Heritability of Selective Attention and Working Memory in Preschoolers

J. F. Stins,[1,4] Leo M. J. de Sonneville,[2] Alexia S. Groot,[1] Tinca C. Polderman,[1,3] Caroline G. C. M. van Baal,[1] and Dorret I. Boomsma[1]

In this study aspects of selective attention and working memory were tested in a large sample of nearly 6-year old monozygotic and dizygotic twin pairs, using a computerized test battery (Amsterdam Neuropsychological tasks). In the selective attention task the presence of a foil signal (target signal at an irrelevant location) resulted in more false alarms than a non-target signal. In the working memory task an increase in memory load lead to an increase in response times and errors. We analyzed variations in absolute performance parameters (overall speed and accuracy) and relative performance parameters (increase in errors and/or reaction time). The results showed clear familial resemblances on performance. It proved difficult to ascribe these effects to shared genes or to shared environment. An exception was memory search rate, which was clearly heritable.

In order to study the genetics of childhood psychopathologies such as ADHD it is of great importance to identify endophenotypes that predict the liability of the pathology of interest. The advantage of an endophenotypic approach to study complex disorders (or complex traits, such as IQ) is that it allows the identification of genes that by themselves make only a small contribution to the trait under study. A minimal requirement for suitable endophenotypes is that they yield continuously quantifiable measures; another requirement is that endophenotypes are anchored in neuroscience (e.g., Castellanos

and Tannock, 2002, and De Geus and Boomsma, 2001). Indeed, researchers of psychopathologies often make use of neurocognitive tasks that are known to activate particular brain systems, and that yield quantitative performance measures, usually reaction times and percentage of errors. For example, Bush *et al.*, (1999), using a counting Stroop task, found that children with ADHD showed underactivation in the anterior cingulate cognitive division (ACcd), and a concurrent increase in distractor interference, relative to controls. Stroop performance may thus qualify as a suitable (cognitive) endophenotype, indicative of ADHD. However, the search for endophenotypes should also be guided by the fact that many psychopathologies are highly heritable, and a recent line of inquiry has started to focus on the heritabilities of the endophenotypes themselves.

Twin, adoption and family studies have shown that many psychopathologies are highly heritable (for review, see Acosta *et al.*, 2004). For example, the heritability of ADHD is estimated to be around between 50% and 90% (e.g., Thapar *et al.*, 1999). In a similar vein, the heritability of attention problems as

[1] Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands.
[2] Department of Pediatrics, Vrije Universiteit Medical Centre Amsterdam, Postbus, 7057, 1007 MB, Amsterdam, The Netherlands.
[3] Department of Child and Adolescent Psychiatry, Erasmus MC–Sophia, Dr. Molewaterplein 60, 3015 GJ Rotterdam, The Netherlands.
[4] To whom correspondence should be addressed at Department of Biological Psychology, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. Tel: +31-020-5922729. Fax: +31-020-5988832. e-mail: dorret@psy.vu.nl.

established by questionnaires is estimated to be around 70–90% (e.g., Rietveld *et al.*, 2003). Whole genome scans have identified promising chromosomal regions that are associated with ADHD (e.g., Bakker *et al.*, 2003; Fisher *et al.*, 2002). The results from these two genome scans have, however, not identified any overlapping regions. The process of gene finding may benefit from the development of suitable tasks that not only activate the affected brain regions, but that are also heritable. For example, if ADHD is heritable, and if performance on certain response conflict tasks is indicative of ADHD, then it should be the case that performance on conflict tasks itself has a heritable component. This knowledge, combined with an understanding of the brain regions that are involved in resolving response conflict, can guide the search for genes that are involved in abnormal brain functioning of ADHD.

The aim of the current paper is explore which individual differences in neurocognitive performance at a young age are mediated by genetic factors. To this end, a large sample of nearly 6-year old MZ and DZ twin pairs performed a selective attention task and a working memory (WM) task using a computerized test battery. Not a lot is known about neurocognitive performance at this particular age, and studies that examined heritabilities of childhood neurocognitive performance are rare (e.g., Groot *et al.*, 2004; Slaats-Willemse *et al.*, 2003). However, it is known that attention problems are highly heritable at this age (van Beijsterveldt *et al.*, 2004; Groot *et al.*, 2004). Longitudinal and cross-sectional studies have shown that heritability of some traits, such as IQ, increases with increasing age (e.g., Bartels *et al.*, 2002), and that heritability of other traits tends to be relatively stable across different ages (e.g., attention problems; Rietveld *et al.*, 2004). If we want to gain a better understanding of the genetic pathways of childhood psychopathologies, it is of importance to study not only neurocognitive performance at different ages, but also heritability of performance indices at different ages. In this study we took the first step in examining the heritability of attentional skills in a population of healthy young children.

The subject group consisted of 474 nearly 6-year old twins. Groot *et al.*, (2004) have already published results from the same group on the ability to maintain the alert state and the ability to withhold a prepotent response. Deficits in these functions are related to ADHD (e.g., Berger and Posner, 2000; Slaats-Willemse *et al.*, 2003), and one of the key findings in our previous study was that there were clear familial resemblances on these functions, although it proved impossible to ascribe these resemblances to shared environmental influences or to shared genes. In the present study we focus on another deficit in ADHD, viz. deficits related to executive functions (EF). According to some authors, the core deficit in ADHD is at the level of the behavioral inhibition system that, in turn, affects executive functions (e.g., Barkley, 1997). Two skills that are subserved by the EF system were assessed, viz. the ability to select task-relevant information, and the ability to compare items stored in working memory with external input.

