The specificity of associations between cognition and change in English, maths and science attainment during adolescence

Authors:

Georgina Donati[1], Emma L Meaburn[1], and Iroise Dumontheil[1]

[1]Department of Psychological Sciences, Birkbeck, University of London, London, UK.

Word Count: 8,813

Corresponding author:

Dr Iroise Dumontheil

Department of Psychological Sciences

Birkbeck, University of London

Malet Street

WC1E 7HX

Email: i.dumontheil@bbk.ac.uk

Tel: 0044 203 073 8008

Declarations of interest: none

Abstract

Executive functions (EFs) are predictive of early academic attainment. However, there is little research investigating whether academic outcomes are differentially associated with cognitive abilities during adolescence, when EFs are still developing. Using a large population-based sample, three latent components, working memory, inhibitory control, and processing speed, were characterised from ten cognitive tasks. These components were used in structural equation models alongside measures of IQ (vocabulary, matrix reasoning) to assess specific relationships with English, maths and science attainment at 16 years of age while controlling for socio-economic status (SES) and previous attainment at age 11. Cognitive measures and SES contributed to individual differences in change in academic performance across adolescence, and specific associations between cognitive abilities and academic subjects could be observed. These results show that SES and cognitive abilities, in particular working memory, continue to influence academic progress beyond childhood, and that these associations are specific to individual academic subjects.

Keywords: Executive Function, Adolescence, Academic attainment, Cognition, ALSPAC

1. Introduction

Beyond attitude and motivational factors, a number of cognitive abilities have been proposed to explain individual differences in academic attainment, including executive functions (EFs), attention, and IQ (Best, Miller, & Naglieri, 2011; Cragg & Gilmore, 2014; St Clair-Thompson & Gathercole, 2006). EFs are a set of cognitive processes, partially distinct from IQ (Friedman & Miyake, 2017; Friedman et al., 2006; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003), which are necessary for the voluntary control of behaviour and the successful achievement of goals. Studies investigating EFs as predictors of academic attainment have tended to focus on preschool and primary school years (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Bull & Scerif, 2001; Espy et al., 2004; Gilmore et al., 2013) and few have sought to investigate the differential effects of cognitive abilities on academic subjects. The present study aimed to examine specific relationships between IQ, EFs and academic subject attainment in adolescence, a period when EFs continue to develop and differentiate, as do schooling demands, potentially changing the relationship between cognitive abilities and academic attainment (Best & Miller, 2010; Huizinga, Dolan, & van der Molen, 2006; Lee, Bull, & Ho, 2013; Luna, Marek, Larsen, Tervo-Clemmens, & Chahal, 2015).

Dominant adult EF frameworks such as the Miyake-Friedman model distinguish between: working memory (WM), the ability to hold and manipulate information in mind; shifting, the ability to flexibly switch attention between different tasks, rules, or mental states; and inhibitory control (IC), the ability to suppress distracting information and unwanted responses. These are considered key aspects of executive functioning, which, although correlated when measured experimentally, can also be separated (Lee et al., 2013; Miyake et al., 2000) - although more recent work suggests that IC may not necessarily load into a distinct factor (Friedman & Miyake, 2017). Developmental research indicates that the unity and diversity of EFs is also observed during development, for example Letho et al. (2003) identified in 8-12 year olds three distinct, but highly correlated, EF factors. However, rather than identifying adult-like executive functioning across age groups, studies suggest a general pattern of increasing specialisation of EFs with age. Indeed, while studies of preschool children are more likely to observe undifferentiated EFs (single factor model), older children and young adolescents

show more specialised EFs, with two or three factor models best fitting the data (see Lee et al., 2013). In their own accelerated design longitudinal study, Lee et al. (2013) found that during childhood and early adolescence (5-13 years) a two-factor model of EF fit the data best, whereas by age 15 a three-factor model had emerged. Another longitudinal study between the ages 17 and 23 years found that individual differences in EFs were quite stable by late adolescence, but that common aspects of EFs were still sensitive to environmental influences (Friedman et al., 2016). These developmental studies typically show prolonged improvements in performance during childhood and adolescence, with somewhat distinct developmental trajectories for each EF factor (Huizinga et al., 2006; Lee et al., 2013), supporting the idea that they reflect separable cognitive processes. Some of the observed changes with age, and possibly the commonality of EFs, may reflect differences in processing speed, which some studies model as a separate factor (e.g. Christopher et al., 2012; Huizinga et al., 2006; Kail, 2000; McAuley & White, 2011, see Lee et al., 2013 for discussion). Other influential developmental models of EF additionally include aspects of attention, which may share some mechanisms with WM (Fougnie, 2008; Wendelken, Baym, Gazzaley, & Bunge, 2011), inhibitory control (Luna et al., 2004), and planning (Anderson, 2002), namely in that selective and sustained attention are critical for maintaining items in working memory in the presence of distraction (Awh & Jonides, 2001; Shimi & Scerif, 2017; Wendelken et al., 2011). Individual task measures and latent factors of EFs have both been used to investigate the role of EFs in academic attainment, independently of IQ.

