest. We gratefully acknowledge support from grants by the Netherlands Organisation for Scientific Research, NWO 051-02-060, 668-772 and NWO/SPI 56-464-14192, and the Centre for Neurogenomics and Cognitive Research, Amsterdam.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30−60.

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*, 21−38.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*, 85−106.

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior Genetics*, *31*, 615−624.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*, 115−129.

Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics*, *32*, 237−249.

Brand, C. R., & Deary, I. J. (1982). Intelligence and "inspection time". In H. J. Eysenck (Ed.), *A model for intelligence* New York: Springer.

Butcher, L. M., Meaburn, E., Dale, P. S., Sham, P., Schalkwyk, L. C., Craig, I. W., et al. (2005). Association analysis of mild mental impairment using DNA pooling to screen 432 brain-expressed single-nucleotide polymorphisms. *Molecular Psychiatry*, *10*, 384−392.

Butcher, L. M., Meaburn, E., Knight, J., Sham, P. C., Schalkwyk, L. C., Craig, I. W., et al. (2005). SNPs, microarrays and pooled DNA: Identification of four loci associated with mild mental impairment in a sample of 6000 children. *Human Molecular Genetics*, *14*, 1315−1325.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31−42.

Casey, B. J., Cohen, J. D., Jezzard, P., Turner, R., Noll, D. C., Trainor, R. J., et al. (1995). Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *NeuroImage*, *2*, 221−229.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*, 241−257.

Censabella, S., & Noël, M. P. (2005). The inhibition of exogenous distracting information in children with learning disabilities. *Journal of Learning Disabilities*, *38*, 400−410.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*, 277−296.

De Geus, E. J. C. (2002). Introducing genetic psychophysiology. *Biological Psychology*, *61*, 1−10.

De Geus, E. J., & Boomsma, D. I. (2001). A genetic neuroscience approach to human cognition. *European Psychologist*, *6*, 241−253.

De Geus, E. J., Wright, M. J., Martin, N. G., & Boomsma, D. I. (2001). Genetics of brain function and cognition. *Behavior Genetics*, *31*, 489−495.

De Jonge, P., & De Jonge, P. F. (1996). Working memory, intelligence and reading ability in children. *Personality and Individual Differences*, *21*, 1007−1020.

De Ribaupierre, A., & Lecerf, T. (2006). Relationships between working memory and intelligence from a developmental perspective: Convergent evidence from a neo-Piagetian and a psychometric approach. *European Journal of Cognitive Psychology*, *18*, 109−137.

Deary, I. J. (2001). Human intelligence differences: Towards a combined experimental-differential approach. *Trends in Cognitive Sciences*, *5*, 164−170.

Deary, I. J., Spinath, F. M., & Bates, T. C. (2006). Genetics of intelligence. *European Journal of Human Genetics*, *14*, 690−700.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, *15*, 157−173.

Eriksen, B. A., & Eriksen, C. -W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143−149.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, *25*, 249−263.

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

Frangou, S., Chitins, X., & Williams, S. C. R. (2004). Mapping IQ and gray matter density in healthy young people. *NeuroImage*, *23*, 800−805.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237−241.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1−34.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, *93*, 265−281.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & Catherine DeSoto, M. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*, 121−151.

Gevins, A., & Cutillo, B. (1993). Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and Clinical Neurophysiology*, *87*, 128−143.

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, *29*, 523−535.

Hewitt, J. K. (Ed.). (2004). *Variance components methods for mapping quantitative trait lociBehavior Genetics*, *Vol. 34*. (pp. 125−215).

Hockey, A., & Geffen, G. (2004). The concurrent validity and test–retest reliability of a visuospatial working memory task. *Intelligence*, *32*, 591−605.

Jacobs, N., Van Gestel, S., Derom, C., Thiery, E., Vernon, P., Derom, R., et al. (2001). Heritability estimates of intelligence in twins: Effect of chorion type. *Behavior Genetics*, *31*, 209−217.

Jansma, J. M., Ramsey, N. F., Coppola, R., & Kahn, R. S. (2000). Specific versus nonspecific brain activity in a parametric n-back task. *NeuroImage*, *12*, 688−697.

Kessels, R., & Wingbermühle, E. (2001). De WAIS-III^{Nl} als neuropsychologisch instrument [The WAIS-III^{Nl} as neuropsychological instrument]. *De Psycholoog*, *36*, 296−299.

Kuntsi, J., Stevenson, J., Oosterlaan, J., & Sonuga-Barke, E. J. S. (2001). Test–retest reliability of a new delay aversion task and executive function measures. *British Journal of Developmental Psychology*, *19*, 339−348.

Lee, K., Ng, S. F., Ng, E. L., & Lim, Z. Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology*, *89*, 140−158.