Lack of selective attention and working memory skills is often seen in children with ADHD (e.g., Barkley, 1997; Swaab-Barneveld, 1998), and it is important to determine the genetic and environmental contributions to these skills in young children. If it can be established that performance on tasks of selective attention and working memory is heritable, then these tasks may be used as a basis to further refine the ADHD cognitive endophenotypes. Furthermore, an insight into the genetics of these functions may ultimately help to unravel the genetic pathways of ADHD. Performance on these tasks may then be used as an early genetic predictor for ADHD, since ADHD cannot be diagnosed yet at this early age. We used a selective attention task that has already been shown to differentiate children with ADHD from controls by Swaab-Barneveld (1998) who reported more errors in children with ADHD than controls. The memory search task is a variant of the Sternberg task, and is thought to index working memory efficiency that, in turn, is affected in ADHD (e.g., Barkley, 1997).

The contributions of additive genetic factors, shared environmental factors and unique environmental factors in explaining the variance observed for these measures were estimated using structural equation modeling. In addition, sex differences in these contributions were explored.

METHOD

Subjects

The sample consisted of 237 twin pairs with a mean age of 5.8 years (SD. 0.1, range 5.67–5.92). All subjects were registered at birth with the Netherlands Twin Registry (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. Of all multiple births in the Netherlands, 40–50% are registered by the NTR (Boomsma, 1998; Boomsma *et al.*, 2002). There were 52 monozygotic

male twin pairs (MZM), 37 dizygotic male twin pairs (DZM), 73 monozygotic female twins pairs (MZF), 36 dizygotic female twin pairs (DZF) and 39 dizygotic opposite-sex twin pairs (DOS) in the sample. In the same sex twin pairs, zygosity was determined on the basis of DNA polymorphisms.

Assessment

In this study the Amsterdam Neuropsychological Tasks (ANT) (de Sonneville, 1999) were used. The ANT consists of a series of tasks, designed especially for the evaluation of attentional control in children as young as five, including sustained, selective and divided attention and visuo-motor coordination paradigms. All children were visited at home, trained testers administered the test on a laptop. Before each test the children received a practice session to ensure that the tasks were well understood and practiced. Prior to assessment, it was determined whether the child was a left-hander or a right-hander. Hand preference was determined by self-report of the subject. When in doubt, the child was asked to draw a doll, and the hand being used for this activity was considered the preferred hand.

The children were tested one by one in random order with regard to birth order. In this paper we report on data of the following subtests of the ANT:

Selective attention task

In this task a fruit basket is presented with four pieces of fruit. Two pieces of fruit are aligned in a vertical fashion (top and bottom) and two pieces in a horizontal fashion (left and right). An example of the stimulus is shown in Figure 1. Subjects have to give a yes-response if the target fruit is shown at one of the two relevant locations, i.e. the top or bottom location of the vertical axis. They have to give a no-response if the target fruit is shown but at an irrelevant location (left or right; horizontal axis), or if the target fruit is absent altogether. The display with the target fruit on
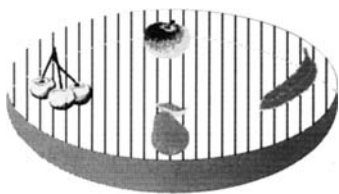
the vertical axis is the target signal; the display with the target fruit on the horizontal axis is the foil signal, and the display that contains only the four non-target fruits is the non-target signal. It is expected that the presence of a foil signal will lead to slower RTs (due to suppression of the incorrect response), and/or to more false alarms (due to a failure to suppress the incorrect response). If the foil manipulation is effective then we can proceed to assessing the heritability of the foil effect.

The three signal types were presented in a random order (28 target signals, 14 foil signals, 14 non-target signals). Yes and no responses were given by pressing one of two buttons (left or right) of a mouse that was attached to the laptop. Which of the two buttons served as the yes and no buttons was a function of the hand preference of the subject, and was determined separately for each subject. For right-handers, the right mouse button served as the yes button and was operated by the index finger of the right hand. The left button served as the no button, and was operated by the index finger of the left hand. This pairing was reversed for left-handers. Following a key-press, the next signal is presented 1200 ms later, preceded the last 500 ms by a warning signal (fixation cross). Prior to the experiment, the children were given verbal instructions, and were shown the stimulus material. Next, they received twelve practice trials that were not analyzed. When in doubt, the experimenter could administer an additional set of practice trials. Subsequently, the Selective Attention experiment began.

Memory search task

In this task children are presented with an image of a house with four animals presented simultaneously in the windows and the door opening. An example of a

Fig. 1. An example of a stimulus used in the selective attention task.

Fig. 2. An example of a stimulus shown in the memory search task.

stimulus is shown in Figure 2. Subjects are instructed to press the yes-key when the signal contains an animal from the memory set, and to press a no-key when this is not the case. On each trial the animals occupy different positions. The task consists of two parts. In part I the memory set contains one animal (e.g., a mouse; Memory Load = 1), in part II the memory set contains two animals (e.g., a bird and a cat; Memory Load 2).