Intelligence is a well-established predictor of academic attainment, with correlations ranging from .3 to .7 (see Roth et al., 2015 for review). The close relationship between academic attainment and non-verbal IQ has been replicated in at least 40 countries across the world (Lynn & Mikk, 2007). A meta-analysis of 162 studies with an international sample of more than 100,000 individuals with a mean age of 13.9 years found a correlation of .54 between IQ and academic attainment, across subjects and ages. Moderator analyses indicated that verbal IQ was a higher predictor of academic attainment (.53) than non-verbal IQ (.44), and that the association between IQ and academic attainment was lower in elementary (.45) than in middle and high school (.54 and .58 respectively, which did not differ) (Roth

et al., 2015). Another large cohort longitudinal study found non-verbal IQ to consistently predict ~40% of the variance in attainment throughout primary and secondary school (Laidra, Pullmann, & Allik, 2007). Verbal IQ, which measures crystallised intelligence, represents what has already been learnt via non-verbal IQ, which is thought to reflect fluid or general intelligence. Using both measures allows us to distinguish between the contribution of previous learning and current ability (Soares, Lemos, Primi, & Almeida, 2015). Deary and colleagues found that a common general intelligence latent factor combining verbal, quantitative, and nonverbal reasoning measures at age 11 (measured by the Cognitive Abilities Test 2nd Edition), predicted 59% of variance in maths and 48% in English attainment at age 16 in more than 70,000 participants, with a smaller amount of variance in English and science, but not maths, additionally explained by a residual verbal factor (Deary, Strand, Smith, & Fernandes, 2007). Earlier in development, Alloway and Alloway (2010) found that at age 5 verbal IQ predicted literacy at age 11, while maths was predicted by non-verbal IQ (Alloway & Alloway, 2010). These two studies demonstrate how verbal and non-verbal IQ can vary in their subject specific role across development.

Although IQ is a strong predictor of academic attainment, EFs have been shown to predict academic attainment independently of IQ both through individual task measures and latent factors (Alloway & Alloway, 2010; Cragg & Gilmore, 2014; Rhodes et al., 2016). A large number of cross-sectional studies have provided evidence that WM and inhibitory control account for unique variance in arithmetic, beyond variance explained by IQ, age, processing speed or reading, in a wide range of age groups (e.g. Monette, Bigras, & Guay, 2011; Bull & Scerif, 2001; see Cragg & Gilmore, 2014 for review). In general, associations between WM and maths and literacy have tended to be more consistent across ages, while IC may be a stronger predictor of pre-school (Blair & Razza, 2007; Espy et al., 2004), but not necessarily later primary school, maths and literacy (Bull & Scerif, 2001). Evidence is more limited regarding predictors of science attainment and compared with maths or reading, considerably less is known about the role of executive functions in science learning (Tolmie, Ghazali, & Morris, 2016). Using a large task battery including measures of both response and semantic inhibition, a cross-sectional study in 10 and 11 year-olds found a relationship between

English, maths and science attainment and IC (St Clair-Thompson & Gathercole, 2006). A study with early adolescents (12-13yrs olds) found spatial working memory and planning ability but not inhibition or shifting to be predictive of learning biology (Rhodes et al., 2014). Visual-spatial working memory has also been shown to predict chemistry (Rhodes et al., 2016) and physics performance in adolescents (Chen & Whitehead, 2009). Although shifting is associated with mathematic attainment and reading skills, meta-analyses suggest these associations are not independent of general intelligence (Yeniad et al., 2013). Some of the associations between EFs and academic attainment are observed across culture. For example, Lan et al. recruited 119 Chinese and 139 American children and found that while WM was the best predictor of complex maths and reading tasks in preschoolers, IC, measured with a response inhibition task, predicted basic maths tasks like counting (Lan, Legare, Ponitz, Li, & Morrison, 2011). There are few longitudinal studies of EFs as predictors of academic attainment. The results support the cross-sectional data, with WM and IQ found to uniquely predict maths and reading outcome in primary and secondary school (Alloway & Alloway, 2010; Dumontheil & Klingberg, 2012; Mazzocco & Kover, 2007). However, as these studies have tended not to control for early academic attainment, it is unclear whether EFs and IQ continue to uniquely influence academic outcomes beyond early effects. One study by Stipek and colleagues suggest that although working memory and attention are important in early attainment, there is a 'fade-out' by adolescence (Stipek & Valentino, 2015). This is an important issue, as a better understanding of the predictors of learning and academic attainment throughout the school years could inform the potential of targeted interventions beyond the early years (Heckman, 2006).

While the studies reviewed above collected various measures of academic attainment, few systematically investigated the potential specific influences of IQ and EFs on different academic subjects. In their large meta-analysis, Roth et al. (2015) found that across age groups IQ predicted maths and science attainment and languages and social sciences to a similar extent, with correlations of .43-.49. Best and colleagues found that while the relationships between EF and academic attainment changed over time from 5-17 years, the pattern of these correlations was similar for maths and reading, leading the authors to conclude that a domain-general mechanism must be operating

across academic subjects (Best et al., 2011). In contrast, Latzman et al. (Latzman, Elkovitch, Young, & Clark, 2010) found different associations between cognitive abilities and different academic subjects in a sample of 11-16 year old males. Using the Delis-Kaplan Executive Functions System (Delis, Kaplan, & Kramer, 2001), the study tested the association between three derived EF variables (monitoring, conceptual flexibility and inhibition) and reading, maths, social studies and science attainment, covarying for IQ. Monitoring was found to be related to reading and social studies, conceptual flexibility to reading and science, and inhibition to maths and science, suggesting some specificity of the relationship between cognition and individual academic subjects' attainment.

Socio-economic status (SES) is also often found to be correlated with academic attainment and is considered a major component influencing academic success (Hackman, Farah, & Meaney, 2010; Reardon, 2011; Sirin, 2005). Not only does the positive association persist into adolescence but it potentially increases. A longitudinal study with four data points from age 7 to 15 found that individual differences in academic achievement (AA) explained by SES remained fairly stable between ages 7 and 11 but increased from 11 to 15 years (Caro, 2009). Another growth curve model starting at kindergarten showed an initial negative association with SES in terms of rate of reading improvement, but this reversed from age 8 onwards (Kieffer, 2012). There is also evidence that SES may influence AA at least partly via executive functions (Lawson et al., 2014).

