The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance

Article in Journal of Child Psychology and Psychiatry · December 2001

Source: PubMed

CITATIONS
237

READS
3,541

6 authors, including:

Tom Manly
Medical Research Council (UK)
86 PUBLICATIONS 3,829 CITATIONS

Vicki Anderson
Murdoch Childrens Research Institute
442 PUBLICATIONS 11,252 CITATIONS

Peter Watson
Medical Research Council (UK)
87 PUBLICATIONS 3,386 CITATIONS

Ian H Robertson
Trinity College Dublin
370 PUBLICATIONS 13,856 CITATIONS

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Ian Nimmo-Smith
Retrieved on: 22 August 2016
"Attention" is not a unitary brain process. Evidence from adult studies indicates that distinct neuroanatomical networks perform specific attentional operations and that these are vulnerable to selective damage. Accordingly, characterising attentional disorders requires the use of a variety of tasks that differentially challenge these systems. Here we describe a novel battery, the Test of Everyday Attention for Children (TEA-Ch), comprising nine subtests adapted from the adult literature. The performance of 293 healthy children between the ages of 6 and 16 is described together with the relationships to IQ, existing measures of attention, and scholastic attainment. This large normative sample also allows us to test the fit of the adult model of functionally separable attention systems to the observed patterns of variance in children's performance. A Structural Equation Modelling approach supports this view. A three-factor model of sustained and selective attention and higher-level 'executive' control formed a good fit to the data, even in the youngest children. A single factor model was rejected.

There are behavioural and anatomical grounds to believe that Attention Deficit Disorder (ADD) is particularly associated with poor self-sustained attention and behavioural control. The TEA-Ch performance of 24 boys diagnosed with ADD presented here is consistent with this view. When performance levels on WISC-III subtests were taken into account, specific deficits in sustained attention were apparent while selective attention performance was within the normal range.

**Keywords:** ADD/ADHD, assessment, attention, executive function, normal development.

**Abbreviations:** ADD: Attention Deficit Disorder; ADHD: Attention Deficit Hyperactivity Disorder; SEM: Structural Equation Modelling; TEA: Test of Everyday Attention; TEA-Ch: Test of Everyday Attention for Children.

**Introduction**

Vast and increasing numbers of children are referred to clinical services with suspected attentional problems. In some studies, referral rates for Attention Deficit Hyperactivity Disorder (ADHD) have reached 6% of all school age boys and 1.5% of girls (U.S. data: Swanson, Learner, & Williams, 1995). Deficits in attention processes have been attributed to many other developmental conditions (Lang, Athanasopoulous, & Anderson, 1988) including autism (Burack et al., 1997), Asperger syndrome (Klin, Sparrow, Volkmar, Cicchetti, & Rourke, 1995) Tourette syndrome (Georgiou, Bradshaw, Phillips, & Chiu, 1996).

Requests for reprints to: Tom Manly, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, U.K. (E-mail: tom.manly@mrc-cbu.cam.ac.uk).

leukaemia (Brouwers, Riccardi, Fedio, & Poplak, 1985) Turner syndrome (Skuse et al., 1997), and acquired brain injuries (Anderson & Pentland, 1998; Kaufmann, Fletcher, Levin, Miner, & Ewing Cobbs, 1993).

Although the ADHD construct has become increasingly cognitive in its emphasis (moving from the more behaviourally descriptive "hyperkinetic" terminology), the diagnosis continues to rest exclusively on reports of behaviour and inferences about underlying processes. Although parents or teachers may be asked whether a child "has difficulty in sustaining attention", no performance-based measures of such capacities are recommended in the latest DSM revision (American Psychiatric Association, 1994).

Although clearly problematic, this reliance on report and inference reflects both our apparently acute sensitivity to the direction, intensity, and limitations in
another’s attention, and the difficulties experienced by psychometrists in capturing this construct with any precision.

The fundamental problem in measuring attention is that, as a postulated central process, it is at once everywhere and nowhere. The influence of attention cannot be measured unless a person is asked to do something. That something will inevitably involve many other perceptual, cognitive and output systems that may be as—or more—influential on performance than attention.

Separation of Attentional Functions in Adults

Experimental psychology has predominantly operationalised attention as the systematic variation in performance of a single task under different “attentional conditions”. For example, if people are asked to respond to visual targets presented on a computer screen, then their response times to left- or right-sided targets are broadly equivalent. If, however, they receive a useful hint about where the target is most likely to appear, then their performance improves. The absence of eye movements means that the visual stimulation under each condition is identical, as is the task instruction and goal. The difference in reaction times can therefore be interpreted as reflecting a hidden, “top-down” process that modulates detection efficiency in a particular location—attention (Posner, 1980). In a similar way, controlling for reading speed and colour naming allows the conflict condition in the colour-word Stroop task to be interpreted in terms of selective processing and response suppression.

The application of such approaches to studies of adult patients with acquired brain lesions, and more recently, within functional imaging studies, has enhanced understanding of the neural basis of attention and separations between different attentional systems. Reviewing across these areas in 1990, Posner and Petersen argued for three basic characteristics of attention functions within the brain. The first was that specific attention systems existed. The second was that these were separable from more “basic” perceptual, cognitive, and output systems. The third was that within the attention system, specific regions and networks performed different types of operation. On the basis of current evidence these distinct systems were provisionally characterised as: (a) a capacity to move attention within space (spatial attention); (b) a capacity to enhance the processing of particular target characteristics regardless of spatial location (selective attention); and (c) a capacity to maintain a particular processing set over time (sustained attention).

These conclusions have important clinical implications. The functional/anatomical separation of attention systems from basic perception means that—through acquired brain damage or developmental difficulty—it is possible to have deficits that are exclusively or predominantly attentional in nature. The functional/anatomical separation within the attention system means that—again depending on the locus of any damage—quite different profiles of attentional deficits can be seen, each with different implications for problems in everyday life.

For clinicians working with adults this has led to increasingly specific and differential assessments, for example of spatial attention (Driver & Halligan, 1991; Posner & Petersen, 1990; Wilson, Cockburn, & Halligan, 1987), selective or focused attention (Bench et al., 1993; Duncan et al., in press; Pardo, Pardo, Janer, & Raichle, 1990; Stam et al., 1993; Tranel & Hyman, 1990), sustained attention (Brouwer & Van Wooldaer, 1985; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Rueckert & Grafman, 1996; Rueckert, Sorensen, & Levy, 1994; Wilkins, Shallice, & McCarthy, 1987), the capacity to switch (Owen et al., 1993) and divide attention (Baddeley, Bressi, Della Salla, Logie, & Spinnler, 1991) and, of course, a range of higher-level “executive” measures (Burgess & Shallice, 1996; Duncan, 1986; Shallice & Burgess, 1991; Wilson, Alderman, Burgess, Emslie, & Evans, 1997).

In line with this premise of separability, Robertson and colleagues developed the Test of Everyday Attention (TEA; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994, 1996), a single battery for adults with eight subtests designed to make differential demands on sustained attention, selective attention, divided attention, and attentional switching capacities.

Standardising neuropsychological measures requires that large numbers of the healthy population are assessed in order that a given individual’s performance can be meaningfully compared with the population mean. The standardisation of the TEA also provided an opportunity to test Posner and Petersen’s (1990) postulation of a separation between selective and sustained attention, which had been developed on the basis of limited numbers of subjects in functional imaging and lesion studies, on a much larger sample of the neurologically healthy population. If a common, unitary process is thought to contribute heavily to the performance on one group of measures but not to another—and there are individual differences in the efficiency of that process—then the variance in those measures should group together and diverge from others.

The results of a factor analysis conducted on the normalisation sample of the TEA indeed lent statistical support to Posner and Petersen’s proposal. For example, two sustained attention tasks with very different ostensible demands (one counting tones over relatively brief periods and the other monitoring for rare targets within a 10-min stream) loaded onto a common factor and diverged from selective attention measures. Similarly, two visual search tasks had a common loading with the Stroop colour-word task (Trerenery, Crosson, DeBee, & Leber, 1989), supporting the characterisation of a selective attention factor.