Per task 20 target and 20 non-target signals were presented in random order. Reaction times and errors are expected to be higher when the memory load equals 2 (i.e., 2 items are stored in memory) than when the memory load equals 1. The reaction time increase can be used as an index of the rate of memory search. If the memory load manipulation is effective then we can proceed to assessing the heritability of the load effect.

The children pressed one of two mouse keys, as described above, to signal a yes and no response. After a key press, the next stimulus was presented after 1200 ms, preceded the last 500 ms by a warning signal (small fixation square). Prior to the experiment, the children were given verbal instructions, and were shown the stimulus material. Next, they received twelve practice trials that were not analyzed. When in doubt, the experimenter could administer an additional set of practice trials. Subsequently, the Memory Search experiment began.

SCORES

Selective attention task

The reaction times for hits (RT hits), correct rejections on foils (RT CR [foil]) and correct rejections on non-targets (RT CR [non-target]) are sampled separately. The following error scores were derived: percentage of misses (P-misses), the percentage of false alarms (P-FA) which can be divided further in percentage of false alarms on foils (P-FA [foil]) and the percentage of false alarms on non-targets (P-FA [non-target]). Of special interest is the difference between RT CR [foil] and RT CR [non-target], which can be used as an index of the size of the foil effect on RT, and the difference between P-FA [foil] and P-FA [non-target], which can be used as an index of the size of the foil effect on accuracy. Also, summary variables were computed: the percentage of errors (% errors = $(2 \times$ P-misses + P-FA [foil] + P-FA [non-target])/4), giving an indication of the accuracy of task performance, and the mean reaction time, which gives an indication of overall processing speed.

Memory search task

The reaction time for hits (RT hits) and for correct rejections (RT CR) are sampled separately under load 1 (i.e., one item in the memory set) and under load 2 (i.e., two items in the memory set). The following error scores were computed: percentage of misses (P-misses) and percentage of false alarms on non-target signals (P-FA) per load condition. In addition, the total error% was computed as (P-misses$_{load1}$ + P-FA$_{load1}$ + P-misses$_{load2}$ + P-FA$_{load2}$)/4.

The increase in search time when the number of items to be stored in memory increases from 1 to 2 was calculated as ((RT hits + RT CR)$_{load2}$−(RT hits + RT CR)$_{load1}$)/2. We called the variable Load [RT], which can be used as an index of the memory search rate. In a similar vein, we computed the effect of increasing memory load on accuracy: Load [Acc] = ((P-misses + P-FA)$_{load2}$ − (P-misses + P,FA)$_{load1}$)/2.

Data of children who had an extremely high error rate (> 40% false alarms or misses) were excluded from the study. In addition, we excluded data of children whose overall RT was more than 3 standard deviations above the sample mean. These criteria resulted in 14 individuals being excluded from the selective attention task (11 of which due to their high error rates), and 8 individuals from the memory search task (4 of which due to their high error rates). Due to these criteria the number of complete twin pairs dropped from 237 to 225 in the selective attention task, and from 237 to 231 in the memory search task.

TEST OF MEANS

In the selective attention task, we performed *t*-tests to investigate whether the foil manipulation was effective, and whether there were differences in RTs and accuracy between boys and girls. For the memory search task, we performed ANOVA's to test the effects of memory load, response type, and sex on the RTs and the accuracy variable. We adopted a significance level of 0.05 throughout. All tests were done separately for the first born and the second born twins. We also calculated the effect sizes; partial eta squared(η_p^2) for ANOVA's, and Cohen's *d* for *t*-tests.

GENETIC ANALYSIS

Data from monozygotic and dizygotic twins were used to decompose the variance in performance on the different tasks of the ANT into a contribution of the additive effects of one or more genes (A),

environmental influences that are shared by twins (C) and environmental influences that are not shared by twins (E). If monozygotic twins, who share all their genetic material, are more alike than dizygotic twins, who share on average half of their genetic material, genetic effects are indicated.

Pearson correlations were calculated for the different attention measures between first born and second born twins separately for all sex by zygosity groups. An indication of the heritability can be derived by doubling the difference between correlations for MZ twins and those for DZ twins $[h^2 = 2(r_{MZ} - r_{Dz})]$ (Falconer and Mackay, 1996).

Influences of genetic and common environmental factors were tested using Structural Equation Modelling in the program Mx. Raw data were analyzed to make it possible to include data from incomplete twin pairs. For each genetic model, the relative importance of genetic, common environmental and unique environmental influences were expressed in h^2, c^2 and e^2 respectively. The effect of sex on these estimates was assessed by likelihood-ratio tests (3 df) by comparing the fit of an ACE model with parameter estimates constrained to be equal across sexes (ACEnosex) to one in which they were allowed to vary (ACEsex). Genetic models were compared to saturated models, to obtain an indication of goodness of fit. The chi-squared statistic is computed as twice the difference between the likeli-hood for the full model (LL_0) and that for a reduced model (LL_1) ($\chi^2 = -2(LL_0 - LL_1)$). Subsequently, the same approach was used to compare the ACE model without sex differences with the AE and CE models without sex differences ($\Delta df = 1$). These were then compared to the E model without sex differences ($\Delta df = 1$).

Prior to the genetic tests we examined—using the Levene test—whether the variances of the RTs in both test were of equal magnitude across the sex and zygosity groups.

RESULTS

Descriptive statistics

In Table I the number of subjects, mean and standard deviation are shown for girls and boys per variable for the first-born twin (twin 1) and the second born twin (twin2) for the selective attention task.

