Academic Outcomes following a school-based RCT for ADHD: 6 year follow-up

Kapil Sayal,1,2* Christine Merrell,3 Peter Tymms3 & Adetayo Kasim4

1 Division of Psychiatry & Applied Psychology, School of Medicine, University of Nottingham
2 Centre for ADHD and Neurodevelopmental Disorders Across the Lifespan (CANDAL), Institute of Mental Health, Nottingham
3 Centre for Evaluation and Monitoring (CEM), Durham University
4 Wolfson Research Institute for Health and Wellbeing, Durham University

Correspondence:
Kapil Sayal
Division of Psychiatry & Applied Psychology, University of Nottingham
E Floor, South Block,
Queen's Medical Centre
Nottingham, NG7 2UH, UK
Email: kapil.sayal@nottingham.ac.uk
Tel: +44 1158230264
Fax: +44 1158230270
ABSTRACT

Objective
For children with high levels of ADHD symptoms, to investigate the impact of early school-based interventions on academic outcomes in mid-childhood.

Methods
A 6 year follow-up of 4-5 year olds (n=52,075) whose schools participated in a cluster randomized controlled trial for children at risk of ADHD. School-level interventions involved the provision of a booklet with evidence-based information (book) and/or feedback of names (identification) of children with high levels of ADHD symptoms. At ages 10-11 years, outcome measures were scores in English and mathematics tests.

Results
For children with high levels of ADHD symptoms, the interventions had no impact on academic outcomes. When all children were analyzed, the book intervention had a positive impact on mathematics. Baseline inattention was associated with poorer academic outcomes, whereas impulsiveness was associated with better academic outcomes.

Conclusions
The provision of evidence-based information about helping children with ADHD at school may have wider academic benefits.

Keywords: ADHD, schools, interventions, follow-up
Attention-deficit/hyperactivity disorder (ADHD) affects around 5% of school-aged children world-wide (Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007), although higher rates are reported in some parts of the US (Rowland et al., 2002). In addition, many children who have high levels of inattention, hyperactivity and/or impulsiveness symptoms may not reach threshold to meet full diagnostic criteria (Willoughby, 2003). These children are also at risk of adverse outcomes in adolescence and adulthood (Moya, Stringaris, Asherson, Sandberg, & Taylor, 2014; Polderman, Boomsma, Bartels, Verhulst, & Huizink, 2010; Washbrook, Propper, & Sayal, 2013). As ADHD is a neuro-developmental disorder with onset at a young age, early interventions that aim to reduce the likelihood of later problems are of interest (Charach et al., 2013; Coates, Taylor, & Sayal, 2014). Potentially, interventions delivered through schools provide an approach to optimize access and encompass the majority of at-risk children (DuPaul & Eckert, 1997).

This paper reports on a long-term follow-up of a large cluster randomized controlled trial (RCT) of school-based interventions for ADHD. The rationale for the initial RCT reflected uncertainty (equipoise) about the potential for beneficial impacts. Although it has been argued that intensive interventions are needed to change classroom practices (Joyce & Showers, 2002; Adey, Hewitt, Hewitt, & Landau, 2004), lighter touch interventions such as providing basic information to teachers might have small effects at the individual level but greater impact at a population level. Major programmes of different types of in-service work need to be tested to see whether these can make differences to schools and pupils. The provision of research-based advice for teachers might help any teacher who wanted to improve their own teaching
and even if a proportion of teachers followed part of the advice it can be expected to enhance children's learning.

Specifically, this follow-up study investigates the impact of interventions delivered during school Year 1 (children aged 5 years) on the academic outcomes at the end of school Year 6 (children aged 10-11 years) of children attending primary (elementary) schools in England. Given that the initial trial and a follow-up relying on parent ratings have respectively suggested the possibility of some adverse academic outcomes at age 7 (Tymms & Merrell, 2006) and behavioral outcomes at age 10-11 (Sayal et al., 2010), we aimed to assess whether there were positive or negative effects of the interventions on academic outcomes at the age of 10-11 years. We also assessed whether the interventions had any wider impact on academic outcomes for all children, regardless of their levels of ADHD symptoms at baseline.