Children’s Attention

The main question that we pose in this article—and which underpinned the development of the Test of Everyday Attention for Children (TEA-Ch) battery—is whether the broad views of attention developed predominantly with adults have value in thinking about attention and disorders of attention in childhood. Although it cannot be assumed that the processes of the developing brain correspond closely to those seen in maturity, there are potential advantages if such links can be made. If adult models form reasonable approximations (and presumably an increasingly reasonable approximation as children get older) then both assessment and rehabilitation of attention disorders in children could benefit from the findings of adult studies.

In developing the TEA-Ch our aim was to adapt measures that had proven effective in adult attention into game-like assessment tools for use with children between the ages of 6 and 16. As with the TEA, in standardising
and collecting norms for a new battery we also had the opportunity to statistically test the patterns of convergence and divergence in performance against a broad model developed within the adult literature.

However, as we can only measure attention indirectly this is a somewhat ambitious aim. As children will vary in many abilities (motor skill, task comprehension, language, and so on) that may be shared between tests hypothesised to make different attention demands, our postulated attention factors would have to be expressed very strongly to be seen amid this "noise". Such problems of "task impurity" have been held to account for the generally low correlations observed between executive measures in adulthood (Burgess, 1997; Miyake et al., in press). As developing children may be expected to show greater variability along these nonattentional dimensions, the challenge may be much greater than with adults.

The first approach to meeting this challenge was in the design of the tests. We aimed to minimise the demands on memory\(^1\), reasoning, task comprehension, motor speed, verbal ability, and perceptual acuity while maintaining the demands on the targeted attentional system. To this end, for example, we controlled for motor speed on a visual search task by comparing performance under high and low attentional demands (see Methods section). Where possible, language was avoided in the stimuli or the required responses and the perceptual demands were reduced to detection rather than complex discrimination. Demonstrations and practice trials with correction were used to reduce the impact of task comprehension differences and to try and improve the reliability of performance.

Such attempts can only ever have limited success. The second approach to minimising the impact of "task impurity" was to use Structural Equation Modelling (SEM) in our analysis. This technique is related to the more conventional exploratory factor analysis method but has important differences. In conventional factor analysis, the model is determined in a "bottom-up" fashion from the data. SEM allows one to test an a priori specified model against the observed variance.

If one takes results from three very different tests that are theoretically designed to tap a common underlying process, then that process should be better represented by the covariance between the tests than the variance in any specific measure—particularly if the "peripheral" features of the tests differ. It is therefore possible to differentiate between the hypothesised "latent variables" of the model and the "manifest variables" of each of the measures. Elegant SEM approaches to executive and attentional processes in adults are described in, for example, Burgess, Vetch, Costello, and Shallice (2000) and in Miyake et al. (in press). In a study that has only recently been drawn to our attention, Shapiro and colleagues (Shapiro, Morris, Morris, Flowers, & Jones, 1998) examined the patterns of children performing computerised tests of attention and general ability using SEM. In line with the predictions of this study, among 106 second-grade students they found that a unitary model of attention formed a poor fit to the observed variance, while a three-factor model of sustained attention, selective attention, and "central capacity" formed a significant and parsimonious fit.

The final approach that we describe here is to examine the performance of a clinical group on the TEA-Ch battery and other measures. Although the etiology and unitary nature of ADHD remains controversial, there is some convergence on frontal abnormalities that may be more pronounced within the right hemisphere (Barkley, 1997; Castellanos, & DuPaul, 1992b; Brumback & Staton, 1982; Castellanos et al., 1996; Filipek et al., 1997; Heilman, Voller, & Nadeau, 1991; Lou, Henriksen, & Bruhn, 1984; Swanson, Castellanos, Murias, LaHoste, & Kennedy, 1998; Voeller & Heilman, 1988). In adults, right prefrontal function has been associated with a number of functions including sustained attention (Cohen & Semple, 1988; Cohen, Semple, Gross, King, & Nordahl, 1992; Lewin et al., 1996; Pardo, Fox, & Raichle, 1991; Rueckert & Grafman, 1996; Wilkins et al., 1987) and response inhibition (Garavan, Ross, & Stein, 1999). Consistent with this view—and with links between children's and adults' attention—previous ADHD studies have emphasised poor performance in tasks of sustained attention (Barkley, 1997; Douglas, 1972; Hooks, Milich, & Lorch, 1994; Shue & Douglas, 1992) and response inhibition (Barkley, 1997; Barkley, Grodzinsky, & DuPaul, 1992a; Logan, Schachar, & Tannock, 1997; Swanson et al., 1998).

In the first study presented here we describe the patterns of performance seen in 293 healthy children on nine subtests of the novel TEA-Ch battery. We report the reliability of the assessments, the relationships between the TEA-Ch and existing tests of attention (in addition to IQ and scholastic attainment tests), and the relative fits of a unitary vs. theoretically derived three-factor SEM model of separable attention processes. In the second study we will examine the performance of 24 children with a diagnosis of ADHD.

**Study 1: Normative Data and Structural Equation Model**

**Method**

Participants. A total of 293 children between the ages of 6 and 16 were recruited from state schools in Melbourne, Australia. Equal numbers of boys and girls were tested in each of six age-bands, band 1 = 6–7 years, 2 = 7–9 years, 3 = 9–11 years, 4 = 11–13 years, 5 = 13–15 years, and 6 = 15–16 years (the second digit in each range reflecting higher cutoff point for the band). Exclusion criteria were: previous head injury or neurological illness; developmental delay or sensory loss; and collecting norms for a new battery we also had the opportunity to statistically test the patterns of convergence and divergence in performance against a broad model developed within the adult literature.

However, as we can only measure attention indirectly this is a somewhat ambitious aim. As children will vary in many abilities (motor skill, task comprehension, language, and so on) that may be shared between tests hypothesised to make different attention demands, our postulated attention factors would have to be expressed very strongly to be seen amid this "noise". Such problems of "task impurity" have been held to account for the generally low correlations observed between executive measures in adulthood (Burgess, 1997; Miyake et al., in press). As developing children may be expected to show greater variability along these nonattentional dimensions, the challenge may be much greater than with adults.

The first approach to meeting this challenge was in the design of the tests. We aimed to minimise the demands on memory\(^1\), reasoning, task comprehension, motor speed, verbal ability, and perceptual acuity while maintaining the demands on the targeted attentional system. To this end, for example, we controlled for motor speed on a visual search task by comparing performance under high and low attentional demands (see Methods section). Where possible, language was avoided in the stimuli or the required responses and the perceptual demands were reduced to detection rather than complex discrimination. Demonstrations and practice trials with correction were used to reduce the impact of task comprehension differences and to try and improve the reliability of performance.

Such attempts can only ever have limited success. The second approach to minimising the impact of "task impurity" was to use Structural Equation Modelling (SEM) in our analysis. This technique is related to the more conventional exploratory factor analysis method but has important differences. In conventional factor analysis, the model is determined in a "bottom-up" fashion from the data. SEM allows one to test an a priori specified model against the observed variance.

If one takes results from three very different tests that are theoretically designed to tap a common underlying process, then that process should be better represented by the covariance between the tests than the variance in any specific measure—particularly if the "peripheral" features of the tests differ. It is therefore possible to differentiate between the hypothesised "latent variables" of the model and the "manifest variables" of each of the measures. Elegant SEM approaches to executive and attentional processes in adults are described in, for example, Burgess, Vetch, Costello, and Shallice (2000) and in Miyake et al. (in press). In a study that has only recently been drawn to our attention, Shapiro and colleagues (Shapiro, Morris, Morris, Flowers, & Jones, 1998) examined the patterns of children performing computerised tests of attention and general ability using SEM. In line with the predictions of this study, among 106 second-grade students they found that a unitary model of attention formed a poor fit to the observed variance, while a three-factor model of sustained attention, selective attention, and "central capacity" formed a significant and parsimonious fit.