For the selective attention task, a t-test revealed that girls were somewhat faster than boys (oldest twin only), and that girls were more accurate than boys (youngest twin only), although the effect sizes were modest. Our main interest is with the effects of the foil signal on performance. As can be seen from Table I, there was little evidence of an effect of the foil on speed of responding, but there was substantial evidence for an effect on accuracy. Both boys and girls, and both

Table I. Number of subjects (*N*), mean and standard deviation (SD) for the variables of the selective attention task for girls and boys, separately for the first born twin (Twin1) and the second born twin (Twin2)

Variable	Sex	*N* (twin1)	*N* (twin2)	Mean twin1	Mean twin2	SD twin1	SD twin2
Mean RT	Girl	118	133	1843* ($d=0.27$)	1870	370	389
	Boy	112	97	1955	1937	444	390
Mean Accuracy	Girl	118	133	3.41	2.98* ($d=0.32$)	3.91	3.28
	Boy	112	97	3.19	4.05	3.37	3.45
Foil [RT]	Girl	118	113	41	14	366	288
	Boy	112	97	67[#] ($d=0.31$)	−13	306	324
Foil [Ace]	Girl	118	133	2.18[#] ($d=0.34$)	2.95[#] ($d=0.62$)	9.17	6.76
	Boy	112	97	2.61[#] ($d=0.49$)	3.53[#] ($d=0.72$)	7.61	6.93

Effect sizes of significant effects (Cohen's *d*) are shown in parentheses.
RT = Overall Reaction Time, Accuracy = percentage of errors, Foil [RT] = the size of the foil effect for the RT variable, Foil [Acc] = the size of the foil effect for the accuracy variable (all RT-variables in ms).
Note: a foil is a target signal at an irrelevant location, thus requiring a no-response.
*t-test showed a significant difference between boys and girls, $p < 0.05$.
[#]t-test showed a significant deviation from 0.
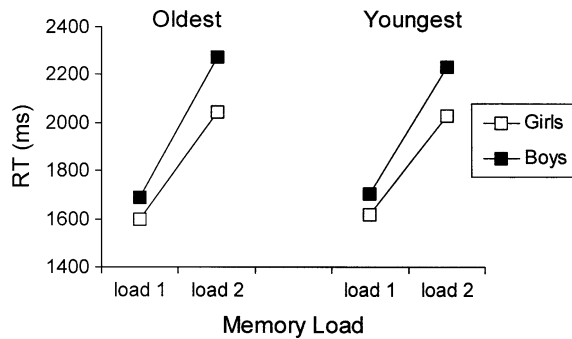
Fig. 3. Mean RTs (in ms) for the Memory Search task, as a function of memory load (1 vs. 2), response type (hits vs. correct rejections) and sex (boys vs. girls).

first and second born twins, made more errors (false alarms) to the foil signal than to the non-foil signal. It thus seems to be the case that the foil manipulation was effective, but only for accuracy. Therefore, with regard to the foil effect, we will investigate the heritability of accuracy only, not RT.

The mean RTs for the memory search task are shown in Figure 3 for the oldest twins (left panel) and the youngest twins (right panel). A 2×2 ANOVA was performed on the RTs with memory load (Load 1 vs. Load 2) as within-subjects factor, and sex (boys/girls) as between-subjects factor. This ANOVA was performed separately for the youngest and oldest twins. For both groups we found near-identical results. First, both groups showed an effect of memory load (youngest: $F(1, 231) = 582.2$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.716$; oldest: $F(1, 231) = 603.8$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.723$). It took less time to identify the target when one item was stored in memory (mean RT = 1652 ms) than when 2 items were in memory (2146 ms). Thus Load [RT] (the search rate) had a value of 494 ms. Second, both groups showed an effect of sex (youngest: $F(1, 231) = 12.8$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.052$); oldest: $F(1, 231) = 14.9$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.061$), indicating that girls were on average faster than boys. Finally, in both groups sex interacted with memory load (youngest: $F(1, 231) = 9.8$, $p < 0.01$, $\eta_{\mathrm{p}}^2 = 0.041$); oldest: $F(1, 231) = 10.2$, $p < 0.01$, $\eta_{\mathrm{p}}^2 = 0.042$). As can be seen from the figure, the increase in RTs with increasing load was greater for boys than for girls.

A 2×2 ANOVA was performed on the error percentages with memory load (Load 1 vs. Load 2) as within-subjects factor, and sex (boys/girls) as between-subjects factor. This ANOVA was performed separately for the youngest and oldest twins. Again, for both groups we found near-identical results. Both

groups showed an effect of memory load (youngest: $F(1, 231) = 9.1$, $p < 0.01$, $\eta_{\mathrm{p}}^2 = 0.038$; oldest: $F(1, 231) = 6.1$, $p < 0.05$, $\eta_{\mathrm{p}}^2 = 0.026$). There was an overall increase in errors with increasing memory load; averaged over both groups, the error rate increased from 3.5% (Load 1) to 4.4% (Load 2), yielding a value of load [Acc] of 0.9%. No effects involving sex were significant.