The present study therefore sought to assess specific associations between cognitive abilities and academic attainment during adolescence in a large developmental population-based sample, the Avon Longitudinal Study of Parents and Children (ALSPAC). The first step was to establish latent cognitive ability measures from a series of cognitive tasks collected during adolescence, with a particular focus on EFs related measures. Measures were derived from performance on the N-back, Counting Span and Stop Signal tasks, the Test of Everyday Attention for Children (Manly et al., 2001), a Digit Vigilance task and simple and choice reaction time (RT) tasks. As the N-back and Stop-signal tasks are two well established EF tasks, it was predicted that WM and IC factors would be found (Miyake et al., 2000). Including measures from the other tasks allowed for the identification of

more general cognitive components related to processing speed and attention. From the literature above it was possible to develop hypotheses about the likelihood of observing working memory, inhibitory control, processing speed and attention latent measures. However, as the ALSPAC variables were collected at different time points across adolescence, a period when the structure of EF and various aspects of cognition are still developing (Best, Miller, & Jones, 2009; Kail, 2000; Lee et al., 2013; Lehto et al., 2003; Luna et al., 2015), it was unclear whether we could fully constrain our latent factor analyses based on the literature. Indeed variables collected closer in time may have more in common than those purporting to measure similar constructs. Therefore, a data-driven exploratory approach was used to establish the latent factors.

These latent cognitive measures were subsequently used in a structural equation model alongside vocabulary (verbal) and matrix reasoning (non-verbal) measures of IQ, subject attainment at age 11 and social economic status (SES) to predict variance in academic achievement in English, maths and science. It was expected that SES would have general effects whereas WM, vocabulary and reasoning would show more specific associations with achievement in the three academic subjects. It was unclear whether IC, if derived, would explain any variance in the different academic subjects due to the mixed evidence regarding its role in later academic attainment. As all but one of the ALSPAC cognitive measures were collected only once during adolescence, it was not possible to perform longitudinal analysis on the cognitive ability measures, therefore the analyses presented in this paper were performed on data combined across ages 10 to 20 years, controlling for age and sex within measure. Although this may hide some age specific associations, it has been found previously that individual differences in EF across development are fairly stable (Miyake & Friedman, 2012), and combining EF measures in latent traits to test associations with academic achievement-related measures over the course of adolescence has been successful used before (Christopher et al., 2012) and therefore we felt the approach justifiable.

2. Methods

2.1 Study cohort

ALSPAC (http://www.bristol.ac.uk/alspac/) is an on-going population-based study investigating factors influencing development and health. Initial recruitment included 14,541 mothers with 13,988 children alive at age one. Another round of recruitment at around age 7 left the total sample size for data collected after this age at 15,247 (Boyd et al., 2013). The study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/). The final sample for the current study includes 5,838 participants (2,784 males) aged 9 years 10 months to 20 years 0 months. Within this sample, 27 were reported to having ever had any developmental delay from birth. The process of selection from the full sample is described below. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee.

2.2 Cognitive Measures

The data dictionary was used to identify cognitive measures available during adolescence, broadly defined for this study as between the ages of 10 and 20 years, with a focus on measures of IQ, EFs, attention and processing speed. Measures from the affective Go/No-go task were considered but excluded as the literature has mainly found this task to measure positive/negative bias rather than response inhibition or shifting (Erickson et al., 2005). Also excluded was the Probability Reversal Learning task due to the amount of missing data. The final set of cognitive measures were derived from ten tasks across five time points.

2.2.1 Age 10

The **Counting Span task** (Case, Kurland, & Goldberg, 1982) (N = 5,347) is a WM task where the child is shown a series of red and blue dots on a screen and they are asked to count the red dots. The child is asked to point and count the number of red dots on each screen. At the end of each set the participant is asked to recall in order the number of red dots presented on each screen of that set. The participant had two practice blocks of two screens, then three sets each of two, three, four and five screens. A *Counting Span score* was calculated from the number of sets where the information was correctly recalled, weighted by the number of screens within each set (M = 18.9, range = 0 – 42). The

Stop Signal task (Logan & Cowan, 1984) is an IC task where the participant must respond to X's and O's on the screen by pressing the corresponding button as quickly as possible (Go trials). This establishes a mean RT baseline. On Stop trials a beep played randomly 150ms or 250ms before the participant's baseline RT indicates the participant should refrain from responding. The task started with two practice blocks: first a block of 30 Go trials, then a block of 16 Go trials and eight Stop trials. There were then two experimental blocks of 48 trials, 16 of which were Stop trials (33%). As the number of correct Stop trials in the 150ms and 250ms delay conditions were highly correlated, an average *Stop Signal number of correct Stop trials* across delays was calculated for each individual (N = 5,266) (*M* = 13, range = 4 – 16). The *Stop Signal number of correct Go trials* (N = 5,280) (*M* = 54, range = 23 – 64) and the *Stop Signal Go trials RT* (N = 5,307) (*M* = 599 ms, range = 388 – 818 ms) were also included in our analyses to reflect performance on this task more broadly (for example to consider potential trade-offs between speed, correct stops and correct key presses) and feed into a potential processing speed component.