Luciano, M., Posthuma, D., Wright, M. J., De Geus, E. J., Smith, G. A., Geffen, G. M., et al. (2005). Perceptual speed does not cause intelligence, and intelligence does not cause perceptual speed. *Biological Psychology*, *70*, 1−8.

Luciano, M., Smith, G. A., Wright, M. J., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). On the heritability of inspection time and its covariance with IQ: A twin study. *Intelligence*, *29*, 443−457.

Luciano, M., Wright, M., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behavior Genetics*, *31*, 581−592.

Neubauer, A. C., Spinath, F. M., Riemann, R., Borkenau, P., & Angleitner, A. (2000). Genetic and environmental influences on two measures of speed of information processing and their relation to psychometric intelligence: Evidence from the german observational study of adult twins. *Intelligence*, *28*, 267−289.

Neyens, L. G. J., & Aldenkamp, A. P. (1996). Stability of cognitive measures in children of average ability. *Child Neuropsychology*, *3*, 161−170.

Nokelainen, P., & Flint, J. (2002). Genetic effects on human cognition: Lessons from the study of mental retardation syndromes. *Journal of Neurology, Neurosurgery and Psychiatry*, *72*, 287−296.

Plomin, R. (2003). Genetics, genes, genomics and g. *Molecular Psychiatry*, *8*, 1−5.

Plomin, R., DeFries, J. C., Craig, I. W., & McGuffin, P. (2002). Behavioral genetics. In R. Plomin, J. C. DeFries, I. W. Craig, & P. McGuffin (Eds.), *Behavioral genetics in the post genomic era* (pp. 3−15). Washington D.C: American Psychological Association.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral genetics*, 4 ed. New York: Worth Publisher.

Plomin, R., & Spinath, F. M. (2002). Genetics and general cognitive ability (g). *Trends in Cognitive Sciences*, *6*, 169−176.

Polderman, T. J. C., Stins, J. F., Posthuma, D., Gosso, M. F., Verhulst, F. C., & Boomsma, D. I. (2006). The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence*, *34*, 549−560.

Posthuma, D., De Geus, E. J., Baare, W. F., Hulshoff Pol, H. E., Kahn, R. S., & Boomsma, D. I. (2002). The association between brain volume and intelligence is of genetic origin. *Nature Neuroscience*, *5*, 83−84.

Posthuma, D., De Geus, E. J., & Boomsma, D. I. (2001). Perceptual speed and IQ are associated through common genetic factors. *Behavior Genetics*, *31*, 593−602.

Posthuma, D., Luciano, M., Geus, E. J., Wright, M. J., Slagboom, P. E., Montgomery, G. W., et al. (2005). A genomewide scan for intelligence identifies quantitative trait loci on 2q and 6p. *American Journal of Human Genetics*, *77*, 318−326.

Posthuma, D., Mulder, E. J., Boomsma, D. I., & De Geus, E. J. (2002). Genetic analysis of IQ, processing speed and stimulus-response incongruency effects. *Biological Psychology*, *61*, 157−182.

Raven, J. C. (1960). *Guide to the standard progressive matrices.* London: H.K. Lewis & Co. Ltd.

Rietveld, M. J. H., Dolan, C. V., Van Baal, G. C. M., & Boomsma, D. I. (2003). A twin study of differentiation of cognitive abilities in childhood. *Behavior Genetics*, *33*, 367−381.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*, 199−223.

Sowell, E. R., Trauner, D. A., Gamst, A., & Jernigan, T. L. (2002). Development of cortical and subcortical brain structures in childhood and adolescence: A structural MRI study. *Developmental Medicine and Child Neurology*, *44*, 4−16.

Stins, J. F., Van Baal, G. C., Polderman, T. J., Verhulst, F. C., & Boomsma, D. I. (2004). Heritability of Stroop and flanker performance in 12-year old children. *BioMed Central Neuroscience*, *5*, 49.

Swanson, H. L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory & Cognition*, *32*, 648−661.

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *96*, 471−491.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*, 623−628.

Van der Sluis, S., Van der Leij, A., & De Jong, P. F. (2005). Working memory in Dutch children with reading- and arithmetic-related LD. *Journal of Learning Disabilities*, *38*, 207−221.

Vickers, D., & McDowell, A. (1996). Accuracy in the frequency accrual speed test (FAST), inspection time and psychometric intelligence in a sample of primary school children. *Personality and Individual Differences*, *20*, 463−469.

Wechsler, D., Kort, W., Compaan, E. L., Bleichrodt, N., Resing, W. C. M., Schittekatte, M., et al. (2002). *Handleiding WISC-III^{Nl} [Manual WISC-III^{Du}].* London: The Psychological Corporation Limited, Nederlands Instituut van Psychologen Dienstencentrum.