

Psychological Mechanisms in Hyperactivity: I Response Inhibition Deficit, Working Memory Impairment, Delay Aversion, or Something Else?

Jonna Kuntsi

Institute of Child Health, University College London Medical School, U.K.

Jaap Oosterlaan

Free University, Amsterdam, The Netherlands

Jim Stevenson

University of Southampton, U.K.

This study tested the predictions of three different theories of hyperactivity: response inhibition deficit, working memory impairment, and delay aversion. A sample of 51 pervasively hyperactive children and 119 control children, identified by screening a general population sample of 1316 twin pairs, were assessed on tests relating to each of these theories. The hyperactive group performed worse than the control group on the delay aversion measure and some of the working memory tasks. Controlling for IQ removed the significant group differences on the working memory measures, however. There were no significant group differences on the inhibition variables derived from the stop task. However, there was evidence of a pattern of responding on the stop task that was strongly characteristic of hyperactivity: hyperactive children were variable in their speed, generally slow and inaccurate in responding. This pattern of responses may indicate a nonoptimal effort/activation state. Hyperactive girls were indistinguishable from hyperactive boys in their performance on the tasks.

Keywords: ADD/ADHD, delay aversion, hyperactivity, response inhibition, state regulation, working memory.

Abbreviations: ADD: attention deficit disorder; ADHD: attention deficit hyperactivity disorder; CPRS-48: Conners' Parent Rating Scale; CTRS-28: Conners' Teacher Rating Scale; DRA: Delayed Response Alternation Task; MRT: mean reaction time; SD of RTs: standard deviation of reaction times; SSRT: stop signal reaction time.

Introduction

The psychological mechanisms that underlie hyperactivity have proved very difficult to identify. Numerous studies have reported poorer performance among hyperactive children on various tasks (for reviews, see Kuntsi & Stevenson, 2000; Sergeant, Oosterlaan & Van der Meere, 1999; Van der Meere, 1996); the debate centres on the interpretation of these group differences. Alternative interpretations include a response inhibition deficit, a working memory impairment, and "delay aversion".

The response inhibition hypothesis (e.g. Barkley, 1997; Quay, 1988, 1997) is a widely held theory of hyperactivity. The most systematic research testing the hypothesis of an impairment in the ability to withhold a prepotent response has employed the stop task. This task has advantages over other measures of inhibition, in that it is based on an explicit model of the inhibitory process (the

race model; Logan, 1994; Logan & Cowan, 1984) and it allows a distinction between inhibitory control and the processes involved in the execution of the primary task. The primary task in the stop signal paradigm is a two-choice reaction-time task. On the inhibition trials an auditory stop signal is presented, which indicates that the child should withhold responding.

A recent meta-analysis of stop task studies on hyperactivity (Oosterlaan, Logan, & Sergeant, 1998) concluded that children with ADHD (attention deficit hyperactivity disorder) have a slower inhibitory process, but they are not less likely to trigger the inhibitory process nor is their inhibitory process more variable than that of other children. On the primary task children with ADHD are more variable (Oosterlaan, unpublished data from the meta-analysis) and slower in their speed, and they also make more errors. Rather than indicating a specific response inhibition deficit, the overall pattern of the findings may therefore suggest a *generally slow mode of information processing* (see also Sergeant et al., 1999; Tannock, 1998). The finding of a high variability in speed of responding may indicate lack of consistent effort (Oosterlaan & Sergeant, 1996).

Requests for reprints to: Jonna Kuntsi PhD, Behavioural Sciences Unit, Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, U.K. (E-mail: j.kuntsi@ich.ucl.ac.uk).

Another approach to investigating the presumed cognitive deficits in hyperactivity focuses on the findings that hyperactive children in general perform poorly on a variety of executive function measures (Pennington & Ozonoff, 1996). Pennington and colleagues (Pennington, Bennetto, McAleer, & Roberts, 1996) argue that the ability the so-called executive function measures share is that of working memory. In this model, inhibition is seen as intrinsic to the operation of the working memory.

Only a few studies on hyperactivity have employed working memory measures. Two studies of those included in a review of executive function studies by Pennington and Ozonoff (1996) reported group differences between hyperactive/ADHD and control groups on a working memory measure: Gorenstein, Mammato, and Sandy (1989) using the sequential matching memory task and Shue and Douglas (1992) using the self-ordered pointing task. Siegel and Ryan (1989) reported, however, no differences on most comparisons between children with and without attention deficit disorder (ADD) on two working memory measures, the sentence and counting span tasks. Children with ADD performed worse than the control children at the younger age group (7–8-year-olds) and only on the sentence span task. The 9–13-year-olds showed no significant differences.

Other research challenges the notion that the “core” deficit is one of response inhibition or a more general executive function impairment. By manipulating variables that influence task performance, Sonuga-Barke and colleagues (e.g. Sonuga-Barke, Houlberg, & Hall, 1994; Sonuga-Barke et al., 1992) have obtained evidence in favour of an alternative hypothesis: hyperactive children could be *delay averse*.

On a task in which the children had to make a choice between a small immediate reward and a large delayed reward, hyperactive children chose the small immediate reward more often than control children, earning fewer points, but only when this strategy reduced the overall delay period (Sonuga-Barke et al., 1992). When choosing the small immediate reward led to a post-reward delay, such that the overall delay was the same as when choosing the large delayed reward, the hyperactive children waited as well as the control children for the larger reward. Hyperactive children were not impulsive in the sense that they were *unable to wait*, instead they preferred not to wait. The argument therefore is that the cognitive “deficits” hyperactive children are thought to show could in fact be more motivational in nature: they could be due to an aversion to delay. Sonuga-Barke (1994) suggests that in situations where it is not possible to reduce the delay period, hyperactive children may be attempting to reduce the *subjective* experience of delay (by spending time off-task) or to increase the level of stimulation (by fidgeting and moving around).

The aim of the present study was to test the predictions of these three theories of hyperactivity—response inhibition deficit, working memory impairment, and delay aversion—within a single study. We wanted to contrast the relative efficacy of tasks related to each of these theories in differentiating children with hyperactivity from control children.

A specific aim with regard to the response inhibition hypothesis was to confirm that the stop task results from previous studies are not restricted to clinic-based samples. Clinic-referred samples are unlikely to be representative of the total population of hyperactive children. The only study of those included in the meta-analysis of stop task

studies (Oosterlaan et al., 1998) that did not find evidence of a slower inhibitory process among hyperactive children was the only study that obtained their sample from the general population (Daugherty, Quay, & Ramos, 1993). The evidence regarding hyperactive children’s performance on working memory measures is limited and inconsistent overall, and therefore we included several different working memory measures in the test battery to enable a detailed investigation of this hypothesis.

The present study also aimed to investigate the effects of sex, IQ, conduct problems, and anxiety on the children’s performance on the tasks. Hyperactivity often co-occurs with other behavioural difficulties such as conduct problems, learning difficulties, and anxiety (Angold, Costello, & Erkanli, 1999; Hinshaw, 1992). Childhood hyperactivity predicts conduct problems in adolescence, whereas the opposite is not true, suggesting that conduct disorder among hyperactive children is a complication of their hyperactivity (Taylor, Chadwick, Heptinstall, & Danckaerts, 1996). This raises the question of whether the poor performance of hyperactive children on some tasks is related to the “core” hyperactivity or the co-occurring conduct problems. Little is known about possible sex differences in hyperactive children’s performance on psychological tasks, as the studies focusing on clinic-referred samples have tended to include boys only. The screening of the present sample was specifically designed to ensure the inclusion of hyperactive girls.

Whereas the present paper reports results from our study as they relate to the theories of hyperactivity, our companion paper (Kuntsi & Stevenson, 2001) presents results on the role of genetic factors in influencing aspects of hyperactivity. Twin study methodology enables an investigation of genetic factors and psychological mechanisms in a single study. The sample therefore consisted of twins.