The final approach that we describe here is to examine the performance of a clinical group on the TEA-Ch battery and other measures. Although the etiology and unitary nature of ADHD remains controversial, there is some convergence on frontal abnormalities that may be more pronounced within the right hemisphere (Barkley, 1997; Castellanos, & DuPaul, 1992b; Brumback & Staton, 1982; Castellanos et al., 1996; Filipek et al., 1997; Heilman, Voller, & Nadeau, 1991; Lou, Henriksen, & Bruhn, 1984; Swanson, Castellanos, Murias, LaHoste, & Kennedy, 1998; Voeller & Heilman, 1988). In adults, right prefrontal function has been associated with a number of functions including sustained attention (Cohen & Semple, 1988; Cohen, Semple, Gross, King, & Nordahl, 1992; Lewin et al., 1996; Pardo, Fox, & Raichle, 1991; Rueckert & Grafman, 1996; Wilkins et al., 1987) and response inhibition (Garavan, Ross, & Stein, 1999). Consistent with this view—and with links between children's and adults' attention—previous ADHD studies have emphasised poor performance in tasks of sustained attention (Barkley, 1997; Douglas, 1972; Hooks, Milich, & Lorch, 1994; Shue & Douglas, 1992) and response inhibition (Barkley, 1997; Barkley, Grodzinsky, & DuPaul, 1992a; Logan, Schachar, & Tannock, 1997; Swanson et al., 1998).

In the first study presented here we describe the patterns of performance seen in 293 healthy children on nine subtests of the novel TEA-Ch battery. We report the reliability of the assessments, the relationships between the TEA-Ch and existing tests of attention (in addition to IQ and scholastic attainment tests), and the relative fits of a unitary vs. theoretically derived three-factor SEM model of separable attention processes. In the second study we will examine the performance of 24 children with a diagnosis of ADHD.

**Study 1: Normative Data and Structural Equation Model**

**Method**

Participants. A total of 293 children between the ages of 6 and 16 were recruited from state schools in Melbourne, Australia. Equal numbers of boys and girls were tested in each of six age-bands, band 1 = 6–7 years, 2 = 7–9 years, 3 = 9–11 years, 4 = 11–13 years, 5 = 13–15 years, and 6 = 15–16 years (the second digit in each range reflecting higher cutoff point for the band). Exclusion criteria were: previous head injury or neurological illness; developmental delay or sensory loss; referral for attentional or learning problems; assessment as having special educational needs.

The sex and age distribution of the sample are presented in Table 1 below. Socioeconomic status (SES) was determined for each child using the Daniel's Scale of Occupational Prestige (Daniel, 1983) on which parental occupation is rated between 1 and 7, with a score of 1 reflecting high and a score of 7 indicating low SES. The mean SES coding for the group was 4.04 (SD = 1.23, range = 1.7–6.6).

A subgroup of 16 children selected at random from the full sample (mean age = 10.80, SD = 2.83, range = 6.14–15.96) were also assessed on four subtests of the WISC-III and the Wide Ranging Achievement Test (Revised) (Jastak & Wilkin-
Table 1

<table>
<thead>
<tr>
<th>Age band</th>
<th>Total</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>6–7</td>
<td>38</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>7–9</td>
<td>56</td>
<td>30</td>
<td>26</td>
</tr>
<tr>
<td>9–11</td>
<td>54</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>11–13</td>
<td>58</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>13–15</td>
<td>58</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>15–16</td>
<td>29</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>293</td>
<td>147</td>
<td>146</td>
</tr>
</tbody>
</table>

* Upper cutoff of each band (e.g. only children under 7 years would appear in band 1).

Son, 1984). This group did not differ from the remainder of the sample in terms of age ($t = 0.67$, $p = .51$), sex distribution ($\chi^2 = 1.51$, $p = .22$) or SES ($t = 0.20$, $p = .84$).

A subgroup of 96 children also completed the colour-word Stroop, the Trails Test, and the Matching Familiar Figures Test (see measures below). The mean age of this subsample was 11.04 years ($SD = 2.92$, range = 6.14–15.96). Again this group was representative in terms of age ($t = -0.06$, $p = .95$) and sex distribution ($\chi^2 = 0.17$, $p = .68$). Although the subsample tended to be from slightly higher socioeconomic groups, the difference did not reach statistical significance (mean SES for subsample = 4.26, $SD = 1.29$; mean SES for remaining sample = 3.94, $SD = 1.20$, $t = -1.82$, $p = .07$; see below for influence of SES on TEA-Ch performance).

Measures

The Test of Everyday Attention for Children (TEA-Ch):

Sustained attention subtests.

(1) **Score!** Sustained attention requires the active maintenance of a particular response set under conditions of low environmental support (e.g. when there are few triggers to the relevant behaviour, when the task lacks interest or reward).

The Score! subset is a 10-item tone-counting measure based on a task originally described by Wilkins et al. (1987). In each item, between 9 and 11 identical tones of 345 ms are presented, separated by silent interstimulus intervals of variable duration (between 500 and 5000 ms). Children were asked to silently count the tones (without assistance from fingers) and to give the total at the end—as if they are “keeping the score by counting the scoring sounds in a computer game”.

If a child was unable to count to 15 or was unable to pass two practice trials (with relatively few tones) the test was not given. The requirement to pass practice items as a way of ensuring task comprehension, checking on possible sensory problems and improving the reliability of the measures, was a feature of each of the tasks. The duration of the test was approximately 5 min 40 s (some variability occurring due to the need to repeat instructions and so forth).

In their 1987 study, Wilkins and colleagues reported the sensitivity of an analogous task to focal right frontal lesions. Subsequent work with adults using the clinical version in the Test of Everyday Attention (Elevator Counting subtest—Robertson et al., 1994) has shown that right-hemisphere strokes disproportionately impair performance (Robertson, Manly, Beschin, et al., 1997). Similar measures have also produced predominantly right frontal activation in healthy adult participants in functional imaging studies (Lewin et al., 1996; Pardo et al., 1990).

(2) **Score DT.** Stuss, Shallice, Alexander, and Picton (1995) have argued that in sustained attention tests, the low level of exogenous activation for the goal relevant schema (listening to the test, counting etc.) made it vulnerable to possible competitors for processing resources. In conventional tone-counting or vigilance tasks, the nature of the competitors will vary with the individual and environment. The Score DT measure was designed to increase the sensitivity of the basic Score! task by including a “built-in” distractor. The tone-counting aspect of the task is identical to that of Score! described above. In addition, meaningful, auditory speech—in the form of news bulletins—was simultaneously presented.

Children were asked to keep a count of the tones whilst at the same time keeping “an ear out” for the mention of an animal during the news broadcast. They were asked to “concentrate most on the counting. If you concentrate too much on the news, the counting is very difficult”.

Two practice items were given before the 10 items of the test. Test duration was approximately 5 min 40 s.

(3) **Code Transmission.** Vigilance tasks, in which people are asked to monitor a stream of information for the occurrence of a rare target, have been predominantly used in the study of sustained attention in both normal and clinical populations (Koellea, 1996; N. H. Mackworth, 1948; J. F. Mackworth, 1970; Parasuraman, Warm, & See, 1998; Rosvold, Mirskey, Sarason, Bransome, & Beeck, 1956; See, 1995). Although a disproportionate decrement in performance over time is often considered the hallmark of a sustained attention deficit, the value of this measure in clinical groups has been called into question (Manly & Robertson, 1997; Stuss et al., 1989; Van Zomeren & Ven den Burg, 1985; Wilkins et al., 1987). In particular, it has been argued that such analysis may miss fluctuations in attention over much briefer time-scales (Stuss et al., 1989), and that, as performance on simple detection tasks becomes more routine or automated with practice, the demands on anterior attentional circuits can actually be reduced (Fisk & Schneider, 1981; Paus et al., 1997; Robertson, Manly, Andrade, et al., 1997). Findings of impaired “vigilance level” performance in adult brain-injured patients at any stage within such tasks have, however, been ubiquitous (Brouwer & Van Wolfelaar, 1985; Godefroy, Cabaret, & Rousseau, 1994; Loken, Thornton, Otto, & Long, 1995; Ponsford & Kinsella, 1992; Rueckert & Grafman, 1996; Spikman, Van Zomeren, & Deelman, 1996; Whyte, Polansky, Fleming, Branch-Coletti, & Cavallucci, 1995).