2.2.2 Age 11

At age 11 three attention tasks from the **Tests of Everyday Attention for Children** (adapted from Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996) were performed. The **Sky Search task** (N = 5,587) assesses selective attention and motor control. Participants had to circle pairs of identical space ships from a large number of similar pairs (distractors) as quickly as possible. The *Selective attention speed* was calculated as the average time spent trying to find a pair minus a motor score, estimated by having asked participants to circle pairs of space ships with no constraints (*M* = 3.5 s, range = −0.4 – 7.8 s). The **Dual task** assesses divided attention (N = 5,534). Participants were required to repeat the task above while counting spaceship noises played throughout the task. A *Dual task decrement score* was calculated as the difference in numbers of pairs correctly identified in the Dual task compared with the Sky Search task (*M* = 0.9, range = −7.2 – 21.6). Finally, the **Opposite Worlds task** assesses attentional control (N = 5,431) (Strauss, Sherman, & Spreen, 2006). In the same world condition participants have to read aloud a sequence of 1s and 2s, while in the opposite world condition participants have to say 1 for the number 2 and 2 for the number 1. Participants read

four sequences of 24 numbers each in the order: same world, opposite world, opposite world, same world. The *Opposite World RT cost* was calculated as the proportional difference between opposite and same world RTs and represents the cost of the verbal/visual interference, controlling for same world RT (M = 0.3, range = −0.2 – 0.8). For these three measures, higher values correspond to poorer attentional control.

2.2.3 Age 13.5

At age 13.5 years participants were assessed on the Cognitive Drug Research (CDR) computerised cognitive assessment system (United BioSource Corporation). Participants performed the **Digit Vigilance task** which measures sustained attention. A number was shown on one side of the screen and remained constant while a sequence of different numbers were shown in the middle of the screen. Participants pressed a key when the side and middle numbers matched. A total of 450 numbers were presented over 3 minutes, with 45 targets (10%). Measures on this task were the *Digit Vigilance d-prime* (z-score of number of targets – z-score of number of false alarms) (N = 5,030, M = 0.1, range = −5.5 – 1.8) and *Digit Vigilance RT* (N = 5,072, M = 428 ms, range = 303 – 572 ms) for correctly detected targets. In the **Simple RT task** (N = 5,041), participants pressed a key labelled YES every time the world YES appeared on the screen. There were 30 trials, presented with varying inter-stimulus intervals. In the **Choice RT task** (N = 5,030), participants pressed keys labelled YES or NO depending on which word was presented on the screen. There was an equal probability of YES/NO trials, with 30 trials presented with varying inter-stimulus interval. Measures were respectively *Simple RT* (M = 294 ms, range = 209 – 486 ms), *Choice RT* (M = 443 ms, range = 260 – 660 ms) and *Choice RT task number of correct trials* (M = 27, range = 21 – 30).

2.2.4 Age 15.5

The **Stop Signal task** from age 10 was repeated, with the same practice and test blocks but slightly different delay times between stimulus and stop signal presentations for different participants, hence a residual score covarying for delay duration was calculated for the purpose of this study. As at age 10, the measures included in our analyses were *Stop Signal number of correct Stop trials (residual)* (N =

4,769, *M* = 0.1, range = −7.1 – 2.6). *Stop Signal number of correct Go trials* (N = 4,811, *M* = 50, range = 10 - 64), and *Stop Signal Go RT* on correct trials (N = 4,831, *M* = 566 ms, range = 309 – 818 ms). *Vocabulary* (N = 4,859, *M* = 45.7, range = 8 – 71) and *Matrix Reasoning* (N = 4,854, *M* = 24.7, range = 5 – 80) raw scores were taken from the **Wechsler Abbreviated Scale of Intelligence** (WASI, Wechsler, 1999) interview performed at age 15 (*M* = 15 years 5 months, range=14 years 3 months – 17 years 5 months).

2.2.5 Age 17

An **N-back** task was used at age 17 years to test WM, more specifically updating. Participants were presented with numbers 0-9 for 500 ms and had 3000 ms to judge whether the current number was the same as the number shown either 2-back or 3-back. The practice block consisted of 12 trials with two targets, and there were single blocks of the 2-back and 3-back conditions each consisting of 48 trials with eight targets. Measures on this task were *2-back accuracy* (N = 3,230, *M* = 77%, range = 15% – 100%) and *2-back RT* (N = 3,226, *M* = 680 ms, range = 82 – 1385 ms). In addition, as with the Opposite Worlds task, additional scores were created to represent the added WM cost of 3-back in relation to 2-back. *Accuracy 3-back - 2-back* (N = 3,048, *M* = −10%, range = −60% – 46%) and *RT (3-back – 2back)/2-back*, a proportional difference score which takes into account means overall difference in RT (N = 3,041, *M* = 0.1, range = −0.9 – 1.1).

2.3 Analysis of cognitive measures

Whilst it is of interest to examine cognitive abilities and their relationship to AA within a developmental framework, only the Stop Signal task was repeated at two different time points and therefore cognitive measures across ages 10-20 years were included in a single analysis. Principal component analysis (PCA) was used to identify measures that could be combined into latent measures of EF, processing speed, attention, or other aspects of cognition. The *Vocabulary* and *Matrix Reasoning* WASI standardised scores were not entered into the PCA and were used as stand-alone measures of IQ. A cut-off criterion for item loading was fixed at 0.3 and with the condition that the overall and individual Kaiser-Meyer-Olkin (KMO) values, a measure of sampling adequacy, be over

0.5. Oblimin oblique rotation was used as previous studies indicate it is appropriate when the inter-correlation between components is high (Field, Miles, & Field, 2012).