Genetic analysis

In Tables II and III the correlations for the five zygosity groups for the selective attention task and the memory search task, respectively, are shown. Inspection of Table II (selective attention) reveals that the correlations for overall RT were all moderately high, with the exception of DZF twins. This pattern of correlation indicates possible influences of genes and common environment. For accuracy twin correlations tend to be higher for MZ pairs than for DZ pairs, suggesting that genetic factors are important. However, no such pattern was observed for the size of the foil effect on accuracy. This suggests that unique environmental factors account for the larger part of the observed individual differences.

Inspection of Table III (memory search) showed a pattern of twin correlations for overall RT and memory search rate (Load [RT]) that was indicative of genetic factors on these variables. In addition, the accuracy appeared to be influenced by common environmental and genetic factors, but hardly any evidence for familial influences was found on the Load [Acc] variable.

We conducted a Levene test across sex (male and female) and zygosity group (monozygotic and dizygotic) on all variables listed in Tables II and III. We found that the variances of the mean RT in the

Table II. Twin correlations for the variables of selective attention task for all zygosity groups

Zygosity (N)	RT	Accuracy	foil [Acc]
MZM (48)	0.54	0.42	0.04
DZM (34)	0.52	0.25	0.22
MZF (69)	0.44	0.32	0.04
DZF (36)	0.21	−0.24	0.12
DOS (38)	0.51	−0.13	0.11

Number of subjects (N) in parentheses.
RT = overall reaction time, foil [Acc] = the size of the foil effect on accuracy, MZM = monozygotic male, DZM = dizygotic male, MZF = monozygotic female, DZF = dizygotic female, DOS = opposite sex twins.
Note: a foil is a target signal at an irrelevant location, thus requiring a no-response.

Table III. Twin correlations for the variables of the Sternberg memory search task for all zygosity groups

Zygosity (N)	RT	Accuracy	Load [RT]	Load [Acc]
MZM (50)	0.64	0.54	0.40	0.24
DZM (35)	0.02	0.44	0.15	0.07
MZF (71)	0.52	0.26	0.33	−0.05
DZF (36)	0.20	0.11	−0.15	0.18
DOS (39)	0.39	0.15	0.01	0.23

Number of subjects (N) in parentheses.
RT = Overall Reaction Time, Accuracy = percentage of errors, Load [RT] = the increase in search time with increasing memory load, Load [Acc] = increase in errors with increasing memory load, MZM = monozygotic male, DZM = dizygotic male, MZF = monozygotic female, DZF = dizygotic female, DOS = opposite sex twins.

memory search task was statistically unequal across the groups, both for the first born twins, $F(3, 227) = 5.355$, $p < 0.001$, and for the second born twins, $F(3, 227) = 4.536$, $p < 0.01$, which seemed to be caused by a somewhat larger variance in the monozygotic males than in the other groups. For the remaining variables the Levene test yielded no significant effects.

Univariate genetic models were fitted to the selective attention (RT, accuracy and foil) and memory search variables (RT, accuracy, load-RT and Load-accuracy). For none of these variables, sex differences in variance components were significant ($\Delta\chi^2(3)$ ranged from 1.061 to 6.749). However, the highest χ^2-value was obtained with the memory search task, and was marginally significant ($p = 0.08$). This is consistent with the results of the Levene test, which showed significant differences in variance across the groups for that particular variable.

Table IV shows model fitting outcomes and parameter estimates. For MS-RT and MS- Load-RT, common environmental influences could be discarded from the models, but additive genetic influences were significant, resulting in heritabilities of 54% (Confidence interval = 42% to 63%) for MS-RT and 29% (Confidence interval = 14% to 43%) for MS- Load-RT (difference between ACEnosex and CEnosex: $\Delta\chi^2(1) = 3.914$ and 4.300 respectively). At least part of the variance of SA-RT, SA-accuracy and MS-accuracy was familial, but no conclusions could be drawn from the analyses as to whether this source of variation was genetic or common environmental. SA-RT resulted in a model with either moderately high heritability (50%, CI = 38% to 61%) or large influences of common environment (44%, CI = 33% to 54%). SA-accuracy showed somewhat lower heritability (29%, CI = 14% to 42%) or influences of common environment (21%, CI = 8% to 33%), as did

MS-accuracy, for which the AE model showed a heritability of 35% (CI = 22% to 47%) and the CE model resulted in c^2 of 33% (CI = 21% to 44%). Variance of SA-foil and MS-Load-accuracy was completely explained by unique environmental factors.

DISCUSSION

The aim of this paper was to examine childhood performance on two neurocognitive tasks, and to examine whether individual differences in performance could be explained by genetic factors. If so, this would aid in refining the cognitive endophenotypes of behavior problems, such as ADHD. In this study, a large sample of nearly 6-year old MZ and DZ twins performed a selective attention task and a working memory task.

In the selective attention task, subjects had to respond to a target item, but only when it was present on a certain location. The crucial condition was the one in which the target was present, yet at the task-irrelevant location. It was expected that the presence of this foil signal would bias for a yes-response, which needs to be suppressed in order to produce the correct response. This would then lead to a relatively high proportion of false alarms and/or slower RTs. It was found that the foil induced more false alarms than when the target was absent, but that RT was essentially unaffected by the foil.

The working memory task consisted of a variant of the Sternberg memory search task. Subjects had to memorize one animal (Load 1) or two animals (Load 2), and to respond to the presence of an animal from the memory set. It was expected that memory Load 2 would lead to higher RTs and/or more errors than Load 1. Both these expectations were confirmed: when 2 items were stored in memory subjects responded on average nearly 500 ms slower, and with a nearly 1%-increase in errors, than when just 1 item had to be stored.