Method

Instruments to Identify Hyperactivity

Conners’ Parent (CPRS-48) and Teacher (CTRS-28) Rating Scales (Goyette, Conners, & Ulrich, 1978). The Conners’ scales are widely used to obtain ratings from parents and teachers on a range of problem behaviours in children, and on hyperactivity in particular. The parent scale provides the following dimensions: conduct problem, learning problem, psychosomatic, impulsive-hyperactive, anxiety, and hyperactivity index. The dimensions obtained from the teacher scale are the following: conduct problem, hyperactivity, inattentive-passive, and hyperactivity index. As the hyperactivity index is based on items from the other dimensions and is therefore not an additional independent measure, we did not use this subscale in the present study.

Participants

Recruitment of sample. The twins were recruited from a general population sample of same-sex twins aged between 7 and 11 years. It is a common procedure in twin studies to exclude opposite-sex twins. Figure 1 illustrates the various stages in the recruitment of the sample. We approached all primary schools ($N = 2439$), including special schools, in 16 Local Education Authorities in southern England. We asked the class teachers of any twins fulfilling the criteria for our study (same-sex twins; date of birth between 1 September 1985 and 1 September 1990) to complete the CTRS-28 (Goyette et al., 1978), separately for each twin.

The next stage of the screening process involved choosing the twin pairs in which *at least one twin* scored above the cutoff

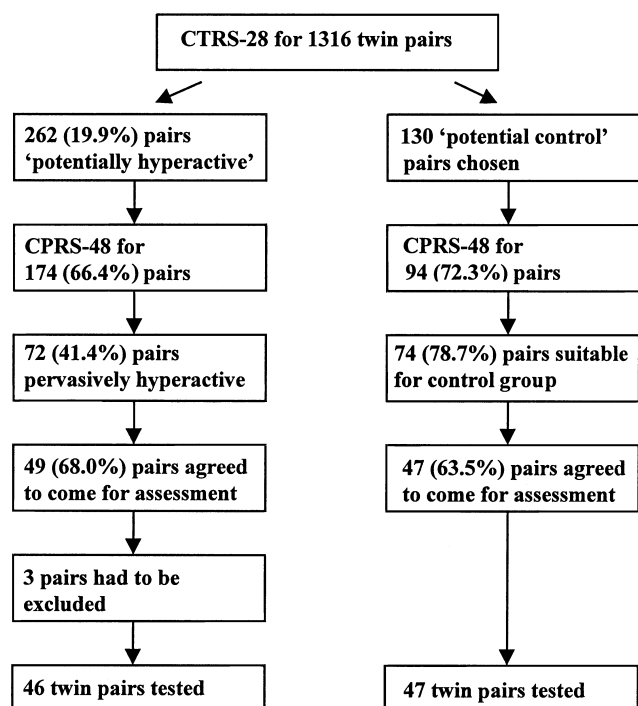


Figure 1. Recruitment of the sample.

point (T-score of 64 or higher) on the hyperactivity dimension of the CTRS-28. This cutoff was selected so that we would be likely to achieve our aim of including the 5% of *pervasively* hyperactive children in the study. In deciding on the cutoff point we took into account the likely rate of agreement between parents and teachers and the estimated percentage of cases who would drop out at any stage during the screening process. We were therefore not aiming to include the highest scoring 5% in the "potentially hyperactive" group at this stage, as this would have resulted in including only a very small percentage of the sample after the second stage of the screening. We also chose children who were candidates for the control group, stratifying them (i.e. not matching one by one) with the hyperactive twin pairs on age, sex, and zygosity. To be chosen as controls, both twins had to score below the cutoff point on the hyperactivity dimension.

We then asked one of the parents, or both parents together, to complete the CPRS-48 (Goyette et al., 1978) for each twin. Only if the *same* twin scored at or above the cutoff point (T-score = 64) on the hyperactivity dimension of the CTRS-28 and on the impulsive-hyperactive dimension of the CPRS-48, was the twin pair invited for an assessment session. Thus for a "potentially hyperactive" pair to be included in the last stage of the study, at least one of the twins had to fulfil our criteria for pervasive hyperactivity. Similarly, for a control pair to be included in the last stage of the study, both twins had to score below the cutoff point on the hyperactivity dimension on both the CTRS-28 and the CPRS-48.

We excluded twin pairs if one or both of them had serious disabilities or medical conditions or if they were on stimulant medication. In total we excluded 11 twin pairs: 2 with autism, 4 with physical disabilities (e.g. cerebral palsy), 2 with medical syndromes, 2 with severe learning difficulties, and 1 twin pair in which a twin was on stimulant medication (Ritalin).

Response rates. We received replies from 66.8% (1629) of the schools we contacted. Of those schools that replied to our letter, only 59 indicated that they did not wish to take part in the study. In 858 of the schools there were no twins fulfilling the criteria for our study. We received the CTRS-28 for 1316 twin pairs. In 19.9% (262) of the twin pairs at least one twin scored above the cutoff point on the hyperactivity dimension. To get an estimate of the proportion of a population of singletons who would score above this cutoff point (to remove the twin effect of

an increased likelihood of *at least* one of the twins being a "case"), it is useful to consider this result separately for "twins A" and "twins B". That is, the twins were arbitrarily divided into twin A and twin B within each pair: 13.7% of "twins A" and 12.6% of "twins B" scored above the cutoff point on the hyperactivity dimension on the CTRS-28.

The response rates at the following stages of the recruitment of the sample are shown in Fig. 1. The parents completed the CPRS-48 on average 3 months after the teachers had completed the CTRS-28. The CPRS-48 was completed on average 4 months ($SD = 2.2$ months) prior to the assessment. The sample whom we tested on the various tasks consists of 93 twin pairs—46 pairs in which at least one twin was pervasively hyperactive and 47 control pairs.

Twin characteristics. In the nongenetic analyses reported in this paper we focus on individual children rather than twin pairs. How *severely* hyperactive the children who met our criteria for pervasive hyperactivity can be assessed as follows. There were 94 individual children who met our criteria for pervasive hyperactivity. Based on the information given in Fig. 1, we can estimate that, in addition, approximately 47 children of the potentially hyperactive children whose parents did not return the CPRS-48 would have met our criteria for pervasive hyperactivity. These 141 children represent 5.4% of the total population of 2632 children (1316 twin pairs), a figure very close to the expected value of 5%.

For the analyses reported here we excluded children who were situationally hyperactive (above the hyperactivity cutoff point only on teacher or parent questionnaire). In the present study, the criterion for inclusion in the control group was for the child to score below the cutoff point on both questionnaires; the criterion for inclusion in the hyperactive group was for the child to be pervasively hyperactive. The sample in the present paper consists of 51 pervasively hyperactive children and 119 non-hyperactive control children (except where there were missing data). The mean hyperactivity scores of the hyperactive group were above the clinical cutoff of a T-score of 70 and, as expected, significantly higher than the mean scores of the control group: CPRS-48 impulsive-hyperactive: hyperactive mean = 71.80, $SD = 6.72$, control mean = 45.74, $SD = 7.79$; $t(168) = 20.80$, $p < .001$; CTRS-28 hyperactivity: hyperactive mean = 75.37, $SD = 10.49$, control mean = 47.18, $SD = 6.27$; $t(65.82) = 17.88$, $p < .001$.

The average age for children in the hyperactive group was 8.8 years ($SD = 1.2$ years) and for the children in the control group 9.0 years ($SD = 1.4$ years). The sex distributions were very similar: 52.9% of the hyperactive children and 55.5% of the control children were girls. (Note that the use of T-scores from the CTRS-28 and CPRS-48 gives sex- and age-adjusted scores.) An independent *t*-test and a chi-square test showed that these differences between the groups on age or sex were not significant.

We used the Standard Occupational Classification from the Office of Population Censuses and Surveys (1990) to classify the families' social class. The family's social class was based either on mother's or father's occupation, whichever was the highest. The percentage of families classified in each of the six categories was as follows: 2% professional etc. occupations, 40% managerial and technical occupations, 23% skilled occupations (non-manual), 27% skilled occupations (manual), 6% partly skilled occupations, and 2% unskilled occupations. In terms of ethnic origin, 92% of the children were Caucasian, 1% were Indian/Pakistani, 1% Asian, 2% African/Caribbean, and 3% were classified as "other".