2.4 Socio-Economic Status (SES)

SES (N = 5,417) was calculated using the mother and partner's occupational social classes as classified by the Office of Population Census and Surveys (OPCS). Scores were reversed so that the higher the occupational status, the higher the score (i.e., 6 = professional, 1 = unskilled manual labour; OPCS, 1990). For each individual, an average score was created from the parental scores averaged across three times points: the first taken before the birth of the child (mother), then when the child was four years of age (mother and partner) and eight years of age (partner) (M = 4.0, range = 1.0 – 6.0).

2.5 Outcome measures

Academic attainment was assessed using national curriculum standardised tests at age 11 and 16 years old (**Table 1**). Age 11 English exams assess reading, grammar, punctuation and spelling. Age 16 English includes a language test assessing reading and writing ability, and a literature test which examines knowledge and understanding of a novel or play, poetry and a previously unseen text. For English at age 16, an average score between the literature and language exams was created for each individual. Maths is assessed in all years with written tests covering all areas of mathematics with the aim of testing conceptual understanding, mathematical reasoning and problem solving. The age 11 assessment also includes a 'mental maths' component in which the children are asked questions orally and under timed conditions they must write down their answers having worked them out in their heads. The science tests assess the development of scientific thinking and knowledge, experimental skills and strategies, analysis and evaluation as well as scientific vocabulary, units, symbols and nomenclature. Science at age 16 is either taken as three separate subjects (chemistry, physics and biology), or as one condensed subject (**Table 1**). Those who did the subjects separately were given an average score and those who did single science kept this score. This was considered equivalent due to the extra teacher time and cross-over between the three separate subjects. For more information about

the syllabus see https://www.gov.uk/government/collections/national-curriculum#programmes-of-study-by-subject. Scores were created at each time point for English, maths and science and were age and sex regressed. At Key Stage 2 (11 years old) children are given a curriculum level from 1-9, this score was used directly. For the General Certificate of Secondary Education (GCSE) at age 16, scores from 1-9 represent GSCE grades U-A* (U=1, G=2, F=3, E=4, D=5, C=6, B=7, A=8, A*=9).

Table 1: List of the raw Standardised Assessment Tasks scores of the ALSPAC sample used in this study. EA: AA: academic achievement; GCSE: General Certificate of Secondary Education; SAT: Standardised Assessment Task.

AA variable	**Test**	**N**	**Mean**	**Range**
English age 11	Key Stage 2 National Curriculum SATs English	11,778	5.14	3 - 6
English age 16[a]	English Literature GCSE	9,683	6.23	1 - 9
	English Language GCSE	11,337	6.06	1 - 9
Maths age 11	Key Stage 2 National Curriculum SATs Maths	11,823	6.13	4 - 8
Maths age 16	Maths GCSE	11,231	5.79	1 - 9
Science age 11	Key Stage 2 National Curriculum SATs Science	12,165	6.38	4 - 7
Science age 16[b]	Single Science GCSE	7,319	5.75	1 - 9
	Chemistry GCSE	1,588	7.33	4 - 9
	Physics GCSE	1,596	7.31	4 - 9
	Biology GCSE	1,696	7.23	1 – 9

[a] For all participants English Literature and Language GCSE scores were averaged into a single score.
[b] Participants who took science subjects separately were given an average score across the three subjects.

2.6 Statistical analysis

As expected, both predictor and outcome measures (other than SES) were correlated with age at time of testing and so age was regressed out of each individual measure for all measures other than SES. Some measures also showed associations with sex and therefore all measures (other than SES) were also sex regressed. All measures (other than SES) entered in the analyses were standardised residual

measures (**Table 2**). Outliers further than 3.29 SD from the mean for each cognitive measure were excluded, removing 806 data points across 18 variables. These points were treated as missing data.

Table 2: Summary of the demographic variable and age-regressed standardised residual scores of the cognitive and academic attainment variables. NC: National Curriculum; RT: reaction time.

Age (y)		Standardised residual scores[a] (mean)	
		Males	**Females**
Demographic variable			
Birth, 4 & 8	**Socio-economic status**	3.944[b]	3.950[b]
Cognitive variables			
10	**Counting Span task:** *Score*	-0.159	0.166
10	**Stop Signal task:** *Number of correct Stop trials*	-0.029	0.026
	Stop Signal task: *Number of correct Go trials*	-0.170	0.152
	Stop Signal task: *Go trials RT*	**-3.724***	3.557
11	**Sky Search task**: *Selective attention speed*	0.233	**-0.223***
	Dual task: *Decrement score*	0.086	**-0.082***
	Opposite Worlds task: *RT cost*	0.008	**-0.008***
13.5	**Digit Vigilance task**: *D-prime (targets – false-alarms)*	-0.177	**0.167***
	Digit Vigilance task: *RT*	**-1.596***	1.541
13.5	**Simple RT task:** *Simple RT*	**-1.653***	1.560
	Choice RT task: *Choice RT*	**-6.807***	6.432
	Choice RT task: *Task accuracy*	-0.346	**0.329***
15	**WASI subtest:** *Vocabulary*	**0.389***	-0.344
	WASI subtest: *Matrix Reasoning*	0.013	-0.012
15.5	**Stop Signal task:** *Number of correct Stop trials*	**-0.154***	0.140
	Stop Signal task: *Number of correct Go trials*	0.289	-0.265
	Stop Signal task: *Go trials RT*	**-4.472***	4.066
17	**N-back task:** *2-back accuracy*	**0.011***	-0.009
	N-back task: *2-back RT*	-4.715	3.568
	N-back task: *Accuracy 3-back - 2-back*	-0.004	0.003
	N-back task: *RT (3-back – 2back)/2-back)*	0.028	**-0.022***
Academic attainment			

11	**NC Levels:** *English age 11*	-0.112	**0.111***
	NC Levels: *Maths age 11*	**0.039***	-0.039
	NC Levels: Sc*ience age 11*	-0.000	0.000
16	**NC Levels:** *English Literature age 16*	-0.287	**0.263***
	NC Levels: *English Language age 16*	-0.296	**0.291***
	NC Levels: *Maths age 16*	-0.028	0.028
	NC Levels: *Single Science age 16*	-0.063	**0.062***
	NC Levels: *Biology age 16*	-0.048	0.055
	NC Levels: *Chemistry age 16*	-0.060	**0.072***
	NC Levels: *Physics age 16*	-0.014	0.019

[a] Standardised residual scores for all available data after regressing out age.
[b] No regression of age was done on the SES score, which ranged from 1 to 6.
* Significant sex differences after regressing out age, $p < .05$. Bold font highlights better performance.