**METHOD**

**Baseline RCT & Interventions**

Full details of the baseline study and the two-year follow-up are described elsewhere (Tymms & Merrell, 2006) and briefly summarised here. Ethical approvals for the baseline and follow-up studies (including the follow-up reported here) were received from the Research Ethics Committee (IRBs) at Durham University. The baseline sample consisted of 73,367 children from 2040 primary (elementary) schools. At baseline, the teachers in the Reception year (when children were aged 4-5 years) completed a behavior rating scale consisting of the 18 DSM-IV ADHD items on all children in their class just before the end of the academic year (Merrell & Tymms, 2001). The items were rated on a yes/no scale and teachers were asked to consider a
Criterion was only met if the behavior was persistent and frequent. To assess the stability of these ratings, a random sample of children was re-assessed by their new class teachers eighteen months after the baseline assessment. The correlation between the two sets of teachers’ ratings was 0.64 (Merrell and Tymms, 2001). To investigate the validity of the rating scale, a random sample of children was also re-assessed with a Continuous Performance Test (CPT) eighteen months after the baseline assessment. The correlation between the baseline teachers’ ratings and the overall score on the CPT test was 0.7 (Merrell and Tymms, 2001). Similarly, the cross-sectional correlation between the current teachers’ ratings and the CPT score was also 0.7 (Merrell and Tymms, 2001). Collectively, these data highlight the predictive validity and stability of the behavior rating scale.

Baseline 'high scorers' reflected having sufficient (≥6) symptoms relating to one of the three ADHD sub-types in DSM-IV (inattentive, hyperactive/impulsive, or combined). Following this, the interventions were randomly assigned at the school level. Schools received one of three interventions during school Year 1 or no intervention, when the children were aged 5 years:

1) Identification and feedback of the names of high scoring children to the school (Identification).

2) Provision of a research-based booklet of advice containing information about ADHD-like behavior and evidence-based ways on how to help these children in the classroom (Book).

3) Receipt of both identification information and the book.

4) No intervention control group.
There was no evidence of differences in the schools' characteristics across randomized assignment.

In terms of contextual factors, participating schools reflected a wide spread of socio-economic characteristics which aids the generalizability of the findings. In terms of socio-economic data and academic attainment results, participating schools are comparable to the rest of England. Parents were not informed about the identification information as part of the research but teachers may have shared this information with them. In England, teacher knowledge about ADHD is variable and tends to be limited in relation to the inattentive subtype (Moldavsky, Groenewald, Owen, & Sayal 2013). Usual school supports reflect 'School Action' involving additional within-school support (such as small group work or sometimes a limited number of hours of individual input) or 'School Action Plus' whereby professionals from outside the school may be involved. In terms of medication use, national data at the time suggested that around half of children who met criteria for ADHD were prescribed medication (Sayal, Ford, & Goodman, 2010). A five-year follow-up of a sub-sample of children who participated in the initial RCT indicated that, amongst children who met criteria for ADHD and had seen specialist health services, one-third had been prescribed stimulant medication (Sayal, Mills, White, Merrell, & Tymms, 2014).

**Predictor measures**

1) Intervention group in the RCT.

2) Symptoms of inattention (score range 0-9), hyperactivity (0-6), impulsiveness (0-3).
Outcome measures

Academic outcomes were assessed using results obtained in the Key Stage 2 (KS2) statutory tests taken in the final year at primary school (at ages 10-11 years). These provide an objective 'real-world' measure of academic achievement. In England, the KS2 period encompasses the school years 3-6 and formal statutory tests in English and Mathematics are taken at the end of school Year 6. The results reflect the 'Level' achieved in these tests and range from Levels 2 to 5; the majority of children are expected to achieve Level 4. Outcome data were available at one decimal point gradations with these ‘Fine Level’ grades ranging from 2.5–5.9. Further details about the KS2 curriculum and exams are available at: http://curriculum.qcda.gov.uk/key-stages-1-and-2/index.aspx; http://www.bristol.ac.uk/cmpo/plug/support-docs/ks2userguide2011.pdf.

The Cronbach’s alpha values of the KS2 exams for the relevant academic year were:

Reading test $\alpha = 0.89$; spelling test $\alpha = 0.89$; mathematics test $\alpha = 0.92$; mental mathematics test $\alpha = 0.89$; and science test $\alpha = 0.84$ (Merrell, 2009).