Procedure

The parents were first interviewed, after which they left the testing room. Two testers then assessed the twins simultaneously in two separate rooms. The order of task presentation was fixed, but different for "twin A" and "twin B". The order of task presentation was counterbalanced across the testers, that is, each tester tested "twin A" equally often. The twins were given

an approximately 20-minute break in the middle of the testing and also several shorter breaks. In total, the testing session took approximately 3 hours. The testers were "blind" with regard to the group status of the twins. The families did not receive any financial reward for participation.

Measures

The Maudsley Index of Childhood Delay Aversion (Kuntisi, Stevenson, Oosterlaan, & Sonuga-Barke, in press). This is a new computer task designed to test the delay aversion theory of hyperactivity. The full task involves several conditions, but in the present study we included only the condition that is predicted to show differences between hyperactive and control children (Sonuga-Barke, personal communication, 1996). In this task the child has to make a choice, 20 times, between a small immediate reward (1 point involving a 2-second pre-reward delay) and a large delayed reward (2 points involving a 30-second pre-reward delay). If the child chooses the small reward, the next trial starts immediately afterwards; this of course reduces the overall length of the session.

The task is presented as a space game, in which the child, as a captain of a space ship, has to destroy enemy spacecraft (using the computer mouse). The aim of the game is to earn as many points as possible and to motivate the children they are told that they will receive a small prize in the end (in this study the children received pencils). Before the experimental trials, the child first practises using the mouse and choosing each of the rewards. The tester also asks the child questions about the game, to ensure that he or she has understood the rules and aims of the game correctly. The delay aversion variable used in the analyses is the percentage of choices for the 2 points delayed reward.

The tester also rated the child's apparent delay aversion (the extent to which the child continued talking or doing something else while waiting) on a simple 3-point scale (*not at all, a little, very much*). We obtained good agreement between testers for these ratings ($\kappa = .60, p = .003$).

Stop Task (Logan & Cowan, 1984; Logan, Cowan, & Davis, 1984—the original version; Oosterlaan & Sergeant, 1998—the version used in this study). This computer task measures inhibition and is based on Logan and Cowan's (1984) "race model" of inhibition. This particular version of the task is presented as a game in which the child has to perform tasks similar to those of an air-traffic controller. The child is first taught to respond to aeroplanes appearing on the computer screen by pressing the response button that is on the same side as the plane (a two-choice reaction-time task). The child is then instructed to withhold responding whenever he or she hears a tone on headphones (the "stop" trials), but otherwise to keep on responding to the planes as quickly as possible (the "go" trials). The tones are presented at four different intervals after the presentation of the planes. All children did two practice and four experimental blocks (with 64 trials in each) on this task and were given short breaks between the blocks.

Twenty-five per cent of the trials are stop trials. The stop signals are presented equally often at each of the four stop signal intervals (50, 200, 350, and 500 ms before the child's expected response). The expected moment of responding is estimated from the child's mean reaction time (MRT) in the preceding block of trials. MRT is calculated across correctly executed responses on go trials. The duration of the stop signals, which are 1 kHz tones produced by a function generator, is 100 ms.

Each trial begins with a 350 ms presentation of a fixation point (+ sign presented at the centre of the screen). The presentation of the stimuli (an airplane, displayed for 1500 ms) then follows. The intertrial interval is 1000 ms. A Keithley PIO-12 digital interface board enables the stimuli to be presented and the data to be collected with millisecond accuracy. The stimuli appear equally often on either side of the screen within each block and the stop signals are presented equally often after left- and right-sided presentations of the stimuli. A go trial always follows a stop trial, except once in each block, where two stop

signals are presented in succession (to prevent children expecting a predictable pattern).

The following stop task variables were used in the analyses: slope of the inhibition function, stop signal reaction time (SSRT), mean reaction time (MRT), standard deviation of reaction times (SD of RTs), total number of errors, number of omission errors, and number of commission (left-right) errors (see Appendix for an explanation of the inhibition variables).

Delayed Response Alternation Task (DRA; Carpenter & Gold, 1994; Gold, Faith Berman, Randolph, Goldberg, & Weinberger, 1996). This task is a computerised working memory measure which, to our knowledge, has not previously been used in a hyperactivity study. The task was validated as a prefrontal measure in a study with adults that involved PET scanning (Gold et al., 1996).

In this task two boxes, one coloured (yellow) and the other uncoloured, are first presented on the screen for 1 second. After a 2-second presentation of an empty screen, two uncoloured boxes appear on the screen and the child has to choose one of these boxes, either the one on the side where the coloured box was or the one on the side where the uncoloured box was. The computer gives feedback as to whether the choice was correct or incorrect (the word right or wrong is presented on the screen for 1.5 seconds immediately after the child has responded). New stimuli (another two boxes, one coloured and the other uncoloured) then appear on the screen after a 1.5-second delay.

The task for the child is to find out the rule that the computer uses to decide which box is the correct one each time. If the child does not find out the rule on his or her own, the rule is then taught explicitly. The rule involves choosing the coloured and the uncoloured box (whatever side they appear on) on alternate trials. The position of the coloured box varies randomly. All children performed the task twice, regardless of whether they found out the rule on their own.

We modified the original instructions rather extensively, as these had been written for adults and were not appropriate for children. Before the children started the task proper, they first practised responding (pressing the numbers 1 and 2 on the keyboard) with a practice version of the task. In this practice version the correct rule was always to choose the coloured (blue) box. The children were told after the practice that the rule might be different in the "real game".

The DRA variables used in the analyses are the percentage of correct choices pre- and post-instruction.

Sentence Span (Daneman & Carpenter, 1980—the original version; Siegel & Ryan, 1989—the version used in this study) and *Counting Span* (Case, Kurland, & Goldberg, 1982) tasks. These tasks are working memory measures. In the sentence span task, the tester reads sentences out to the child who has to supply the missing last word for each sentence. In the end of each set, the child is asked to repeat all the words that he or she had supplied, in the correct order. The tester first gives the child a practice sentence and then, in order also to practise recalling the supplied words, two further sentences. The task proper begins with two-sentence sets and, unless the child fails all three sets of any level, finishes with five-sentence sets. The sentences for the task have been chosen so that the missing word is virtually predetermined. However, the particular word that the child supplies is not important. We made some modifications to the sentences in order for them to be more appropriate for British children.

The counting span task is similar to the sentence span task except that the child is asked to count yellow dots on cards rather than to supply words. The tester asks the child to touch each yellow dot with his or her finger and to count out loud. The child is asked to ignore blue dots on the cards, which are arranged randomly with the yellow dots to prevent counting by subitising. The practice starts with counting the yellow dots on one card. The tester then, presenting one card at a time, asks the child to count the dots on two cards and, when presenting a blank card, to recall the numbers of dots on the previous cards. The testing proper starts with two-card sets and, unless the child fails all three sets of any level, finishes with five-card sets. The

size of the cards was 14 cm × 21 cm and the dots were 0.9 cm in diameter. The possible scores range from 0 to 12 on both working memory tasks.

Wechsler Intelligence Scales for Children (WISC-III^{UK}; Wechsler, 1992). Four subtests from the WISC were used to obtain an estimate of the child's IQ: Picture Completion and Block Design provided an estimate of Performance IQ, and Vocabulary and Similarities an estimate of Verbal IQ. The four subtests together provided an estimate of the child's Full Scale IQ.

Ratings of behaviour during testing. After the testing session, the tester rated the child on hyperactive behaviours (fidgeting, lower limb movements, bottom shuffling movements, and gross motor activity) on a 4-point scale, ranging from "not at all" to "very much".

The test-retest reliabilities for the tasks used in the present study are reported in a separate paper (Kuntsi, Stevenson, et al., in press).