All statistical analyses were conducted using R (R Core Team, 2013). Missing data is an inevitable problem when working with longitudinal data due to participant drop-out or non-response. Restricting analyses to only participants with complete data risks introducing potential biases (Schafer & Graham, 2002) and greatly reduces statistical power. In the present study 1,726 participants had complete data. The decision was made to impute missing data using the current best method, multiple imputation using chained equations (Schafer & Graham, 2002), for participants who had more than 50% complete data, with the constraint that each variable should have less than 50% data missing (Buuren & Groothuis-Oudshoorn, 2011; White, Royston, & Wood, 2011). Measures of vocabulary and matrix reasoning, SES, and academic attainment were included in the imputation model of the cognitive data to address the missing at random (MAR) assumption and improve the imputation model (Buuren & Groothuis-Oudshoorn, 2011; White et al., 2011). Imputed and non-imputed distributions were compared to ensure they were similar before rank normalisation (Beasley, Erickson, & Allison, 2009). Rank normalisation is a common method of normalising data so that it conforms to the assumptions of parametric statistical tests. All subsequent analyses using imputed data (N = 5,838) were compared with analyses of the complete case data (N = 1,726).

A structural equation model was fit using the Lavaan version 0.5-23.1097 (Rosseel, 2012) package in R. The Robust Maximum Likelihood estimator with Yuan-Bentler scaled test statistic (MLR) was used to account for any violations of multivariate normality. We assessed the overall fit of the model with the chi-square test, the Root Mean Square Error of Approximation (RMSEA) and confidence interval, the Comparative Fit Index (CFI) and the Standardised Root Mean Squared Residuals (SRMR). A good model fit was defined as RMSEA < .06, SRMR < .08 and CFI ≥ .95 (Hu & Bentler, 1999; Schreiber, Nora, Stage, Barlow, & King, 2006).

3. Results

Analyses were run on imputed data to increase statistical power and reduce bias (N = 5,838). Similar results were obtained in analyses of the complete case data (N = 1,726, see **Supplementary material Table S1, Table S2** and **Figure S1**).

3.1 Principal component analyses of EF tasks data

The Bartlett's test of sphericity, performed with all 19 cognitive variables, indicated that correlations between items were sufficiently large for PCA, (χ^2 (171) = 22513.69, $p < .001$). The *Accuracy 3-back − 2-back* measure (age 17) was removed from the analysis as it had a KMO value of less than 0.5; this left an overall KMO of 0.68. A scree plot indicated that two, three or four components would be possible. The models were compared and a three-factor solution was the most parsimonious while explaining maximum variance with a good model fit (2-factor fit= .64, root mean square of the residuals (RMSR) = .11; 3-factor fit = .80, RMSR = .08; 4-factor fit = .78, RMSR = .08). In the next step, *Opposite World RT cost* at age 11 was removed as it did not load above 0.3 and *Stop Signal number of correct Stop trials* at age 10 was removed as it loaded both on the second and third components. Finally *2-back RT* (age 17), *Stop Signal number of correct Go trials* at age 10 and *Choice RT task number of correct trials* (age 13) were removed from the analysis for double loading on components one and two. Fourteen variables were retained for a final three-component solution which explained 47% of the variance in the data and had a model fit of 0.8 with correlations between

principal components one and two being r = 0.13, between one and three r = -0.21, and finally between two and three r = -0.03. **Table 3** shows the final component loadings after rotation.

Table 3: Component loadings in the final cognitive measures principal component analysis. PC: principal component; RT: reaction time.

Variable description	*PC1 Slow Processing*	*PC2 Working Memory*	*PC3 Inhibitory Control*
Digit vigilance RT (age 13)	**0.80**	0.20	−0.07
Choice RT (age 13)	**0.74**	−0.13	0.07
Simple RT (age 13)	**0.66**	−0.20	0.04
Stop Signal Go trials RT (age 10)	**0.49**	0.01	0.18
Selective attention speed (age 11)	**0.41**	−0.11	−0.09
2-back accuracy (age 17)	−0.08	**0.75**	0.00
Digit Vigilance d-prime (age 13)	0.12	**0.72**	0.01
Counting Span score (age 10)	−0.12	**0.60**	0.04
RT (3-back – 2-back)/(2-back) (age 17)	0.00	**0.55**	0.01
Dual task decrement score (age 11)	−0.09	**−0.42**	0.04
Stop Signal number of correct Stop trials (age 15)	−0.11	0.14	**0.83**
Stop Signal Go trials RT (age 15)	0.17	−0.04	**0.79**
Stop Signal number of correct Go trials (age 15)	0.02	0.18	**−0.66**
Eigenvalues	2.16	2.11	1.82
% of variance explained	17%	16%	14%
α	0.65	0.60	0.66