The Code Transmission subtest is an auditory vigilance-level measure. The children were asked to monitor a stream of monotonous digits (presented at a rate of one every 2 s) for the occurrence of a particular target sequence (e.g. 5 5) and then to report the digit that occurred immediately before. The target sequence was constant throughout the test.

Following a practice sequence to ensure comprehension, 40 targets were presented over the 12 min of the task.

(4) **Walk Don’t Walk.** This measure was adapted from the Sustained Attention to Response Test (SART: Robertson, Manly, Andrade, et al., 1997). The SART was designed to place emphasis on maintaining attention over one’s own actions and to avoid some of the problems of more automatic responding described above. The SART has proved sensitive to traumatic brain injury (Robertson, Manly, Andrade, et al., 1997) and to predict the reports of everyday “action lapses” in clinical and normal populations (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, et al., 1997). Although the task requires the periodic and unpredictable withholding of the routine response, behavioural and electrophysiological evidence suggest that the active maintenance of attention across the task is a strong determinant of the success of response suppression (Manly et al., 1999, 2000).

In the Walk Don’t Walk subtest, children are given an A4 sheet showing “paths” each made up of 14 squares. They are asked to listen to a tape that will play one sound (go tone) if the move to the next square should be made and another (no-go tone) if not. The moves were made by “dotting” each square with a marker pen, the pen being held approximately 2 cm above the page between each tone.

The go and no-go tones were identical for the first 208 ms (329.6 Hz sine tone), the no-go tone being marked by a concluding vocal exclamation (“D’oh!”). The task therefore ideally required children to listen to the entire sound before making their response. The go tones were presented in a regular,
rhythmic fashion with the no-go tone occurring unpredictably within the sequence (between the 2nd and 12th steps). Inter-tone intervals began at 1500 ms for item 1. Although held constant within each item, the intervals were systematically reduced with each new item, reaching a minimum of 300 ms at item 20.

Two demonstration trials and two practice trials were given before the test items. The dependent variable was the number of items correct out of 20. The total duration of the test items was approximately 6 min 16 s.

**TEA-Ch: Selective attention measures.**

The colour-word Stroop task has frequently been used as a paradigmatic definition of (nonspatial) selective attention, for example, in adult functional imaging studies (Bench et al., 1993; George et al., 1994; Pardo et al., 1990). In children as young as 6, where differences in reading ability may be considerable, this task may be less appropriate. Following the common loading with the Stroop task in the adult TEA factor analysis, we therefore adopted two nonlinguistic visual search tasks to represent this factor.

1. **Sky Search.** In this measure the children were given a laminated A3 sheet depicting rows of paired spacecraft (see Fig. 1). Four distinctive types of craft were presented, with most pairs being of mixed type. They were instructed to try and find all of the target items, defined by a pair of identical craft, as quickly as possible. Twenty targets were distributed among 108 distractors. Termination of the task was self-determined with the child marking a box in the lower left corner when they had finished. Both speed and accuracy were emphasised. Prior to completing the main test, children first completed a practice A4 sheet to ensure comprehension of target identity and the self-completion procedure.

In order to control for differences that are attributable to motor speed rather than visual selection, the children then completed a motor control version of the task. The A3 stimulus sheet was identical to that of the Sky Search test with the exception that all of the distractor items were removed (see Fig. 2). The task therefore consisted of circling all 20 target items as quickly as possible and then indicating completion.

Time taken to completion and accuracy were recorded for each part of the test. A time-per-target score was calculated (time/targets found). Subtraction of the "motor control" time-per-target from the more attentionally demanding Sky Search time-per-item produced an "attention" score that was relatively free from the influence of motor slowness or clumsiness. The time required for this test clearly depends upon the speed of an individual child. The mean time spent on the test item for this sample was 84.7 s ($SD = 39.25$) (see results for changes with age).

2. **Map Mission.** A potential confound within the Sky Search task is the influence of strategic differences due to the requirement to self-determine when the task is complete. A further "timed-out" visual search measure was therefore incorporated into the battery. In the Map Mission, the children were given a printed A3 laminated city map. Eighty targets (small restaurant knife-and-fork symbols measuring 4 mm × 3 mm) were randomly distributed across the map. Distracting symbols of a similar size (depicting supermarket trolleys, cups, and cars) were also present. The children were instructed to find and circle with a pen as many target symbols as possible within 1 minute—a period in which the detection/marking of all 80 targets would be extremely unlikely. The score was the number of targets correctly marked.

**TEA-Ch: Attentional control.**

1. **Creature Counting.** Switching from one task or mental set to another is almost invariably associated with a temporary slowing, which may reflect a transitory loss of efficiency (Rogers & Monsell, 1995). Clinically, the capacity to switch attention has often been tested using complex tasks such as the Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtiss,
Sky Search motor control stimulus sheet. The stimulus sheet is identical to the visual search version, save for the removal of distractors.

On each page of the Creature Counting stimulus booklet, a variable number of "creatures" were depicted in their burrow (see Fig. 3). Interspersed between the creatures were arrows either pointing up or down. The children were asked to begin counting the creatures from the top down but to use the arrows as a cue to switch the direction of their count. A correct response to the item shown in Fig. 3 below would therefore be, "1, 2 (up arrow) 3, 4, 5, 6 (down arrow) 5 4 (up arrow) 5 6 7 8".

The children's ability to count up to and down from 15 was assessed prior to beginning the test. Following two practice items on which feedback was given, the children completed seven test items. The accuracy of the response and the time taken to complete the page were recorded. If children scored 3 or more items correct, a timing score was calculated (seconds-per-switch) by dividing the time taken to complete correct items by the number of switches within those items. Again, the duration of this test varies with the speed of the child, but would usually be approximately 5 min.

(2) Opposite Worlds. Deficits in the inhibition of a verbal response have been noted in developmental disorders (Georgiou et al., 1996) and in adults with acquired brain lesions (Burgess & Shallice, 1996, 1997).

In investigating verbal inhibition in children, Passler, Isaac, and Hynd (1985) used a set of dark and light stimulus cards. Children were asked to respond "night" to the light card and
“day” to the dark card, this being viewed as a reversal of the most automatic or congruent response. The incongruent condition indeed took longer to complete than a congruent version. Using more explicit stimulus-verbal response associations, Gerstadt, Hong, and Diamond (1994), adapted this task using pictures of the moon and the sun. In the Opposite Worlds task, the aim was to make the association as explicit as possible by using the digits 1 and 2 as the stimuli and the words “one” and “two” as the response options.

In the task the children were presented with a stimulus sheet showing a mixed, quasi-random array of the digits 1 and 2 (see Fig. 4). In the “Sameworld” condition they were asked to read out the digits aloud as quickly as possible in the conventional manner. The purpose of the Sameworld condition was to reinforce the “prepotent” set of naming the numbers in the conventional manner in the context of the test materials, and also to identify any unexpected difficulties a child may experience with the task. In the “Oppositeworld” condition they were asked to say the opposite for each digit (“one” for 2 and “two” for 1) as quickly as possible, inhibiting the prepotent verbal response.

In the task, the examiner pointed to each digit in turn, only moving onto the next when a correct response was given, thus turning errors into a time penalty. Following practice in each condition, four test pages were run in the order; Sameworld, Oppositeworld, Oppositeworld, Sameworld. The time taken to complete each condition was recorded. Total time for the Oppositeworld condition was taken as the dependent variable. The average time spent on the test items in this study was 24.2 s (SD = 8.99) for the Sameworld and 30.67 (SD = 11.86) for the Oppositeworld.