Results

We first investigated whether IQ, conduct problems, or anxiety should be included in subsequent analyses as covariates. There were significant IQ differences, multivariate $F(2, 166) = 4.35$, $p < .01$, in the mean Performance IQ: hyperactive mean = 93.90, $SD = 20.08$, control mean = 100.20, $SD = 16.29$; univariate $F(1, 167) = 4.85$, $p < .03$, and mean Verbal IQ: hyperactive mean = 93.49, $SD = 16.66$, control mean = 101.81, $SD = 17.72$; univariate $F(1, 167) = 8.12$, $p < .01$. The groups also differed significantly on ratings of both conduct problems—average of teacher and parent ratings: hyperactive mean = 71.28, $SD = 13.56$, control mean = 49.50, $SD = 8.36$; $t(66.85) = 10.64$, $p < .001$ —and anxiety: hyperactive mean = 57.49, $SD = 14.62$, control mean = 53.05, $SD = 10.43$; $t(72.72) = 1.98$, $p < .05$. The task variables

were significantly correlated with Full Scale IQ and ratings on conduct problems, but not with ratings on anxiety. We therefore decided to include IQ and conduct problems as covariates in the analyses, whilst also reporting the results from analyses with no covariates.

Working Memory Tasks

The delayed response alternation task (DRA) pre- and post-instruction scores were analysed together in a MANOVA (see Table 1). Before teaching the children the rule, both groups performed at chance levels. After the children had been taught the rule, the control children performed significantly better than the hyperactive children. Whereas controlling for conduct problems does not alter the pattern of findings, controlling for IQ removes the significant group difference on the "post-instruction" variable.

On the DRA we asked the children, after they had performed the task for the first time, to identify the rule. A fifth (21%) of the control children and 14% of the hyperactive children found out the rule on their own, $\chi^2(1) = 1.29$, $p = .26$.

Table 2 summarises the results for the counting span and sentence span tasks, which were similarly analysed together in a MANOVA. The means for the two groups on the sentence span and counting span tasks show that hyperactive children tended to perform less well than the control children. However, the only significant group difference was that for sentence span without controlling for IQ. When IQ was included in the analysis as a covariate, this difference between the groups was no longer significant. With conduct problems as a covariate, the group comparisons are similarly nonsignificant. The mean scores on these two tasks suggest that both groups

Table 1
Group Comparisons for DRA Scores

Variable	MANOVA results and effect sizes								MANOVA results with covariates					
	Hyperactive (<i>N</i> = 51)		Control (<i>N</i> = 118)		Univariate			Effect size ^a	ANCOVA IQ as a covariate			ANCOVA CP as a covariate ^b		
	Mean	<i>SD</i>	Mean	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>		<i>F</i>	<i>df</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>p</i>
DRA <i>pre</i>	50.25	11.75	51.11	14.19	0.14	1,167	.71	.06	0.02	1,165	.89	2.04	1,166	.16
DRA <i>post</i>	67.94	14.92	74.70	16.00	6.61	1,167	.01	.42	2.64	1,165	.11	6.00	1,166	.02
Multivariate					3.37	2,166	.04		1.38	2,164	.26	3.25	2,165	.04

^a Effect size = difference in the group means/*SD* in the control group.

^b CP = an average of parent and teacher ratings on the conduct problem subscales (T-scores).

Table 2
Group Comparisons for Sentence Span and Counting Span Tasks

Variable	MANOVA results and effect sizes								MANOVA results with covariates					
	Hyperactive (<i>N</i> = 51)		Control (<i>N</i> = 119)		Univariate			Effect size ^a	ANCOVA IQ as a covariate			ANCOVA CP as a covariate ^b		
	Mean	<i>SD</i>	Mean	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>		<i>F</i>	<i>df</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>p</i>
Sentence span	3.33	1.86	4.24	1.83	8.61	1,168	.004	.50	2.82	1,166	.10	0.68	1,167	.41
Counting span	4.71	3.12	5.66	3.06	3.39	1,168	.07	.31	0.50	1,166	.48	0.47	1,167	.49
Multivariate					4.29	2,167	.02		1.41	2,165	.25	0.37	2,166	.69

^a Effect size = difference in the group means/*SD* in the control group.

^b CP = an average of parent and teacher ratings on the conduct problem subscales (T-scores).

Table 3
Group Comparisons for Delay Aversion Task and Stop Task Variables

Variable	<i>t</i> -test results and effect sizes								ANCOVA results with covariates					
	Hyperactive (<i>N</i> = 49–51)		Control (<i>N</i> = 118)		<i>t</i> -test			Effect size ^a	ANCOVA IQ as a covariate			ANCOVA CP as a covariate		
	Mean	<i>SD</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>		<i>F</i>	<i>df</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>p</i>
Delay aversion	40.29	5.54	53.22	7.61	−2.86	167	.005	.47	3.82	1,165	.05	1.60	1,166	.21
Inhibition slope	.136	.042	.142	.046	−0.80	165	.43	.13	0.03	1,163	.86	1.71	1,164	.19
SSRT	238.67	80.90	222.27	68.49	1.33	165	.18	.24	0.67	1,163	.41	1.64	1,164	.20
MRT	527.41	95.76	475.55	101.76	3.05	165	.003	.51	6.72	1,163	.01	1.19	1,164	.28
<i>SD</i> of RTs	142.37	40.32	114.30	33.74	4.62	165	.001	.83	14.57	1,163	.001	4.46	1,164	.04
Commission errors	4.90	4.97	2.85	5.06	2.40	165	.02	.41	7.56	1,163	.007	3.55	1,164	.06
Omission errors	5.29	6.36	2.36	4.32	2.95	67.12	.004	.68	3.45	1,163	.07	6.59	1,164	.01
Total errors	10.18	9.64	5.20	7.70	3.21	74.66	.002	.65	7.84	1,163	.006	7.21	1,164	.008

^a Effect size = difference in the group means/*SD* in the control group.

^b CP = an average of parent and teacher ratings on the conduct problem subscales (T-scores).

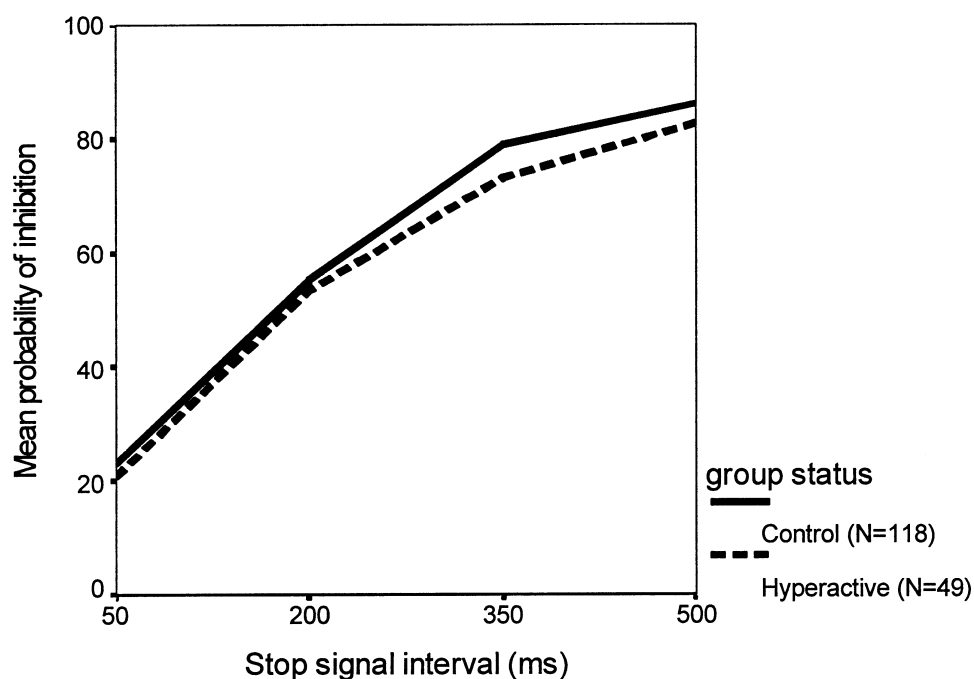


Figure 2. Inhibition functions: probability of inhibition as a function of the stop signal interval (MRT – stop signal delay).

performed relatively better on the counting span than on the sentence span task. Paired *t*-tests confirmed that these within-group differences were significant; hyperactive: $t(50) = 4.06$, $p < .001$; control: $t(118) = 6.08$, $p < .001$.