The first component (PC1) consisted of all positive loadings of RT measures across tasks and ages. As slow RTs were associated with a high PC1 score, we named this component *Slow Processing.* The second component (PC2) included measures from the N-back and Counting Span tasks, which are both WM tasks. Digit vigilance is a sustained attention task primarily, however much literature argues that attention and WM are overlapping concepts (e.g. Awh & Jonides, 2001; Fougnie, 2008; Wendelken et al., 2011). Finally, the Dual task could also be conceptualised as reflecting WM abilities, as participants needed to keep and update a number in verbal WM while performing a

visuospatial task. This component was therefore named *Working Memory*, with a high PC2 score reflecting better WM. Variables constituting the last component (PC3) all came from the Stop Signal task administered at age 15. Within this component the highest loading variable quantified the ability to 'Stop' when a beep was heard, followed by a slow RT in Go trials and poorer accuracy in Go trials, reflecting the common result of slowing responses to increase stopping accuracy (Verbruggen & Logan, 2009). We therefore named this component *Inhibitory Control.*

3.2 Correlation analyses

A score for each cognitive component was computed for every individual using the imputed dataset (N = 5,838) and weighted by variable component loadings. **Table 4** reports the Pearson's correlation coefficients between demographic, cognitive and academic attainment variables. A higher SES was associated with higher academic attainment at age 11 (r = 0.30 – 0.33) and age 16 (r = 0.41 – 0.43), and to a similar extent higher vocabulary and working memory (r = 0.29 – 0.36). Associations were weaker with reasoning, inhibitory control and processing speed. The *Inhibitory Control* and *Slow Processing* components also showed weaker associations with IQ than the *Working Memory* component. *Inhibitory Control* showed small or non-significant associations with academic attainment, while significant associations were observed between academic attainment and vocabulary (r = 0.47 – 0.66) and reasoning (r = 0.23 – 0.35) IQ measures, and the *Working Memory* (r = 0.46 – 0.63) and *Slow Processing* components (r = -0.17 – -0.24). AA variables were more strongly correlated with each other at age 16y (r = 0.76 – 0.85) than age 11y (r = 0.59 – 0.63).

Table 4: Correlation matrix between the demographic, cognitive and academic attainment variables. Values are Pearson's r. SES: socioeconomic status.

	1	2	3	4	5	6	7	8	9	10	11
1. SES											
2. Vocabulary	0.36										
3. Reasoning	0.14	0.20									
4. Inhibitory	-0.07	-0.13	-0.01[a]								

Control											
5. Working Memory	0.29	0.46	0.30	-0.03							
6. Slow Processing	-0.05	-0.14	-0.06	0.13	-0.21						
7. English 11	0.33	0.53	0.23	-0.03	0.50	-0.24					
8. Maths 11	0.30	0.47	0.29	-0.02[a]	0.58	-0.22	0.59				
9. Science 11	0.30	0.52	0.25	-0.03[a]	0.46	-0.17	0.59	0.63			
10. English 16	0.43	0.66	0.25	-0.07	0.56	-0.23	0.67	0.59	0.59		
11. Maths 16	0.41	0.58	0.35	-0.03	0.63	-0.19	0.60	0.73	0.63	0.76	
12. Science 16	0.42	0.65	0.32	-0.05	0.60	-0.19	0.63	0.67	0.66	0.81	0.85

[a] Non-significant correlation; all other p-values < .05

3.3 Structural Equation Model

A structural equation model (**Figure 1**) of cognitive predictors of English, maths and science attainment at age 16 controlling for socio-economic status and attainment at age 11 fit the data well: $\chi^2(152) = 2574.401$, $p < .001$; CFI = .951; RMSEA = .053 [.051, .054]; SRMR = .047 with DF = 152. Total variance explained by this model in English at age 16 was 65%, in maths 71% and in science 70%. This model was compared against a one factor model with a common cognitive factor combining working memory, inhibitory control and processing speed and the fit was significantly worse (CFI = .835, RMSEA = .090 [.089, .092], SRMR = .072, DF = 173, $\chi^2_{diff} = 6003.1$, $df_{diff} = 21$, $p < .001$) as was the fit of a model with a common academic attainment factor (CFI = .925, RMSEA = .062 [.059, .063], SRMR = .048, DF = 170, $\chi^2_{diff} = 1159$, $df_{diff} = 18$, $p < .001$). The model showed that SES, vocabulary (verbal IQ) and the *Working Memory* component predicted variance in all age 16 AA measures, reasoning (non-verbal IQ) predicted variance in maths and science only, the *Slow Processing* component predicted variance in English and maths only, and the *Inhibitory Control* component did not account for any specific variance of the change in AA between 11 and 16. All AA measures at age 11 predicted AA at age 16. The observed associations varied in their sizes, with the smallest associations found between *Slow Processing* and AA and the across subject AA longitudinal associations, and the strongest associations found between *Working Memory* and AA at age 16.