Confounder variables

1) Child gender

2) Baseline academic assessments - these were conducted individually, usually by the class teacher. Children were assessed at the start of school in the Reception year (First academic assessment score) and again at the end of the Reception year (Second academic assessment score) by their teachers. The assessment (Performance Indicators in Primary School (PIPS)) On-entry Baseline Assessment and follow-up) included measures of early reading and early mathematics. The assessment has good psychometric properties (internal reliability of the full scale is 0.94 (Cronbach’s
alpha), test/re-test reliability between the first and second assessment scores is 0.98 (Merrell and Tymms, 2005) and correlation between the total scores and cognitive development at age 11 is 0.68 (Tymms, Merrell, Henderson, Albone & Jones, 2012)).


4) Whether English is the pupil’s first language (data collected at follow-up)

5) Whether the pupil is entitled to free school meals (data collected at follow-up)

Analysis

The following sets of analyses were carried out:

1) Attrition analyses comparing those with and without follow-up data using baseline data (gender, academic assessment scores, ADHD symptom scores (all 3 domains), and RCT intervention group).

2) Multi-level models were used to analyze KS2 academic outcomes for: a) children with high levels of ADHD symptoms at baseline and b) the whole sample (adjusting for baseline inattention, hyperactivity and impulsiveness scores to assess whether risk increases with each one-point increase in symptoms). This latter set of analyses assessed whether the interventions had any wider impact on academic outcomes for all children, regardless of their level of ADHD symptoms at baseline. The models also adjusted for child gender, baseline academic assessments scores, socio-economic deprivation index score, free school meal status, and English as first language. The interaction between Identification and Book was included in the models based on the factorial design for the intervention. Estimated differences in mean scores or
gradients, standard errors and p-values were reported for the interventions and the confounder variables included in the multi-level models. The multi-level models accounted for heterogeneity between schools and used robust standard errors for testing for associations between academic outcomes and the covariates. Hence the analyses account for students nested within schools. Multi-level models minimise the risk of Type 1 errors that could result from ignoring heterogeneity between schools. Effect sizes for the interventions were calculated as the ratio of differences in mean scores and the standard deviation of within-school variance (Hedges, 2007). The models were fitted using SAS 9.3.

RESULTS

Sample attrition

Key Stage 2 follow-up data were available on 52,075 (71%) of 73,367 children. There were no systematic differences between those with and without follow-up data in terms of gender, baseline academic assessment scores, ADHD symptom scores (all 3 domains), and RCT intervention group.

Children with high levels of ADHD symptoms

At baseline, 8.14% of children were above cut-off for inattention and 4.61% for hyperactive/impulsiveness. For children at risk of ADHD (high levels of inattention or high levels of hyperactivity/impulsiveness) there was no main effect of the interventions at follow-up (Tables 1 & 2). Baseline academic assessment scores, gender, socio-economic deprivation index score, and free school meals and English first language status were associated with academic outcomes at follow-up.
Whole sample

For the full sample, in adjusted analyses, there was a positive impact of the book intervention on KS2 Mathematics scores (Table 3). The effect size was 0.06. There was no interaction between gender and intervention. Baseline academic assessment scores, socio-economic deprivation index score, and free school meals and English first language status were all associated with academic outcomes at age 11. There were gender differences in attainment in Mathematics and English tests. The findings also highlighted that a higher number of inattentive symptoms was associated with worse academic outcomes. In contrast, a higher number of impulsiveness symptoms was associated with slightly better scores in these tests.

DISCUSSION

There was no evidence of beneficial or worse outcomes, following the interventions, for children at risk of ADHD at 6 year follow-up. The lack of any long-term effects for this group might reflect the nature of this low intensity intervention. However, although the intervention was focused on children at risk of ADHD, the main positive findings were for the whole sample rather than for the at-risk children. Significantly positive effects of the book intervention for scores in mathematics were found for the whole sample. Although it is possible that this may reflect a chance finding, the large sample size allows for small effects to be detected at a statistically significant level. The effect size was small which may reflect the nature of a light-touch intervention.
For children at risk of ADHD, although findings at the two year follow-up demonstrated that the book intervention was associated with improved academic attainment (Tymms & Merrell, 2006), there was also a suggestion that the receipt of the combined book and identification interventions was associated with possible adverse negative impact on attainment in reading and mathematics. This possible adverse negative impact was not evident at the 6 year follow-up. Collectively, these follow-up findings demonstrate that there are wider positive effects of the book intervention for longer-term academic attainment and, in contrast, there was no effect for the identification intervention on academic attainment at either the two or six year follow-up.