Only small differences (small effect sizes) between males and females were found, with females being somewhat faster and more accurate than males. This finding was replicated in both first-born and second-born twins.

Having established that the experimental manipulations were effective, the heritabilities of the different performance indices were assessed. The patterns of twin correlations, and the model fitting results strongly suggested familial influences of absolute performance parameters (overall reaction time and accuracy) and relative performance parameters (memory search rate and the size of the foil effect). However, it was difficult

Table IV. Model fitting and parameter estimates of heritability (h^2), common environment (c^2), and unique environment (e^2), for three selective attention variables (overall RT, accuracy and foil effect-accuracy) and for four memory search variables (overall RT, accuracy, load-RT and load-accuracy)

	Model	−2LL	df	c.t.m.	χ^2	df	p	h^2	c^2	e^2
Selective Attention: RT ($N=460$)										
1	saturated	2490.076	430							
2	ACEsex	2509.140	442	1	19.064	12	0.087	0.16/0.05	0.25/0.49	0.59/0.46
3	ACEnosex	2512.227	445	2	3.087	3	0.378	0.14	0.34	0.53
4	**AEnosex**	2515.556	446	3	3.329	1	0.068	**0.50**	–	**0.50**
5	**CEnosex**	2512.690	446	3	0.463	1	0.496	–	**0.44**	**0.56**
6	Enosex	2561.481	447	4	45.925	1	0.000	–	–	1.00
				5	48.791	1	0.000	–	–	1.00
Selective Attention: Accuracy ($N=460$)										
1	saturated	2373.244	430							
2	ACEsex	2428.681	442	1	55.437	12	0.000	0.22/0.45	0.00/0.00	0.78/0.55
3	ACEnosex	2432.872	445	2	4.191	3	0.242	0.29	0.00	0.71
4	**AEnosex**	2432.872	446	3	0.000	1	1.000	**0.29**	–	**0.71**
5	**CEnosex**	2436.558	446	3	3.686	1	0.055	–	**0.21**	**0.79**
6	Enosex	2446.301	447	4	13.429	1	0.000	–	–	1.00
				5	9.743	1	0.002	–	–	1.00
Selective Attention: Foil ($N=460$)										
1	saturated	3122.457	430							
2	ACEsex	3167.042	442	1	44.585	12	0.000	0.00/0.00	0.06/0.11	0.94/0.89
3	ACEnosex	3169.353	445	2	2.311	3	0.510	0.00	0.08	0.92
4	AEnosex	3170.030	446	3	0.677	1	0.411	0.06	–	0.94
5	CEnosex	3169.353	446	3	0.000	1	1.000	–	0.08	0.92
6	**Enosex**	3170.767	447	4	0.737	1	0.391	–	–	**1.00**
				5	1.414	1	0.234	–	–	1.00
Memory Search: Load-RT ($N=466$)										
1	saturated	2274.652	436							
2	ACEsex	2299.404	448	1	24.752	12	0.016	0.32/0.43	0.17/0.13	0.51/0.44
3	ACEnosex	2306.153	451	2	6.749	3	0.080	0.41	0.12	0.47
4	**AEnosex**	2306.456	452	3	0.303	1	0.582	**0.54**	–	**0.46**
5	CEnosex	2310.067	452	3	3.914	1	0.048	–	0.46	0.54
6	Enosex	2365.955	453	4	59.499	1	0.000	–	–	1.00
				5	55.888	1	0.000	–	–	1.00
Memory Search: Accuracy ($N=466$)										
1	saturated	2329.180	436							
2	ACEsex	2384.552	448	1	55.372	12	0.000	0.04/0.18	0.18/0.34	0.78/0.49
3	ACEnosex	2390.076	451	2	5.524	3	0.137	0.00	0.33	0.67
4	**AEnosex**	2392.604	452	3	2.528	1	0.112	**0.35**	–	**0.65**
5	**CEnosex**	2390.076	452	3	0.000	1	1.000	–	**0.33**	**0.67**
6	Enosex	2417.166	453	4	24.562	1	0.000	–	–	1.00
				5	27.090	1	0.000	–	–	1.00
Memory Search: Load-RT ($N=466$)										
1	saturated	2315.738	436							
2	ACEsex	2342.430	448	1	26.692	12	0.001	0.24/0.08	0.00/0.26	0.76/0.66
3	ACEnosex	2346.793	451	2	4.363	3	0.225	0.29	0.00	0.71
4	**AEnosex**	2346.793	452	3	0.000	1	1.000	**0.29**	–	**0.71**
5	CEnosex	2351.093	452	3	4.300	1	0.038	–	0.20	0.80
6	Enosex	2360.342	453	4	13.549	1	0.000	–	–	1.00
				5	9.249	1	0.002	–	–	1.00
Memory Search: Load-Accuracy ($N=466$)										
1	saturated	2770.843	436							
2	ACEsex	2790.175	448	1	19.332	12	0.081	0.01/0.19	0.04/0.03	0.95/0.78
3	ACEnosex	2791.236	451	2	1.061	3	0.786	0.00	0.09	0.91
4	AEnosex	2791.996	452	3	0.760	1	0.383	0.09	–	0.91
5	CEnosex	2791.236	452	3	0.000	1	1.000	–	0.09	0.91
6	**Enosex**	2793.293	453	4	1.297	1	0.255	–	–	**1.00**
				5	2.057	1	0.152	–	–	1.00

Best fitting models are shown in bold.
LL = Log Likelihood, df = degrees of freedom; c.t.m. = Compare to model.

to ascribe the individual performance variation either to common genes or to common environment, with the possible exception of overall RT, overall accuracy, and memory search rate, where the heritability reached significance. Similar findings were reported by Groot *et al.* (2004), who found strong familial influences on attentional skills. In that study it also was difficult to determine whether these influences were genetic or shared environmental in origin, and the authors explained this state of affairs by the high inter- and intra-subject variability in performance.