Delay Aversion and Stop Task

Table 3 presents the group comparison results for the delay aversion measure and stop task variables, with Full Scale IQ or conduct problems as covariates (ANCOVAs) and without covariates (*t*-tests). On the delay aversion task, hyperactive children chose the larger reward significantly less often than the control children, and this group difference remained significant after controlling for IQ. With conduct problems as a covariate, this difference between the groups is no longer significant.

We also rated the children's apparent aversion to delay during the task for a subgroup of the sample: this rating was instigated after the first 19 twin pairs had been tested. A Mann-Whitney U-test on these data was highly significant, $U = 1015.50$, $p < .001$, which indicates that the groups differed on the degree of "delay aversion"

they showed during the task. More than half of the control children, but only 16% of hyperactive children, did not appear at all aversive to delay. In contrast, 44% of the hyperactive children, but only 13% of the control children, obtained the highest rating of "very much" or consistently chose the small, immediate reward. We also asked the children how they had decided whether to choose the smaller or the larger reward ("how did you decide whether to fire your phasers the first time or the second time that the box changed from green to red?") and grouped the responses into five categories. The groups did not differ significantly in the reasons given for the choices made on the task, $\chi^2(4) = 1.18$, $p = .88$.

Of the stop task variables, the group comparisons were nonsignificant for the slope of the inhibition function and SSRT and significant for MRT, *SD* of RTs, and the error variables (Table 3). As the group comparison for the slope of the inhibition function was nonsignificant, we did not carry out the ZRFT-correction (see Appendix). The effect size is highest for the *SD* of RTs, which is noticeably higher than for any other variable: for example, the effect size is .47 for the delay aversion

Table 4
Time-on-task Effects: MRT and SD of RTs

	MRT				SD of RTs			
	First half		Second half		First half		Second half	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Hyperactive ($N = 49$)	538.11	103.46	516.64	92.15	139.35	39.92	140.17	42.25
Control ($N = 118$)	480.86	107.26	470.35	98.82	112.60	35.83	112.88	33.69
ANOVA	F	df	p		F	df	p	
Group	9.27	1,165	.003		20.03	1,165	.001	
Time-on-task	27.04	1,165	.001		0.12	1,165	.73	
Group \times Time-on-task	3.17	1,165	.08		0.03	1,165	.86	

variable, whereas it is .83 for the SD of RTs. We also carried out these analyses excluding children who had error rates higher than 10%, as this has been the practice in some previous stop task studies. With these cases excluded from the analyses, the t -test results remained significant for both MRT and SD of RTs. Controlling for IQ did not alter the pattern of findings. With conduct problems as a covariate, the group comparison for MRT is no longer significant and the p -value for commission errors also narrowly misses significance.

Figure 2 shows the inhibition functions for the two groups. For both groups the mean probability of inhibition increased as the stop signal interval increased. The inhibition slopes were calculated by fitting regression lines to the individual inhibition functions.

To investigate whether the speed of responding on the task remains stable over time in each group, a repeated measures ANOVA was carried out, with group as the between-subjects variable (see Table 4). MRT was calculated across the first half (blocks 1 and 2) and the second half (blocks 3 and 4). The results show that both groups were faster during the second half of the task (a practice effect) and that children in the hyperactive group were generally slower than the control children. The group by time-on-task interaction was not significant.

Table 4 also reports the results from a similar analysis for the SD of RTs. Only the main effect for group was significant, indicating that the hyperactive children were more variable in the speed of responding than control children, but there was no change for either group in the variability of speed over time.

Is there evidence of a speed accuracy trade-off? The correlation between MRT and total number of errors was .24. Considering the two types of errors separately, it is clear that there is a relationship between MRT and omission errors, $r = .45$, $p < .001$, but not between MRT and commission errors, $r = -.04$, $p = .6$. As the correlation between MRT and omission errors is positive, it shows an association between *slow* speed and a high number of omission errors. Therefore there was no evidence of a speed accuracy trade-off.

Ratings of Behaviour During Testing

A summary score of the four observation rating variables was derived. On this aggregate score, a higher score indicates more hyperactive behaviour. The result of a t -test was highly significant: hyperactive, $N = 43$, mean = 6.28, $SD = 3.37$; control, $N = 90$, mean = 2.32, $SD = 2.37$; $t(62.63) = 6.92$, $p < .001$, indicating that hyper-

Table 5
Results from the Discriminant Function Analysis: Linear Classification Functions (LCFs) and Linear Discriminant Functions (LDFs)

Variable	LCFs	LDFs
SD of RTs	.901	.813
Omission errors	-.026	.597
Sentence span	-.096	-.561
Delay aversion	-.258	-.545
MRT	-.117	.527
Verbal IQ	-.191	-.507
DRA after teaching	-.056	-.475
Commission errors	.265	.453
Performance IQ	-.200	-.387
Counting span	.204	-.364
SSRT	-.094	.256
Inhibition slope	.434	-.154
DRA pre-instruction	.069	-.077

active children were rated as more active and fidgety than control children.

Correlations between these testers' ratings of hyperactive behaviour during the session with parents', $r = .46$, $p < .01$, and teachers', $r = .57$, $p < .01$, ratings of hyperactivity suggest a degree of cross-informant consistency in hyperactivity ratings. However, as only children from whom we have the observational ratings were included in these analyses, the sample is not representative of the general population. For the larger sample representative of the general population ($N = 250$; see Kuntsi & Stevenson, 2001), the correlation between teacher and parent ratings of hyperactivity was .33 ($p < .01$), whereas for this smaller sample from whom we have the observational ratings the correlation was .71 ($p < .01$). This shows how a higher degree of cross-informant consistency in hyperactivity ratings is obtained for a sample with an excess of hyperactive children.

Discriminant Function Analysis

To contrast, within a single analysis, the efficacy of the various variables in discriminating between the hyperactive and control groups, a discriminant function analysis was carried out. The variables included in the discriminant function analysis were all the main task variables and Performance and Verbal IQs. Overall, the discriminant function significantly distinguished between the hyperactive and control groups, Wilks lambda = 0.837, $\chi^2(13) = 27.73$, $p < .02$.

Table 5 shows two sets of coefficients from the

discriminant function analysis: linear classification functions (the actual weights used to create the discriminant scores) and linear discriminant functions. The latter give the relative magnitude of the contributions to the discriminant score, that is they take into account scaling differences between the measures. The *SD* of RTs on the stop task comes out again as the strongest variable, with a linear discriminant function (LDF) score of .81. Other variables with LDF scores of $+/- .5$ or higher are omission errors, sentence span, delay aversion, MRT, and Verbal IQ.

Sex Effects

Separate ANOVAs for each of the task variables were carried out to investigate possible sex effects on task performance. The results showed that there were no significant main effects for sex or sex \times group interactions.

Discussion

The main research question was whether children with hyperactivity show a response inhibition deficit, impaired working memory, or a tendency to be particularly aversive to delays.

As in the previous study with a similar delay aversion task (Sonuga-Barke et al., 1992), hyperactive children chose the small, immediate reward more often than control children. This finding supports the delay aversion hypothesis: hyperactive children seem to aim to reduce the overall length of the delay period, at the expense of earning higher rewards. The testers' ratings on the apparent delay aversion of the children confirmed this finding, as a similar group difference emerged on these ratings. The hyperactive children not only chose the immediate reward more often, but they also appeared to find the waiting period very aversive (they continued talking or doing something else) when they did wait for the larger reward.