We confirm findings from previous follow-up studies suggesting that inattention symptoms are important in predicting academic risk (Polderman et al., 2010). However, we also found an association between a higher number of impulsiveness symptoms and better academic outcomes. After adjusting for inattention symptoms, a positive correlation has been found (Tymms & Merrell, 2011) between performance in a mathematics assessment (start of year) and impulsiveness symptoms (end of year), particularly relating to blurting out answers. However, this study utilized data collected within a single school-year. In contrast, the present findings are novel in terms of demonstrating an association between impulsiveness at baseline and better academic outcomes at six year follow-up. In terms of possible mechanisms, aspects of impulsivity such as blurting out might be a marker of cognitive engagement (Mayer, 2004). For example, a young child may be so excited by an idea that they cannot stop themselves from blurting out an answer. Such cognitive engagement in itself may either be the result of or could lead to academic progress, with the act of verbalizing
helping to consolidate their learning. Distinguishing between these two hypotheses would require experimental testing (Tymms & Merrell, 2011). Although, at a population-wide level, this finding suggests a possible beneficial aspect to having traits of impulsiveness (Williams & Taylor, 2006), it requires further replication.

**Strengths and Limitations**

This study contributes to the literature as long-term follow-up studies of RCTs of educational or psychological interventions for children with or at risk of ADHD are scarce. Other follow-up studies of interventions have focused more broadly on children at risk of behavioral problems (e.g. the Perry Pre-school Project (Schweinhart et al., 2006); the Chicago Child-Parent Center Program (Reynolds, Temple, Robertson, & Mann, 2001) and the Montreal Longitudinal Experimental study (Vitaro, Brendgen, & Tremblay, 1999)). Particular strengths of the study reflect: 1) the investigation of the roles of all three cardinal symptoms of ADHD; 2) the use of a large school-based epidemiological sample and long period of follow-up; 3) close attention to confounders (including baseline academic levels, socio-economic factors) and 4) the use of an externally-marked ‘real world’ outcome measure.

There are also several limitations to note. First, the interventions were of low intensity and may not be expected to have long-term benefits. Second, there was sample attrition over the follow-up period. However, this was not associated with baseline characteristics. Third, risk status was identified through a single teacher rating and scores may have reflected the child’s relative level of maturity during their first school year or teacher factors.
Practice and Research Implications

For children meeting diagnostic criteria for ADHD, although medication interventions can improve on-task behaviour in the classroom (Prasad et al., 2013), it is unclear whether this translates into improved academic achievement in the long-term (Langberg & Becker, 2012). Our findings suggest that children with high levels of ADHD symptoms are likely to require ongoing input rather than a one-off intervention in their early school years. In contrast, school-based provision of behavioral or educational interventions can maximise the potential reach of interventions. Although the effect size was small, the book intervention was a very low-key and cheap intervention delivered during just one school year within children’s elementary education. As its long-term benefit was not specific to children with ADHD characteristics, further work is required to investigate whether the beneficial effects of the booklet could be optimized. In terms of future research, RCT methodology could be used to investigate the effect of providing schools with age-appropriate strategies each year as the children move through their elementary education or of accompanying the booklet with in-service training models to embed the strategies into teachers’ practice.

Funding:

The interventions described in this paper were funded by the ESRC (grant R000223798).
REFERENCES


Table 1: Outcomes for inattentive children (n = 3892 with complete data)

<table>
<thead>
<tr>
<th>Variables</th>
<th>MATH</th>
<th>ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>Male gender</td>
<td>0.2703</td>
<td>0.0244</td>
</tr>
<tr>
<td>IDACI Score</td>
<td>-0.1199</td>
<td>0.0677</td>
</tr>
<tr>
<td>Free school meals (yes)</td>
<td>-0.0961</td>
<td>0.0290</td>
</tr>
<tr>
<td>English first language (yes)</td>
<td>0.3155</td>
<td>0.0470</td>
</tr>
<tr>
<td>First academic Assessment score</td>
<td>0.0223</td>
<td>0.0020</td>
</tr>
<tr>
<td>Second academic Assessment score</td>
<td>0.0408</td>
<td>0.0020</td>
</tr>
<tr>
<td>Book</td>
<td>0.0433</td>
<td>0.0377</td>
</tr>
<tr>
<td>Identification</td>
<td>0.0117</td>
<td>0.0372</td>
</tr>
<tr>
<td>Book*Identification</td>
<td>-0.0385</td>
<td>0.0531</td>
</tr>
</tbody>
</table>