This study nicely complements a recent series of studies conducted by Fan and co-workers. Using a sample of healthy MZ and DZ twins, Fan *et al.*, (2001) asked whether genetic variation contributed to variations in performance on basic attentional tasks. These tasks were designed to tap distinct attentional brain networks (see also Fan *et al.*, 2002), and damage to each of these networks is associated with distinct neuropsychological deficits (Fernandez-Duque and Posner, 2001). One of the findings of Fan *et al.*, (2001) was that performance on the flanker task, which was supposed to index the efficiency of the dopamine rich frontal executive network, was highly heritable. Damage to this network, in turn, is implicated in ADHD, among others. In a follow-up paper, Fossella *et al.*, (2002) found modest associations between genetic polymorphisms of several genes implicated in frontal (dys)function, such as DRD4 and DAT1, and the efficiency of the frontal executive attention network. These genes, in turn, are implicated in ADHD (e.g., Swanson *et al.*, 2001). Thus, the identification of a heritable component in basic cognitive tasks (response interference or working memory) may guide the search for genes that are involved in normal and abnormal processing in that domain.

This brings us to our main research question; whether neurocognitive performance, as assessed by the different measures from our tasks, may be used to refine the cognitive endophenotypes. De Geus and Boomsma (2001) listed 5 criteria that are ideally possessed by endophenotypes (see Gottesman and Gould, 2003, for a slightly different approach to endophenotypes). Their criteria 1 and 2 state that the endophenotype should be reliable and heritable. With respect to heritability, we have demonstrated that—at least for this age group—there is strong evidence for familial influence, although it is impossible to tell whether this influence represents common genes or common environment. An exception was our measure of memory search rate, which was found to be

heritable. This is noteworthy, because this variable is composed of the RT-difference between Load 1 and Load 2. From test theory we know that, in general, difference scores between two measures (such as interference of difference scores in RT) are characterized by low reliabilities, and hence heritabilities (see also Stins *et al.*, in press).

A remarkable finding is that heritability of attention problems tends to be higher than the heritability of indices that can be considered to be more direct measures of attention, such as accuracy and distractability (the influence of a foil). As argued by Groot *et al.*, (2004), it might be the case that the insight of a parent or teacher into a child's behavior is collected over a longer period of time, as opposed to a neuropsychological test battery. In addition, studies of ADHD have seldom found an influence of shared environment (for an exception, see Rhee *et al.*, 1999).

With respect to reliability, we have no test–retest data, but we can assess reliability by examining the MZ correlations (Falconer and Mackay, 1996). It has been argued that an MZ correlation provides a lower limit to reliability. Inspection of Tables 2 and 3 reveals that many performance measures are characterized by high MZ correlations, which implies high reliabilities.

Criteria 3 to 5 of de Geus and Boomsma (2001) deal with the relationship between phenotype and endophenotype. According to these authors, there should be a high correlation between the phenotype and endophenotype, and this correlation should be based in genetics and—moreover—be theoretically meaningful. Although we have in our Introduction briefly touched upon the relationship with high-level phenotypes, such as psychopathologies, the challenge for researchers is now to combine this knowledge with what is known of the neurobiology of psychopathologies of interest.

REFERENCES

Acosta, M. T., Arcos-Burgos, M., and Muenke, M. (2004). Attention deficit/hyperactivity disorder (ADHD): Complex phenotype, simple genotype?. *Genet. Med.* **6**:1–15.

Bakker, S. C., van der Meulen, E. M., Buitelaar, J. K., Sandkuijl, L. A., Pauls, D. L., Monsuur, A. J., van't Slot, R., Minderaa, R. B., Gunning, W. B., Pearson, P. L., and Sinke, R. J. (2003). A whole-genome scan in 164 Dutch sib pairs with attention-deficit/hyperactivity disorder: suggestive evidence for linkage on chromosomes 7p and 15q. *Am. J. Human Genet.* **72**:1251–1260.

Berger, A. and Posner, M. I. (2000). Pathologies of brain attentional networks. *Neurosci. Biobehav. Rev.* **24**(1):3–5. (review).

van Beijsterveldt, C. E. M., Verhulst, F. C., Molenaar, P. C. M., and Boomsma, D. I. (2004). The genetic basis of problem behavior in 5-year-old Dutch twin pairs. *Behav. Genet.* **34**:229–242.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychol. Bull.* **121**:65–94.

Bartels, M., Rietveld, M. J. H., van Baal, G. C. M., and Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behav. Genet.* **32**:237–249.

Boomsma, D. I. (1998). Twin registers in Europe: an overview. *Twin Res.* **1**:34–51.