Both the current study and the study by Sonuga-Barke et al. (1992) found a significant group difference on a computerised delay aversion measure. The difference in the results between the two studies relates to the magnitude of the difference. In the Sonuga-Barke et al. study hyperactive children chose the larger reward on average in 18 % of the trials (and control children; in 48 % of the trials), compared to 40 % of the trials in the present study (53 % for the control group). A possible explanation for the "stronger" delay aversion tendency of hyperactive children in the original study is the age of the participants: the children in the Sonuga-Barke et al. study were younger (6- and 7-year-olds) than the children in the present study. Some support for this possibility comes from the finding in the present study of a correlation between age and performance on the delay aversion task (.41). In the present study the children also had to perform the task only once, whereas the original study included several different conditions.

Due to time constraints for the testing sessions, we did not include in this study the condition where there was a delay period after choosing the small reward. Sonuga-Barke et al. (1992) showed that in this situation, where choosing the immediate reward did not reduce the overall delay period, the hyperactive and control groups were indistinguishable. This led to the argument that hyperactive children are not truly impulsive, but are delay averse.

Although the hyperactive group obtained lower Performance and Verbal IQs than the control group, the group difference on the delay aversion measure remained significant after controlling for IQ. The hyperactive group's tendency to choose the immediate reward more often was not due to their lower general cognitive ability.

The group difference on the delay aversion measure did not remain significant after controlling for conduct problems. This suggests that co-occurring conduct problems explain most, if not all, of the association between hyperactivity and delay aversion. Sonuga-Barke et al. (1992) did not investigate this issue of whether delay aversion is *specific* to hyperactivity.

Few previous studies have investigated hyperactive children's performance on working memory measures. The study using the counting span and sentence span tasks (Siegel & Ryan, 1989) reported no differences between ADD and control groups on most comparisons. The only comparison on which children with ADD performed worse was that for the younger age group (7- and 8-year-olds) on the sentence span task. In the present study hyperactive children performed significantly worse than control children on the sentence span but not on the counting span task. There was a trend, however, for the hyperactive group to perform worse on the counting span task too.

The sample in the study by Siegel and Ryan (1989) was different from the present hyperactive sample in that they were not required to show *pervasive* symptoms and both groups of ADD children also had unusually high average IQs (108 and 112). Because of the association between IQ and working memory (see below), having a sample of ADD children with unusually high IQs may have "hidden" a working memory impairment, if such an impairment existed. Indeed, the group difference on the sentence span task disappeared in the present study when IQ was controlled for. Controlling for conduct problems similarly resulted in a nonsignificant group difference on the sentence span task.

Taken together, data from the present and the previous (Siegel & Ryan, 1989) studies suggest that the relationship between hyperactivity and poor performance on these working memory tasks is stronger for younger children and for the sentence span rather than the counting span task. In the present study both hyperactive and control children performed relatively better on the counting span task than on the sentence span task. The most obvious difference between the tasks relates to the overall level of difficulty. Although the words in the sentence span task were meant to be virtually predetermined, some of the children had to think for a short while about a suitable answer. This slowed down the pace of the task, which presumably increased the working memory demand. The two tasks also differ on how active a role the child plays: on the counting span task the child is actively involved in counting the dots with his or her finger, whereas on the sentence span task the child simply responds verbally to verbal items the tester presents.

On the computerised working memory measure, the delayed response alternation task, both groups performed at chance levels before they were taught the rule. This was clearly a difficult task for children between the ages of 7 and 11 years. The "pre-instruction" score does not therefore reflect working memory ability. The groups did not differ significantly in the numbers of children reporting to have found out the rule on their own: 14 % of the hyperactive children and 21 % of the control

children did so. To our knowledge this task has neither been used with children nor in a study on hyperactivity in previous research.

When the children performed the task again after having been taught the rule, the hyperactive group performed significantly worse than the control group. The group difference remained significant after controlling for conduct problems. The crucial issue here is the interpretation of this result: does it indicate impaired working memory or could it be a consequence of something else (e.g. a nonoptimal effort/activation state; see below)? The inconsistency in hyperactive children's performance on working memory measures in the present and previous studies is the first reason to question the working memory impairment hypothesis.

Another complicating issue is the failure in the present study to find any group differences on the working memory tasks when IQ was controlled for. Whether controlling for IQ is in fact appropriate is a complex theoretical issue. The argument for including IQ as a covariate in the analyses is that the groups differed significantly on IQ. Any group differences on task performance could be due to differences in general cognitive ability. The counter-argument is that the lower IQ of the hyperactive children is very much part of the disorder and gives clues about the nature of the "core" deficit. Pennington (1994) emphasises the relation between working memory and intelligence: "working memory is clearly one very important mechanism that appears to account for considerable variance in what is called fluid intelligence" (p. 267). Nevertheless, controlling for IQ did not remove the group differences on the delay aversion measure or those stop task variables which differentiated between the groups. It is possible that when hyperactive children show poor performance on working memory tasks, it is a consequence of something else.

With regard to the response inhibition deficit hypothesis, the meta-analysis of previous stop task studies concluded that hyperactivity is associated with a slower inhibitory process (Oosterlaan et al., 1998). That is, the significant group difference on the slope of the inhibition function was due to a slower inhibitory process, as well as greater variability of reaction times on the primary task. The present study did not find this: the groups did not differ significantly on the stop signal reaction time. The only study of those included in the meta-analysis of stop task studies (Oosterlaan et al., 1998), which similarly failed to find a difference between hyperactive and control groups on the stop signal reaction time, was that by Daugherty et al. (1993). Like the present study, and unlike the other stop task studies, Daugherty et al. recruited their sample of hyperactive children from the general population. This raises the possibility that the slow inhibitory process is characteristic only of clinic-referred children with hyperactivity or ADHD.

In the meta-analysis the group difference on the inhibition slope disappeared after the ZRFT correction, leading to the conclusion that hyperactive children are neither less likely to trigger the inhibition process nor are they less variable in their inhibitory process (Oosterlaan et al., 1998). The present results support this conclusion: the failure to find a group difference on the slope of the inhibition function ruled out these possibilities.

There is reason to be somewhat cautious about conclusions regarding the speed of the inhibitory process, however. The test-retest reliabilities were lower for both the inhibition function slope and the stop signal reaction

time than for most other variables (Kuntsi, Stevenson, et al., in press). It may be that the failure to find significant group differences on these variables was due to their lower reliability. The theoretically important issue is whether the groups would have differed on stop signal reaction time, had it been a more reliable measure of the speed of the inhibitory process. If the groups had differed on the inhibition slope too, this would have simply reflected the group differences on the stop signal reaction time and the variability of speed on the primary task. It is extremely unlikely, had the inhibition slope result been significant, that it would have remained significant after the ZRFT-correction. Previous research has shown that the ZRFT-correction removes the significant difference between the groups: hyperactive children are not less likely to trigger the inhibitory process nor do they have a more variable inhibitory process than other children (Oosterlaan et al., 1998).

To summarise the stop task results, the hyperactive children were not less likely to trigger the inhibitory process nor was their inhibitory process more variable, compared to control children. The groups also did not differ significantly in terms of the speed of the inhibitory process, although there is a slight possibility that this was due to the lower reliability of this measure. Even if hyperactivity was associated with a slow inhibitory process, this does not necessarily indicate a specific response inhibition deficit. Given that hyperactive children are also slower in their responses on the primary task (see below), these results together may indicate a generally slow mode of information processing.

Aside from the issue of a response inhibition deficit and the variables measuring it, other stop task variables did pick up significant group differences. Children in the hyperactive group were slower (in terms of their reaction times on the primary task), more variable in their speed of responding (on the primary task), and they made a higher number of both omission and commission errors than the children in the control group. However, neither the hyperactive children, nor the control children, became slower or more variable in their speed of responding over time. In fact, both groups became *faster* on the primary task over time, indicating a practice effect.

The lower IQs of the children with hyperactivity could not explain the pattern of findings that emerged, as the hyperactive children were slower and more variable in their speed and made more errors even when IQ was controlled for. Controlling for conduct problems reduced the size of the group differences, such that two of the significant effects with the smallest effect sizes (mean reaction time and commission errors) became non-significant.