SE = Standard Error; IDACI = Income Deprivation Affecting Children Index
Table 2: Outcomes for hyperactive/impulsive children (n = 2208 with complete data)

<table>
<thead>
<tr>
<th>Variables</th>
<th>MATH</th>
<th></th>
<th></th>
<th>ENGLISH</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>p-value</td>
<td>Estimate</td>
<td>SE</td>
<td>p-value</td>
</tr>
<tr>
<td>Male gender</td>
<td>0.2447</td>
<td>0.0348</td>
<td>&lt;.001</td>
<td>-0.0790</td>
<td>0.0309</td>
<td>0.011</td>
</tr>
<tr>
<td>IDACI Score</td>
<td>-0.1092</td>
<td>0.0869</td>
<td>0.209</td>
<td>-0.3195</td>
<td>0.0773</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Free school meals (yes)</td>
<td>-0.1001</td>
<td>0.0383</td>
<td>0.009</td>
<td>-0.1061</td>
<td>0.0340</td>
<td>0.002</td>
</tr>
<tr>
<td>English first language (yes)</td>
<td>0.2122</td>
<td>0.0647</td>
<td>0.001</td>
<td>0.3004</td>
<td>0.0576</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>First Assessment score</td>
<td>0.0226</td>
<td>0.0025</td>
<td>&lt;.001</td>
<td>0.0201</td>
<td>0.0022</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Second Assessment score</td>
<td>0.0411</td>
<td>0.0024</td>
<td>&lt;.001</td>
<td>0.0418</td>
<td>0.0022</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Book</td>
<td>0.0201</td>
<td>0.0457</td>
<td>0.659</td>
<td>0.0226</td>
<td>0.0408</td>
<td>0.580</td>
</tr>
<tr>
<td>Identification</td>
<td>-0.0271</td>
<td>0.0452</td>
<td>0.549</td>
<td>-0.0163</td>
<td>0.0404</td>
<td>0.686</td>
</tr>
<tr>
<td>Book*Identification</td>
<td>0.0058</td>
<td>0.0645</td>
<td>0.928</td>
<td>-0.0016</td>
<td>0.0577</td>
<td>0.977</td>
</tr>
</tbody>
</table>

SE = Standard Error; IDACI = Income Deprivation Affecting Children Index
Table 3: Outcomes for whole sample (n = 46,369 with complete data)

<table>
<thead>
<tr>
<th>Variables</th>
<th>MATH</th>
<th>ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>Male gender</td>
<td>0.2367</td>
<td>0.0055</td>
</tr>
<tr>
<td>IDACI Score</td>
<td>-0.1746</td>
<td>0.0202</td>
</tr>
<tr>
<td>Inattention (0-9)</td>
<td>-0.0534</td>
<td>0.0017</td>
</tr>
<tr>
<td>Hyperactivity (0-6)</td>
<td>-0.0065</td>
<td>0.0035</td>
</tr>
<tr>
<td>Impulsiveness (0-3)</td>
<td>0.0217</td>
<td>0.0046</td>
</tr>
<tr>
<td>Free school meals</td>
<td>-0.0665</td>
<td>0.0084</td>
</tr>
<tr>
<td>English first language (yes)</td>
<td>0.2707</td>
<td>0.0143</td>
</tr>
<tr>
<td>First academic Assessment score</td>
<td>0.0202</td>
<td>0.0004</td>
</tr>
<tr>
<td>Second academic Assessment score</td>
<td>0.0336</td>
<td>0.0005</td>
</tr>
<tr>
<td>Book</td>
<td>0.0345</td>
<td>0.0158</td>
</tr>
<tr>
<td>Identification</td>
<td>0.0125</td>
<td>0.0154</td>
</tr>
<tr>
<td>Book*Identification</td>
<td>-0.0309</td>
<td>0.0222</td>
</tr>
</tbody>
</table>

SE = Standard Error; IDACI = Income Deprivation Affecting Children Index