**TEA-Ch: Dual task measure.**

(1) **Sky Search DT.** Performance decrements under dual task conditions tend to form sensitive measures of neurological impairment (e.g. Baddeley et al., 1991; Stuss et al., 1989). The TEA combined two of its subtests to form a dual task measure and this model was adopted here.

In the Sky Search DT test children were asked to complete a parallel version of the Sky Search Task (see Fig. 1), which differed only in the location of the targets. As they performed the visual search they were asked to simultaneously and silently count the number of tones presented within each item of an auditory counting task, giving the total at the conclusion of each item. Although the counting task used the same stimuli as the Score! subtest, a regular pacing of one tone per second was used. Following practice, the task and timing were initiated by a countdown played on the tape. The test was ended and timing stopped when the child indicated completion of the visual search component.

As it is possible that a child could completely neglect one of the tasks, scores from both measures were incorporated into a total score. Specifically the time taken to find each visual target was calculated (total time/correctly identified targets)—(a). The proportion of the counting items with correct totals was then calculated (total items correct/total items attempted)—(b). Poor counting performance was then used to inflate the time-per-target scores by dividing (a) by (b). Finally, in order to assess the decrement from single task visual search performance, the raw time-per-target score from the Sky Search task was subtracted from this value. To take an example, a child took 89 s to complete the task during which he found 19 targets. His time-per-target score was therefore 89/19 = 4.68. During this time he gave correct totals to three of the six counting items he was exposed to. His proportion correct score was therefore 3/6 = 0.5. Dividing the time-per-target score by this proportion inflates his time-per-target score to 4.68/0.5 = 9.36. In the original Sky Search test, his time-per-target score was 3.2 s. Subtracting this from the dual weighted time-per-target gives the decrement value, 9.36 - 3.2 = 6.16.

The time taken to complete this measure varies according to the child. In this sample the mean time to completion was 99.3 s (SD = 59.45).

In the factor analysis of the TEA, the analogous measure loaded onto the sustained attention factor. Consequently, it was predicted that a similar relationship would be observed in the children.

**Existing measures.**

(1) **The colour-word Stroop.** (Trenerry et al., 1989). In this paradigmatic measure of nonspatial selective attention, the children were first asked to name the colour of simple colour

**Figure 4.** Item from the Opposite Worlds subtest.
patches. They were then asked to read a list of colour words, in the condition of interest they were asked to name the colour of conflict. The measure was the number correct within a 40-s period.

(2) Trails Test (Spreen & Strauss, 1991). This task is thought to tap selective attention/visual search and the capacity to switch attention. This measure has two parts. In Trails A, children were asked to draw lines between circles randomly arranged on a page. Each circle contained a number and the connection of circles should follow a rule of ascending number sequence. In part B of the test, children were asked to join circles based on an alternating rule of ascending numbers and letters in alphabetical sequence. The time taken to complete each measure and the number of errors made were recorded.

(3) Matching Familiar Figures Test (MFTT). (Arizmendi, Paulsen, & Domino, 1981). In this test, considered a measure of impulsivity, the children were asked to match a single stimulus with one of six similar looking pictures. A speed-accuracy trade-off is generally observed whereby fast, “impulsive” responding tends to produce more errors.

(4) Wechsler Intelligence Scale for Children, 3rd edition. (WISC-III; Wechsler, 1991). Four subtests from this widely used measure of general intellectual function were used; two were measures of verbal knowledge and reasoning (Vocabulary and Similarities) and two were measures of speeded visuospatial and motor function (Block Design and Object Assembly).


Results

Age and sex effects within the TEA-Ch measures. The influence of age on performance of each of the TEA-Ch variables was initially analysed using correlation across the group as a whole (boys and girls). Unsurprisingly, age exerted a significant effect for each measure (Pearson's r in each case being; Sky Search .58, Score! .54, Creature Counting (accuracy) .27, Creature Counting (timing) .63, Sky Search DT .39, Map Mission .75, Score DT .69, Walk Don't Walk .50, Opposite Worlds .73, and Code Transmission .55), the p value in each case being below <.001 needed for corrected statistical significance. As can be seen in Table 2, the absolute difference in performance between each age band tends to diminish in the older groups, suggestive of a developmental plateau and/or, for some measures, the emergence of ceiling effects (see below).

The TEA-Ch was designed to be appropriate for both 6-year-olds and 16-year-olds. As can be seen in Table 2, the first aim of avoiding floor effects in the youngest group was achieved to a considerable degree. No child achieved a zero score on the Score!, Score DT, Code Transmission, or Walk Don’t Walk subtests. Similarly, no child failed to find any of the targets in the Sky Search or Map Tasks. In itself, this suggests that the demonstration and practice items were successful in explaining the tasks and that the perceptual demands of the tasks were relatively insensitive to normal variation in vision and hearing. Only on the Creature Counting subtest were complete failures to score observed (seven children).

Whereas the time-based measures clearly show variation throughout the age range, ceiling levels of performance in older children were observed on some of the item-based measures. Of the adolescents over the age of 15, for example, 72% scored at ceiling on the Score! Subtest—a measure that similarly attracts ceiling performance in most neurologically healthy adults (Robertson et al., 1996). Ceiling levels of performance were less frequently observed in adolescents over the age of 15 on the Score DT, Code Transmission, Walk Don’t Walk, Creature Counting, and Map Mission tests (24%, 37%, 10%, 34%, and 0% respectively).

The effects of sex for the entire sample were examined for each measure (raw scores) using ANOVA. There was no significant difference between the performance of the
147 girls and the 146 boys on any task with the exception of the Creature Counting test (timing score), where the boys performed moderately better than the girls (boys mean switch time = 3.95 s, SD = 1.78; girls mean switch time = 4.48 s, SD = 1.89; F = 5.91, p < .05).

To consider differences between boys and girls at different ages, raw scores were compared for each age band separately. Only one task, the Sky Search visual search measure, produced significant differences, with girls outperforming boys in the age bands 9–11 and 13–15 (F = 6.8, p < .05 and F = 5.6, p < .05, respectively). The results suggest, therefore, that the TEA-Ch is assessing abilities that are not well tapped by (at least these) measures of general ability—and that additional assessment with such measures is far from redundant.

### Table 3

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Pearson's correlation coefficient</th>
<th>Correlation with age partialled out/% agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score!</td>
<td>.64***</td>
<td>76.2%</td>
</tr>
<tr>
<td>Score DT</td>
<td>.74***</td>
<td>71.4%</td>
</tr>
<tr>
<td>Code Transmission</td>
<td>.82***</td>
<td>.78</td>
</tr>
<tr>
<td>Walk Don't Walk</td>
<td>.73***</td>
<td>71.0%</td>
</tr>
<tr>
<td>Sky Search</td>
<td>.90***</td>
<td>.75</td>
</tr>
<tr>
<td>Map Mission</td>
<td>.88***</td>
<td>.65</td>
</tr>
<tr>
<td>Creature Counting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>.69***</td>
<td>.71</td>
</tr>
<tr>
<td>Timing</td>
<td>.73***</td>
<td>.57</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>.92***</td>
<td>.85</td>
</tr>
<tr>
<td>Dual Task Decrement</td>
<td>.81***</td>
<td>.81</td>
</tr>
</tbody>
</table>