Rather than brushing aside the findings regarding the stop task variables that do not measure response inhibition as uninteresting, perhaps they give important clues about the nature of the disorder. Overall, the variability in the speed on the reaction time task (the primary task of the stop task) best discriminated between the hyperactive and control groups. The effect size was highest for this variable and it came out as the strongest variable in the discriminant function analysis.

What could cause some children to be particularly variable in their speed of responding on a reaction time task, sometimes responding fast and at other times slowly? One possibility Oosterlaan and Sergeant (1996) suggested is that it may indicate lack of consistent effort. The children with hyperactivity were not only more

variable in their reaction times, but they were also generally slower and made more errors. This pattern of responding (slow, variable, and inaccurate) conforms to the pattern Van der Meere (1996) identified in his review as characteristic of hyperactivity. These findings could therefore be interpreted as supporting the state-regulation theory of hyperactivity (Van der Meere, 1996), which argues that the core problem in hyperactivity relates to a nonoptimal activation/effort state: "the engine is intact (i.e. the basic information processing capacity is intact), but there is a problem with the supply of petrol" (p. 133). Although Van der Meere emphasises the role of the effort and activation systems in particular, he notes that the possible involvement of the arousal system cannot be excluded either.

The description of hyperactivity as a state-regulation problem not only fits well with the present findings but also with the findings from previous research. Manipulations of state factors, such as presentation rate and external control, have been shown to influence hyperactive children's performance strongly (Van der Meere, 1996; see also Sergeant et al., 1999).

The state-regulation theory emphasises those aspects of ADHD that are most similar to the personality theory concept of novelty seeking or sensation seeking. Sensation seeking refers to characteristics such as boredom susceptibility and thrill and experience seeking (Zuckerman, 1996). Adults with ADHD indeed score higher than other adults on a novelty-seeking scale (Downey, Stelson, Pomerleau, & Giordani, 1997). One of the genes that have been associated with ADHD, the dopamine D4 receptor gene polymorphism (e.g. Faraone et al., 1999; LaHoste et al., 1996; Swanson et al., 1998), has also been associated with novelty seeking (Benjamin et al., 1996; Ebstein et al., 1996; Ebstein, Nemanov, Klotz, Gritsenko, & Belmaker, 1997).

There is a need to confirm that the pattern of responding, which was so strongly associated with hyperactivity, reflects a state-regulation problem and not a cognitive deficit. If it reflects a state-regulation problem, under certain circumstances hyperactive children should not show their typical pattern of responding. That is, the aim would be to set up a task that would optimise their activation/effort state. Research by Douglas and colleagues (e.g. Douglas & Parry, 1983) has shown that simply continuous positive verbal feedback can at least reduce the variability of reaction times and decrease mean reaction times of hyperactive children.

It may not be possible to optimise hyperactive children's activation/effort state, if it in fact reflects "learned helplessness". Aspects of task performance that appear state-related could in fact have a cognitive basis. If hyperactive children found certain tasks difficult because of a cognitive deficit, they might give up even trying to do well on such tasks. Lack of effort would then reflect a past history of failures. The evidence from studies on the effects of medication make this a less plausible explanation, however. For example, hyperactive children did not show their typical pattern of responding on the stop task while on medication (Tannock, Schachar, & Logan, 1995). That is, medication had the immediate effect of accelerating the inhibitory process and the primary-task responses, making the responses less variable and improving error rates.

If hyperactive children indeed perform worse than other children on certain tasks due to a state-regulation problem rather than impaired cognitive functioning, this

has also practical implications. Nevertheless, it is not implausible that hyperactivity is associated both with a state-regulation problem and some cognitive deficit.

A limitation of the present study relates to the measurement of inhibition. The stop task focuses on *momentary* inhibition, the ability to suppress a particular response when it is signalled. It does not measure *ongoing* inhibition, which some accounts (e.g. Barkley, 1997) view as important in hyperactivity (see Sonuga-Barke, 1995, for a further discussion). The inhibition variables also had lower test-retest reliabilities than the other variables, as discussed above. Despite these limitations, the present and previous findings clearly indicate that the response inhibition hypothesis is in need of a revision; it does not explain the *overall* pattern of findings.

Many studies on hyperactivity have not included girls at all. It was one of the aims of the present investigation to explore this neglected issue of the nature of sex effects in hyperactivity. Heptinstall and Taylor (1996) suggest that hyperactivity is truly less common among girls than boys. In the present study, rather than comparing girls who show "clinically significant" hyperactivity to boys who show "clinically significant" hyperactivity, we compared the boys showing extreme hyperactive behaviours (relative to other boys) to a similar proportion of extreme girls. On the performance on the tasks, hyperactive girls were indistinguishable from hyperactive boys. This suggests that even if hyperactive girls show hyperactive behaviours to a lesser extent (Kuntsi, Barrett, Canning, Karia, & Stevenson, 2000), they still show the same pattern of responding on tasks as hyperactive boys.

Some validation for the present definition of hyperactivity came from the observational ratings of hyperactive behaviours during the session, as these blind ratings differentiated well between the hyperactive and control groups. The current data on the hyperactivity ratings on the Conners' scales suggest a similar degree of cross-informant consistency as previous research. Goyette et al. (1978) reported a correlation of .36 between teacher and parent ratings on the hyperactivity dimensions of the Conners' scales; the correlation for the sample representative of the general population was .33 for our data.

In sum, the predicted group differences emerged on the delay aversion measure and some of the working memory tasks, although the group differences on the working memory measures did not remain significant after the effects of IQ were removed. Hyperactive children did not differ from the control children on the inhibition variables. However, other stop task variables picked up significant group differences: hyperactive children were variable in their speed, generally being slow and inaccurate. This pattern of responding may indicate a nonoptimal effort/activation state. Of all the task variables, the variability in speed of responding best discriminated between the groups. The finding that controlling for conduct problems removed significant group differences on some of the variables, including the delay aversion measure, indicates the need for the co-occurrence issue to be tackled from a theory-driven perspective in future investigations.

Acknowledgements—We thank all the children, parents, and teachers who took part in this study; Doug Barrett, Emma Canning, and Nisha Karia for invaluable help during data collection; and Tony Charman, Edmund Sonuga-Barke, and Jaap van der Meere for their helpful comments on an earlier

draft of this manuscript. This research was funded by a Wellcome Prize Studentship to the first author.