Boomsma, D. I., Vink, J. M., van Beijsterveldt, T. C. E. M., de Geus, E. J. C., Beem, A. L., Mulder, E. J. C. M., Derks, E. M., Riese, H., Willemsen, G. A. H. M., Bartels, M., van den Berg, M., Kupper, N. H. M., Polderman, T. J. C., Posthuma, D., Rietveld, M. J. H., Stubbe, J. H., Knol, L. I., Stroet, T., and van Baal, G. C. M., (2002). Netherlands Twin Register: A Focus on Longitudinal Research. *Twin Research* **5**: 401–406.

Bush, G., Frazier, J. A., Rauch, S. L., Seidman, L. J., Whalen, P. J., Jenike, M. A., Rosen, B. R., and Biederman, J. (1999). Anterior Cingulate cortex dysfunction in Attention Deficit/Hyperactivity Disorder revealed by fMRI and the counting Stroop. *Biol. Psychiatr.* **45**:1542–1552.

Castellanos, F. X., and Tannock, R. (2002). Neuroscience of Attention Deficit/Hyperactivity Disorder: the search for endophenotypes. *Nat. Rev. Neurosci.* **3**:617–628.

Falconer, D. S. and Mackay, T. F. C. (1996) Introduction to quantitative genetics. Longan Group LtD., 4th edn.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cognitive Neurosci.* **14**:340–347.

Fan, J., Wu., Y., Fossella, J. A., and Posner, M. I. (2001). Assessing the heritability of attentional networks. *BMC Neurosci.* **2**:14.

Fernandez-Duque, D., and Posner, M. I. (2001). Brain imaging of attentional networks in normal and pathological stated. *J. Clin. Exp. Neuropsychol.* **23**:74–93.

Fisher, S. E., Franks, C., McCracken, J. T., McGough, J. J., Marlow, A. J., MacPhie, I. L., Newbury, D. F., Crawford, L. R., Palmer, C. G. S., Woodward, J. A., Del' Homme, M., Cantwell, D. P., Nelson, S. F., Monac, A. P., and Smalley, S. L. (2002). A genomewide scan for loci involved in attention-deficit/hyperactivity disorder. *Am. J. Human Genet.* **70**:1183–1196.

Fossella, J., Sommer, T., Fan, J., Wu, Y., Swanson, J. M., Pfaff, D. W., and Posner, M. I. (2002). Assessing the molecular genetics of attention networks. *BMC Neurosci.* **3**:14 .

de Geus, E. J. C., and Boomsma, D. I. (2001). A genetic neuroscience approach to human cognition. *Eur. Psychol.* **6**:241–253.

Gottesman, I. I., and Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatr.* **160**:1–10.

Groot, A. S., de Sonneville, L. M. J., Stins, J. F., and Boomsma, D. I. (2004). Familial influences on sustained attention and inhibition in preschoolers. *J. Child Psychol. Psychiatr.* **45**:306–314.

Rhee, S. H., Waldman, I. D., Hay, D. A., and Levy, F. (1999). Sex differences in genetic and environmental influences on DSM-III-R attention-deficit/hyperactivity disorder. *J. Abnormal Psychol.* **108**:24–41.

Rietveld, M. J. H., Hudziak, J. J., Bartels, M., van Beijsterveldt, C. E. M., and Boomsma, D. I. (2003). Heritability of attention problems in children: I. Cross-sectional results from a study of twins, age 3–12 years. *Am. J. Med. Genet. B (Neuropsychiatr. Genet.)* **117B**:102–113.

Rietveld, M. J. H., Hudziak, J. J., Bartels, M., van Beijsterveldt, C. E. M., and Boomsma, D. I. (2004). Heritability of attention problems in children: longitudinal results from a study of twins, age 3–12 years. *J. Child Psychol. Psychiatr.* **45**:577–588.

de Sonneville, L. M. J. (1999). Amsterdam Neuropsychological Tasks: a computer-aided assessment program. In B. P. L. M. Den Brinker, P. J. Beek, A. N. Brand, F. J. Maarse and L. J. M. Mulder (eds.), *Cognitive Ergonomics, Clinical Assessment and Computer-assisted Learning: Computers in Psychology*. Lisse, the Netherlands: Swets & Zeitlinger, pp. 187–203.

Stins, J. F., Polderman, J. C., Boomsma, D. I. and de Geus, E. C. J. (2005). Response interference and working memory in 12-year old children. *Child Neuropsychology* (in press).

Slaats-Willemse, D. S., Swaab-Barneveld, H., de Sonneville, L., van der Meulen, E., and Buitelaar, J. (2003). Deficient response inhibition as a cognitive endophenotype of ADHD. *J. Am. Acad. Child Psychiatr.* **42**:1242–1248.

Swaab-Barneveld, J. T. (1998). Information processing in a child psychiatric population. PhD-thesis, Utrecht.

Swanson, J., Deutsch, C., Cantwell, D., Posner, M., Kennedy, J. L., Barr, C. L., Moyzis, R., Scbuck, S., Flodman, P., Spence, M. A., and Wasdell, M. (2001). Genes and attention- deficit hyperactivity disorder. *Clin. Neurosci. Res.* **1**:207–216.

Thapar, A., Holmes, J., Poulton, K., and Harrington, R. (1999). Genetic basis of attention deficit and hyperactivity. *Br. J. Psychiatr.* **174**:105–111.

Edited by Irwin Waldwan