**p < .001.

### Table 4

<table>
<thead>
<tr>
<th>TEA-Ch subtest</th>
<th>IQ</th>
<th>Vocabulary</th>
<th>Similarieties</th>
<th>Block Design</th>
<th>Object Assembly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sky Search</td>
<td>.14</td>
<td>.05</td>
<td>.05</td>
<td>.15</td>
<td>.19*</td>
</tr>
<tr>
<td>Score!</td>
<td>.14</td>
<td>.14</td>
<td>.10</td>
<td>.12</td>
<td>.09</td>
</tr>
<tr>
<td>Sky Search DT</td>
<td>.14</td>
<td>.12</td>
<td>.07</td>
<td>.12</td>
<td>.16*</td>
</tr>
<tr>
<td>Creature Counting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>.31**</td>
<td>.23**</td>
<td>.20*</td>
<td>.27**</td>
<td>.30**</td>
</tr>
<tr>
<td>Speed</td>
<td>-.01</td>
<td>.12</td>
<td>.07</td>
<td>.01</td>
<td>-.05</td>
</tr>
<tr>
<td>Map Mission</td>
<td>.25**</td>
<td>.15</td>
<td>.16*</td>
<td>.24**</td>
<td>.27**</td>
</tr>
<tr>
<td>Score DT</td>
<td>.15</td>
<td>.10</td>
<td>.14</td>
<td>.09</td>
<td>.16</td>
</tr>
<tr>
<td>Walk Don't Walk</td>
<td>.21*</td>
<td>.14</td>
<td>.05</td>
<td>.24*</td>
<td>.24*</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>.07</td>
<td>-.002</td>
<td>.03</td>
<td>.09</td>
<td>.09</td>
</tr>
<tr>
<td>Code Transmission</td>
<td>.17*</td>
<td>.16*</td>
<td>.08</td>
<td>.18*</td>
<td>.11</td>
</tr>
</tbody>
</table>

aControlled score.
bDual task decrement.
*p < .05; **p < .01 (uncorrected significance levels: p < .001 required for statistical significance when full correction for multiple comparisons is used).
made using partial correlations (age partialled out) on raw scores. The results are presented in Table 5.

Although correction for multiple correlations means that only $p$ values of less than .001 should be strictly taken as statistically significant, a number of patterns emerge that nevertheless offer some support to our characterisation of the primary attentional demands of each subtest. The Sky Search and Map Mission tasks, for example, were characterised as primarily selective attention tests. Both correlated with the superficially very different Stroop measure ($r = .40, p < .001$ and $r = .31, p < .01$ respectively)—often taken as a paradigmatic definition of selective attention processes. This relationship is unlikely to be primarily mediated by speed alone, as the speed component in the Sky Search score was greatly reduced by the motor control manipulation (see Measures). In addition, the Stroop was not significantly related to a number of other speeded measures in the TEA-Ch (Sky Search DT, Creature Counting, Opposite Worlds) nor, indeed, to any sustained attention measure. The Sky Search was not significantly related to a number of other speeded measures in the TEA-Ch battery (Sky Search DT, Creature Counting, Opposite Worlds) nor, indeed, to any sustained attention measure. The Sky Search was not significantly related to a number of other speeded measures in the TEA-Ch battery (Sky Search DT, Creature Counting, Opposite Worlds) nor, indeed, to any sustained attention measure. The Sky Search was not significantly related to a number of other speeded measures in the TEA-Ch battery (Sky Search DT, Creature Counting, Opposite Worlds) nor, indeed, to any sustained attention measure.

It is of note that the Matching Familiar Figures Test, often taken as a measure of behavioural inhibition, had rather low relationships with the Walk Don’t Walk and Opposite Worlds subtests, which feature a requirement to withhold a prepotent motor or verbal response. The relationships were certainly no greater than with other measures without such an obvious requirement, such as the Score! or Score DT tests.

Fit of the adult model of attentional separability to the data: Structural Equation Models. In order to produce a clinical measure, the raw scores of the TEA-Ch subtests were transformed to age-scaled scores (following convention, e.g. Wechsler measures), these were scaled to a mean of 10 and a SD of 3; range 1–19. This was achieved by first analysing each subtest using a one-way ANOVA (for boys and girls separately) with the age bands as factors after applying a Box-Cox (power) transform combined with a linear transform. The ordered residuals were then plotted against the corresponding quantile of a standard normal distribution. The best transform for each subtest was selected by the criterion of minimising the nonlinearity of the regression. The appropriate quantiles were then estimated by application of the inverse power transform to the corresponding normal distributions. This transformation therefore reflects the relationship of an individual’s raw score to the mean and distribution of their age band (see Yandell, 1997, Chapter 13).

In terms of modelling the data, this transformation offers further advantages in normalising the scores and effectively removing the influence of age.

If the TEA-Ch measures make demands on attention, and attention is best thought of as a unitary factor, then the most parsimonious model in accounting for the variance would be a single latent variable. This was investigated and found not to provide an adequate fit to the data, as indicated by the significant $\chi^2$ value, $\chi^2(27) = 95.05, p < .001$.

We therefore investigated whether our broad model based on the adult literature formed a useful fit to the patterns of test performance seen in the children. Each of the variables from the TEA-Ch battery was ascribed to a latent attentional variable based on their theoretical origins (see Measures). Score! Score DT, Code Transmission, and Walk Don’t Walk were ascribed to “sustained attention” (with the Dual Task Decrement measures also being associated with this factor given the results of Robertson et al., 1996). Map Mission and Sky Search were associated with the “selective attention” factor and Creature Counting and Opposite Worlds associated with a broad “executive” or “attentional control” factor.

The scaled scores from the variables and the three “latent” factors were entered into the EQS structural equation modelling software package (Bentler & Wu, 1995). The patterns observed were consistent with this model as indicated by a nonsignificant $\chi^2$ statistic, $\chi^2(224) = 33.431, p = .10$. In addition, and particularly for large samples, Bentler recommends the use of three incremental fit measures; the Comparative Fit Index (CFI), the Normed Fit Index (NFI) and the Non-Normed Fit Index.
These indices represent the improvement in fit between the a priori model and a baseline model of uncorrelatedness between the observed variables. Fit index values of .9 and above are deemed indicative of a good fit to the data (Hair, Anderson, Tatham, & Black, 1995). The values of these three diagnostics for the model were all above this threshold (CFI .973, NFI .913, and NNFI .960), indicating that these three factors alone form a good fit of the patterns of performance observed in a large group of children. The EQS model is represented graphically in Fig. 5 below.

The normalisation of the test data by the transformation to standard scores makes it very unlikely that the observed SEM fit could occur through psychometric properties of the tasks, such as range or ceiling effects, or through test order. The distribution of skew values (the length of the distribution tail) and kurtosis (the grouping of cases within a particular band) across the factors are shown in Table 6 below. (The dual task measure has high values due to the imposition of cutoffs. The use of division in calculating scores can otherwise lead to too great a range.)

A further important question in terms of relating adult models of attention to children is whether the observed fit of the model occurs primarily due to the older children in the sample. The children were therefore divided into two large groups around the median age (Group 1, 6.01–

Table 6

<table>
<thead>
<tr>
<th>Measure</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustained attention tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score!</td>
<td>1.43</td>
<td>1.43</td>
</tr>
<tr>
<td>Score DT</td>
<td>0.82</td>
<td>0.05</td>
</tr>
<tr>
<td>Code Transmission</td>
<td>2.11</td>
<td>5.81</td>
</tr>
<tr>
<td>Walk Don’t Walk</td>
<td>0.79</td>
<td>0.29</td>
</tr>
<tr>
<td>Sky Search DT</td>
<td>5.99</td>
<td>50.54</td>
</tr>
<tr>
<td>Selective attention tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sky Search</td>
<td>2.24</td>
<td>2.24</td>
</tr>
<tr>
<td>Map Mission</td>
<td>0.09</td>
<td>0.91</td>
</tr>
<tr>
<td>Attention control tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creature Counting</td>
<td>1.61</td>
<td>3.25</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>1.66</td>
<td>3.19</td>
</tr>
</tbody>
</table>

10.91, N = 146; Group 2, 10.92–15.98, N = 147) and the strengths of the association between test scores and the latent variables compared. No significant differences emerged, \( \chi^2 (9) = 7.35, p > .6 \). The results therefore support the broad division of attention into separable systems in children as young as 6.

Relationship to SES and scholastic attainment measures.