References

- Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, 40, 57–87.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121, 65–94.
- Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., & Hamer, D. H. (1996). Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nature Genetics*, 12, 81–84.
- Carpenter, C. J., & Gold, J. M. (1994). *Prefrontal functioning in schizophrenia*. Paper presented at the International Neuropsychological Society Annual Meeting, Seattle, February.
- Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386–404.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Daugherty, T. K., Quay, H. C., & Ramos, L. (1993). Response perseveration, inhibitory control, and central dopaminergic activity in childhood behavior disorders. *The Journal of Genetic Psychology*, 154, 177–188.
- Douglas, V. I., & Parry, P. A. (1983). Effects of reward on delayed reaction time task performance of hyperactive children. *Journal of Abnormal Psychology*, 11, 313–326.
- Downey, K. K., Stelson, F. W., Pomerleau, O. F., & Giordani, B. (1997). Adult attention deficit hyperactivity disorder: Psychological test profiles in a clinical population. *Journal of Nervous and Mental Disease*, 185, 32–38.
- Ebstein, R. P., Nemanov, L., Klotz, I., Gritsenko, I., & Belmaker, R. H. (1997). Additional evidence for an association between the dopamine D4 receptor (D4DR) exon III repeat polymorphism and the human personality trait of novelty seeking. *Molecular Psychiatry*, 2, 472–477.
- Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., & Belmaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature Genetics*, 12, 78–80.
- Faraone, S. V., Biederman, J., Weiffenbach, B., Keith, T., Chu, M. P., Weaver, A., Spencer, T. J., Wilens, T. E., Frazier, J., Cleves, M., & Sakai, J. (1999). Dopamine D₄ gene 7-repeat allele and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 156, 768–770.
- Gold, J. M., Faith Berman, K., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1996). PET validation of a novel prefrontal task: Delayed response alternation. *Neuropsychology*, 10, 3–10.
- Gorenstein, E. E., Mammato, C. A., & Sandy, J. M. (1989). Performance of inattentive-overactive children on selected measures of prefrontal-type function. *Journal of Clinical Psychology*, 45, 619–632.
- Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology*, 6, 221–236.
- Heptinstall, E., & Taylor, E. (1996). Sex differences and their significance. In S. Sandberg (Ed.), *Hyperactivity disorders of childhood* (pp. 329–349). Cambridge: Cambridge University Press.
- Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, 111, 127–155.
- Kuntsi, J., Barrett, D., Canning, E., Karia, N., & Stevenson, J. (2000). The Conners' Teacher Rating Scale (CTRS-28) applied to a UK sample. *Journal of Attention Disorders*, 3, 229–237.
- Kuntsi, J., & Stevenson, J. (2000). Hyperactivity in children: A focus on genetic research and psychological theories. *Clinical Child and Family Psychology Review*, 3, 1–23.
- Kuntsi, J., & Stevenson, J. (2001). Psychological mechanisms in hyperactivity: II The role of genetic factors. *Journal of Child Psychology and Psychiatry*, 42, 000–000.
- Kuntsi, J., Stevenson, J., Oosterlaan, J., & Sonuga-Barke, E. J. S. (in press). Test-retest reliability of a new delay aversion task and executive function measures. *British Journal of Developmental Psychology*.
- LaHoste, G. J., Swanson, J. M., Wigal, S. B., Glabe, C., Wigal, T., King, N., & Kennedy, J. L. (1996). Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Molecular Psychiatry*, 1, 121–124.
- Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.
- Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295–327.
- Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology*, 10, 276–291.
- Office of Population Censuses and Surveys. (1990). *Standard occupational classification*. London: HMSO.
- Oosterlaan, J., Logan, G. D., & Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD + CD, anxious, and control children: A meta-analysis of studies with the stop task. *Journal of Child Psychology and Psychiatry*, 39, 411–425.
- Oosterlaan, J., & Sergeant, J. A. (1996). Inhibition in ADHD, aggressive, and anxious children: A biologically based model of child psychopathology. *Journal of Abnormal Child Psychology*, 24, 19–36.
- Oosterlaan, J., & Sergeant, J. A. (1998). Response inhibition and response re-engagement in ADHD, disruptive, anxious and normal children. *Behavioural Brain Research*, 94, 33–43.
- Pennington, B. F. (1994). The working memory function of the prefrontal cortices: Implications for developmental and individual differences in cognition. In M. M. Haith, J. B. Benson, R. J. Roberts Jr, & B. F. Pennington (Eds.), *The development of future-oriented processes* (pp. 243–289). Chicago: The University of Chicago Press.
- Pennington, B. F., Bennetto, L., McAleer, O., & Roberts, R. J. Jr (1996). Executive functions and working memory. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 327–348). Baltimore, MD: Paul H. Brookes Publishing Co.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 37, 51–87.
- Quay, H. C. (1988). The behavioral reward and inhibition system in childhood behavior disorder. In L. M. Bloomingdale (Ed.), *Attention deficit disorder: New research in attention, treatment and psychopathology* (pp. 176–186). Oxford: Pergamon Press.
- Quay, H. C. (1997). Inhibition and attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, 25, 7–13.
- Sergeant, J. A., Oosterlaan, J., & Van der Meere, J. (1999). Information processing and energetic factors in attention-deficit/hyperactivity disorder. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 75–104). New York: Plenum Press.
- Shue, K. L., & Douglas, V. I. (1992). Attention deficit hyperactivity disorder and the frontal lobe syndrome. *Brain and Cognition*, 20, 104–124.
- Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development*, 60, 973–980.

- Sonuga-Barke, E. J. S. (1994). Annotation: On dysfunction and function in psychological theories of childhood disorder. *Journal of Child Psychology and Psychiatry*, 35, 801–815.
- Sonuga-Barke, E. J. S. (1995). Disambiguating inhibitory dysfunction in childhood hyperactivity. In J. Sergeant (Ed.), *Eunethydis: European approaches to hyperkinetic disorder* (pp. 209–223). Zurich: Fotorotar.
- Sonuga-Barke, E. J. S., Houlberg, K., & Hall, M. (1994). When is “impulsiveness” not impulsive? The case of hyperactive children’s cognitive style. *Journal of Child Psychology and Psychiatry*, 35, 1247–1253.
- Sonuga-Barke, E. J. S., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion—I. The effect of delay on choice. *Journal of Child Psychology and Psychiatry*, 33, 387–398.
- Swanson, J. M., Sunohara, G. A., Kennedy, J. L., Regino, R., Fineberg, E., Wigal, T., Lerner, M., Williams, L., LaHoste, G. J., & Wigal, S. (1998). Association of the dopamine receptor D4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): A family-based approach. *Molecular Psychiatry*, 3, 38–41.
- Tannock, R. (1998). Attention deficit hyperactivity disorder: Advances in cognitive, neurobiological, and genetic research. *Journal of Child Psychology and Psychiatry*, 39, 65–99.
- Tannock, R., Schachar, R., & Logan, G. (1995). Methylphenidate and cognitive flexibility: Dissociated dose effects in hyperactive children. *Journal of Abnormal Child Psychology*, 23, 235–266.
- Taylor, E., Chadwick, O., Heptinstall, E., & Danckaerts, M. (1996). Hyperactivity and conduct problems as risk factors for adolescent development. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 1213–1226.
- Van der Meere, J. (1996). The role of attention. In S. Sandberg (Ed.), *Hyperactivity disorders of childhood* (pp. 111–148). Cambridge: Cambridge University Press.
- Wechsler, D. (1992). *Wechsler Intelligence Scale for Children—3rd ed.* U.K. London: The Psychological Corporation.
- Zuckerman, M. (1996). The psychobiological model for impulsive unsocialized sensation seeking: A comparative approach. *Neuropsychobiology*, 34, 125–129.

Manuscript accepted 22 June 2000

Appendix

Stop Task Inhibition Variables (taken from Oosterlaan et al., 1998)

Inhibition slope. The inhibition function is generated by plotting the probability of inhibition against mean go signal reaction time minus stop signal delay (MRT–delay). The inhibition slope is calculated by fitting a regression line to the individual inhibition function.

Stop signal reaction time (SSRT). In theory, we integrate the distribution of go signal reaction times from zero to a point in time at which the integral equals the probability of responding given a stop signal (i.e. 1 minus the probability of inhibition). We treat that point as an estimate of the time at which the stop process finished. This time is defined relative to the onset of the go signal (because we use the distribution of go signal reaction times to define it), thus we subtract out stop signal delay to estimate SSRT.

In practice, SSRT is calculated as follows: first, reaction times on go trials are rank ordered on a time axis. Second, we pick the n th reaction time, where n is defined by the product of the number of reaction times in the distribution and the probability of responding given a stop signal (or 1 minus the

probability of inhibition). For example, if there were 100 reaction times in the distribution and the probability of responding given a stop signal was .3, the n th reaction time would be the 30th in the rank-ordered distribution. The n th reaction time is an estimate of the time at which the stop process runs to completion, relative to the onset of the primary task stimulus. Third, we subtract stop signal delay from the n th reaction time and estimate SSRT. For example, if the n th reaction time was 545 ms and the stop signal delay was 200 ms, SSRT would be 345 ms. SSRT is calculated for each stop signal and then averaged.

ZRFT-slope. If there is a group difference on the inhibition slope, the so-called ZRFT transformation is applied to the inhibition function. The ZRFT transformation corrects for differences in mean reaction time, go signal reaction time variability, and SSRT. The probability of inhibition is plotted as a function of a z score that represents the relative finishing time of the go process and the stop process in standard deviation units, using the SD of RTs on the primary task to define these units. ZRFT is obtained with the following formula: $ZRFT = (MRT - \text{stop signal delay} - SSRT) / SD$ of RTs on the primary task. The slope of the inhibition function plotted against ZRFT is known as the ZRFT-slope.