SES, as indexed using Daniel’s (1983) system, formed a
poor predictor of TEA-Ch performance. No relationship reached a statistical significance corrected for multiple correlations and only one (Creature Counting) reached uncorrected significance ($r = -0.16, p < 0.05$).

The correlations between TEA-Ch measures and the Wide Ranging Achievement Test-Revised Reading, Spelling and Arithmetic scales (Jastak & Wilkinson, 1984) are presented in Table 7. Again, no relationship meets statistical significance once correction for multiple correlations is applied. It is perhaps of note, however, that the four measures of sustained attention show the strongest relationships with the attainment scores.

The results from this study are discussed after the clinical study described below.

**Study 2: The Performance of Boys with an ADHD Diagnosis**

**Method**

**Participants.** Twenty-four boys meeting DSM-IV criteria for Attention Deficit Hyperactivity Disorder were tested on subtests of the TEA-Ch and WISC-III measures. The mean age of the sample was 9.95 (SD = 2.23). The boys were recruited from referrals made to an Outpatient Child and Family Centre. Diagnoses, according to DSM-IV criteria, were made by a Consultant Psychiatrist and Clinical Psychologist based on Conners rating scales (Conners, Sitarenios, Parker, & Epstein, 1998) supplied by the schools and two 1.5-hour assessment sessions. Exclusions from the current study were made if there was (a) a diagnosis of comorbid clinical or psychiatric disorder; (b) peripheral sensory loss, physical disability, epilepsy, or unequivocal brain damage; (c) severe adversity and/or involvement with Child Protection Agencies. The children were seen prior to the prescription of any medication.

All but one of the children attended mainstream school (the exception attended a Special Needs School due to behavioural, rather than cognitive, difficulties). None of the group was identified as having specific learning needs, or indeed was seen as warranting a Statement of Educational Needs.

**Measures.** The boys were administered the Score! Reading, Score DT, Walk Don't Walk, Sky Search, Sky Search DT, and Opposite Worlds subtests of the TEA-Ch (use of the remaining measures Code Transmission, Map Mission, and Creature Counting was not practicable within the time available). A two-test short-form of the WISC-III (Sattler, 1988; Wechsler, 1991), comprising Vocabulary and Block Design subtests, was also administered.

<table>
<thead>
<tr>
<th>TEA-Ch subtest</th>
<th>F value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score! Reading</td>
<td>(1, 158) = 17.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age-scaled score</td>
<td>(1, 158) = 14.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Score DT</td>
<td>(1, 157) = 88.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age-scaled score</td>
<td>(1, 157) = 45.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Walk Don't Walk</td>
<td>(1, 158) = 66.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age-scaled score</td>
<td>(1, 158) = 50.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sky Search DT</td>
<td>(1, 157) = 14.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Raw decrement score</td>
<td>(1, 158) = 5.65</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Sky Search</td>
<td>(1, 158) = 0.39</td>
<td>.983 ns.</td>
</tr>
<tr>
<td>Age-scaled score</td>
<td>(1, 158) = 1.68</td>
<td>.907</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>(1, 158) = 23.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Raw score</td>
<td>(1, 158) = 21.8</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*p values < .001 required to reach statistical significance with correction for multiple comparisons.

**Procedure.** Neuropsychological assessment was completed 2 weeks before any medical treatment was commenced. The same experimenter saw each participant alone. Testing took place at home or at the clinic, and was completed at one sitting. Test administration was completed according to the standardised procedures described above.

**Results**

**Basic comparison with the normative sample.** Analyses of covariance (ANCOVAs), comparing the ADHD boys' results with those of boys from the normative sample (with age covaried), were performed on raw scores and the age-adjusted scaled scores. The F statistics are shown with the corresponding $\alpha$-value in Table 8 below. The results clearly show that the ADHD boys were significantly poorer than control boys at performing all of the TEA-Ch measures with the exception of the Sky Search subtest (although if formal correction for multiple comparisons is adopted, the Sky Search DT measure would also fail to meet corrected significance levels).

Although this comparison indicates that the ADHD boys performed the measures more poorly than controls, it tells us little about whether attention skills were particularly poor in this group. The boys, for example, also performed poorly on other measures with less obvious demands on attention, such as the Vocabulary subtest of the WISC-III. To examine whether attention represents a disproportionate problem in ADHD, children were selected from the TEA-Ch sample who were matched to the ADHD group on age and level of performance on the WISC-III Vocabulary subtest. It is important to note that only age and Vocabulary scores were used in the selection of controls—indeed, given the poor performance of the ADHD boys on this measure, relatively few children from the sample met these conditions ($N = 15$) (see Table 9).

As can be seen in Table 9, although the groups had almost identical performance on the Vocabulary measure, the ADHD children continued to show significantly poorer attention skills on many of the TEA-Ch measures. The Sky Search DT test, which had previously reached uncorrected significance, no longer had this
Table 9
Comparison of 24 ADHD-Diagnosed Boys with 15 Age and WISC-III Vocabulary Matched Controls

<table>
<thead>
<tr>
<th></th>
<th>Vocabulary controls</th>
<th>ADHD group</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>10.46 (2.85)</td>
<td>9.95 (2.23)</td>
<td>n.s.</td>
</tr>
<tr>
<td>WISC-III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary score</td>
<td>5.53 (2.13)</td>
<td>5.42 (2.78)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Block Design</td>
<td>9.53 (2.80)</td>
<td>6.96 (3.06)</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>TEA-Ch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score!</td>
<td>10.47 (3.09)</td>
<td>7.08 (3.66)</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Score DT</td>
<td>10.13 (3.68)</td>
<td>4.96 (2.79)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Walk Don’t Walk</td>
<td>9.33 (3.31)</td>
<td>4.79 (2.86)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Sky Search DT</td>
<td>7.87 (5.36)</td>
<td>6.04 (3.85)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Sky Search</td>
<td>8.93 (3.22)</td>
<td>9.67 (4.00)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>9.73 (3.65)</td>
<td>6.75 (4.21)</td>
<td>&lt; .05</td>
</tr>
</tbody>
</table>

p values < .001 required to reach statistical significance with correction for multiple comparisons.

Table 10
Performance of a Group of 24 ADHD Boys Compared with Control Children Matched on WISC-III Block Design Subtest Scores and Age

<table>
<thead>
<tr>
<th></th>
<th>Block Design controls</th>
<th>ADHD group</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>9.99 (2.77)</td>
<td>9.95 (2.23)</td>
<td>n.s.</td>
</tr>
<tr>
<td>WISC-III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>7.09 (1.54)</td>
<td>6.96 (3.06)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>9.41 (4.09)</td>
<td>5.42 (2.78)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>TEA-Ch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score!</td>
<td>9.05 (2.77)</td>
<td>7.08 (3.66)</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Score DT</td>
<td>9.36 (3.19)</td>
<td>4.96 (2.79)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Walk Don’t Walk</td>
<td>8.76 (3.35)</td>
<td>4.79 (2.86)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Sky Search DT</td>
<td>7.33 (3.45)</td>
<td>5.46 (4.27)</td>
<td>.116</td>
</tr>
<tr>
<td>Sky Search</td>
<td>10.09 (2.76)</td>
<td>9.67 (4.00)</td>
<td>.08</td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>9.19 (2.66)</td>
<td>6.75 (4.21)</td>
<td>&lt; .027</td>
</tr>
</tbody>
</table>

p values < .001 required to reach statistical significance with correction for multiple comparisons.

capacity. The difference between the groups of the Opposite Worlds measure would now also disappear if full correction for multiple comparisons is adopted.

As a final, and more conservative test of the specificity of attentional dysfunction within the group, the ADHD boys were matched with controls on the basis of age and WISC-III Block Design scores. This is a conservative comparison because, as with a number of measures from the TEA-Ch, the Block Design test requires a fast performance and, in common with some aspects of attention, is thought to reflect efficiency of predominantly right hemisphere functioning (Lezak, 1995).

Twenty-two children from the normative sample met the age and Block Design performance level criteria. The comparisons on TEA-Ch subtests, presented in Table 10, show significant and disproportionate deficits on three of the sustained attention factor measures (Score!, Score DT, and Walk Don’t Walk) and on the attentional control measure, Opposite Worlds. If corrected for multiple comparisons, however, only the Score DT and Walk Don’t Walk measures, both from the sustained attention factor, would continue to reach statistical significance.

Discussion

The subtests of the TEA-Ch are not measures of attention. They are measures of auditory and visual detection, of counting, of response speed, and so forth. Separable attention processes are inferred constructs believed to contribute significantly to differences in the efficiency of performance on these tasks. By simplifying instructions, using practice sessions, and reducing the contribution of perception, memory, and reasoning, our aim was to minimise variability due to nonattentional factors.

The results from 293 children from the normal school-age population give some grounds to believe that these aims were met to a reasonable degree. For all but one of the tasks, all of the children—even as young as 6—were able to perform at least one item correctly. Taken together with the mean levels of performance (shown in Table 2), this indicates that the basic comprehension and perceptual demands were met. In the case of the Score! subtest, for example, the children’s task was to maintain a count of the number of identical sounds that they heard within the period of each item. Over 90% of the sample gave correct totals for 5 or more of the 10 items. As a basic capacity to count up to 15 was assessed prior to the task, and the sounds must have been detectable, errors would seem to stem from a difficulty in voluntarily maintaining “one’s mind” on this very tedious activity—in other words, sustained attention (cf. Wilkins et al., 1987).

The nonsignificant relationships between the TEA-Ch
measures and conventional WISC-III IQ tasks (at least within the range we assessed) provide additional support for the relatively focused nature of the tasks and for the value of separate assessment of attentive functions.

Tests of executive function in adults (particularly those that stress novelty) can be inherently unreliable (Burgess, 1997; Wilson et al., 1997)—indeed, it has been argued that frontal/dysexecutive difficulties can themselves be thought of as a deficit in the reliable application of strategy, attention, or behaviour control (Rabbitt, 1997). In this sense, the poorest level of performance seen in a child—whether or not it can be repeatedly observed—may be informative. However, given the influence of factors with day-to-day variation (noise, mood, other preoccupations and so forth), clinically it is useful to have a sense of the stability of a measure. In this respect the predictive value of TEA-Ch scores from one session to another are surprisingly good although, given the rapid development of abilities seen in the younger children, such reliability can only provide a short-range forecast.

Given that underlying attentional processes do contribute to performance on the measures, it is possible to ask whether this looks like the footprint of a single system (say, efficiency of “top-down” control) or, as suggested by Posner and Petersen (1990) among others, like the influence of different systems performing different types of attention process.

Using the SEM approach, the single system model did not form a good fit to the data. Using a broad three-way division into sustained and selective attention and executive or attentional control based on adult models, however, did.

The results showed that the fit of this model was not significantly different for the older and the younger children. This suggests that despite the relative immaturity of these capacities, the differential nature of the processes is still sufficiently expressed to be detected across the tasks.

It remains important to see whether this fit emerged from “peripheral” aspects of the tasks that are of little relevance to the postulated attentional factors. It is certainly true, for example, that the sustained attention tests tended to be based on summed item correct scores and (consequently) are more vulnerable to ceiling effects than the selective attention tasks. The use of the age-scaling technique (on which the model is based), however, considerably reduced the impact of this factor and the skew and kurtosis of the sustained attention scores were unremarkable.

Other factors such as the requirement for speeded responses, or linguistic responses, can also be excluded to a degree. The use of the “motor control” task in the Sky Search test, for example, resulted in the speed contributions to the two selective attention tasks being rather different. Verbal responses were required in the sustained attention and attentional control measures. Further supportive evidence emerges from the relationships between the TEA-Ch and other measures of attention. As with healthy adults (Robertson et al., 1996), the colour-word Stroop showed the strongest relationship to measures of selective attention with very different “surface” characteristics—and relatively poor relationships to sustained attention measures.

The central aim of clinical assessment is to predict and account for difficulties that may arise within everyday life. Perhaps most particularly for attention and executive functions, the conceptual gulf between tests and the real world can be substantial. In the standard assessment setting, the examiner acts to focus the child’s attention on the material, provides clear instructions as to what is expected, is suitably encouraging, and attempts to protect the environment from external distractors. These features, together with the relative novelty of the materials, are important in making a reasonable comparison between different children. It is important to ask, however, whether a child’s performance under these conditions has any bearing on their behaviour in a crowded, noisy classroom, in the cacophonous home, or when the information that is relevant and the nature of the goal are not clearly specified? Clearly the best way of assessing whether such difficulties occur is to ask teachers, parents, or the children themselves. In accounting for such problems, however, consideration of particular cognitive deficits, in conjunction with emotional, social, and environmental factors, may be crucial in selecting the most appropriate form of support.

One group of children who, by definition, are reported to experience difficulties in these everyday situations are those diagnosed with ADHD. In the second study, the performance of such children on the TEA-Ch measures under standard testing conditions was examined.

Twenty-four boys with a DSM-IV diagnosis of ADHD (who were not yet prescribed medication) showed significant deficits across sustained attention and attentional control subtests of the TEA-Ch but, notably, no deficit on the speeded-visual search task. Given the poor performance of the group on a range of measures, including vocabulary, this comparison does little to indicate the specificity of attentional dysfunction within the group. More conservative comparisons with children who were matched on both age and performance levels on WISC-III subtests revealed persistent deficits. The results are consistent with previous findings that have emphasised both a sustained attention deficit and a difficulty in the suppression of prepotent responses (Barkley, 1997; Barkley et al., 1992b; Douglas, 1972; Hooks et al., 1994; Logan et al., 1997; Shue & Douglas, 1992; Swanson et al., 1998). They are also consistent with views of a neurological basis to the ADHD disorder—which have identified abnormalities including within right frontal systems.

There are important questions about the representativeness of our ADHD sample. In particular, although comorbid diagnoses are very common in ADHD, children with comorbid diagnosis were excluded from our study. Our aim, however, without wishing to over-reify the diagnostic category, was to examine the “pure” effects of ADHD as a basis for subsequent evaluation of the effect of comorbid and other factors. Whether such children would show similar profiles on the TEA-Ch remains an open question. A second important issue in this respect concerns the rather poor performance of the group on WISC-III measures, compared with a number of other reported ADHD samples. Although our comparison indicates that specific deficits in attention are still apparent when we control for WISC-III performance levels, further work is again required to examine the pattern in ADHD children who enjoy higher levels of general performance.

Anderson and colleagues (Anderson, Fenwick, Manly, & Robertson, 1998) have previously reported on the TEA-Ch performance of children who had sustained closed head injuries an average of 6 years prior to the
assessment. Although closed head injuries disproportionately compromise frontal/temporal functions, damage to a wide range of cortical and subcortical structures is possible (Mattson & Levin, 1990). As a group, these children showed deficits across the range of TEA-Ch measures including the selective attention factor relative to a control group (matched on both age and WISC-III Block Design subtest performance). Although further work is clearly required to look at a range of disorders, these early studies indicate the value of differentially assessing attentional functions. Perhaps most importantly, in adults the assessment of separable attentional function has led to improved targeting and effectiveness in rehabilitative interventions (e.g. Robertson, Tegner, Tham, Lo, & Nimmo-Smith, 1995; Sturm, Willmess, Orgass, & Hartje, 1997). It is to be hoped that the extension of such measures for children could prove similarly effective—both in rehabilitation and in assessing the specific effects of pharmocological interventions.

Acknowledgements—We are grateful to the University of Melbourne and the U.K. Medical Research Council who supported this research. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study. Our thanks also to Mick Wilson, Sarah Jones, and the Thames Valley Test Company for their assistance in developing the measure and for permission to use examples of stimuli from the test in Figures 1, 4. We are indebted to Rani Jacobs, Tim Godber, and John Gora for their help with the study.

References


Differential Assessment of Attention


Manuscript accepted 7 August 2001