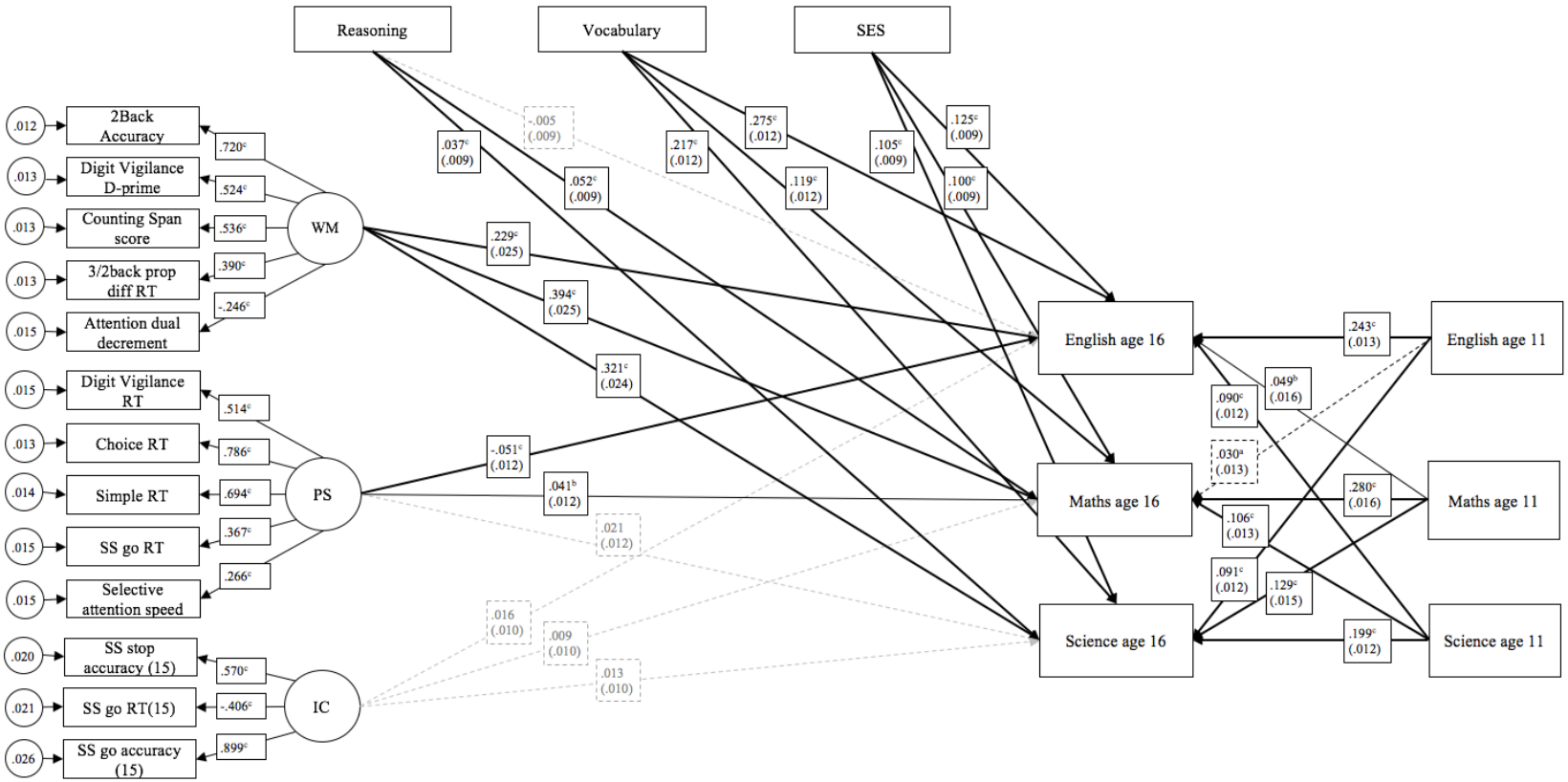

Figure 1: Imputed data structural equation model of cognitive predictors of English, maths and science attainment at age 16 controlling for socio-economic status and attainment at age 11. Figures in boxes are beta values, with standard errors in brackets. [a] $p < .05$ (black dashed lines), [b] $p < .01$ (thin full black lines), [c] $p < .001$ (thick full black lines), non-significant paths are shown with grey dashed lines. IC: inhibitory control, PS: processing speed, RT: reaction time, SES: socio-economic status, SS: Stop-Signal task, WM: working memory.

Discussion

The present study demonstrated that a range of cognitive abilities are uniquely and differentially associated with improvements in English, maths and science during adolescence. The study's strengths include the sample size, controlling for previous attainment to look at adolescence-specific effects and controlling for other subjects' attainment to assess subject-specific effects. In doing this we have shown that cognitive abilities are involved in influencing change in attainment over adolescence and that they do so with some subject specificity.

The PCA derived three cognitive components: working memory, inhibitory control and processing speed. In both the WM and PS components variables from different tasks taken at different time points loaded together suggesting there was more common variance within these two components than there was within task or age. The WM component was dominated by accuracy in the 2-back task and in the Digit Vigilance task. Both of these tasks require holding a number in working memory and updating this information when necessary. Additionally, they both require sustained attention. This is consistent with the view that WM and attention or executive attention are highly overlapping (Fougnie, 2008; Wendelken et al., 2011) or interchangeable constructs which involve the selection and maintenance of certain information in an active accessible state particularly in the presence of interference (Kane, Bleckley, Conway, & Engle, 2001). The *Opposite World RT cost* was the only measure in the dataset representing shifting. In line with previous research suggesting that it is an independent executive function, this measure did not load onto the obtained factors (Miyake et al., 2000), and did not explain sufficient unique variance to emerge as a separate component on its own. Three other variables were removed from the PCA analysis for double loading and all of these variables double loaded with the WM component suggesting it could be representing a more general ability. However, it did not correlate very highly with matrix reasoning (0.3), often considered a measure of fluid intelligence.

The main aim of this study was to investigate associations between cognitive abilities and subject-specific academic attainment during adolescence. The correlation matrix showed that attainment was

highly correlated between academic subjects and became more correlated over time. It could be that as children get older discrete subject knowledge becomes less important to success than a more general ability to apply skills and knowledge appropriately. At age 16 science attainment correlated strongly with English and maths attainment, in line with previous research showing that science involves both reading and spatial abilities (Mayer, Sodian, Koerber, & Schwippert, 2014). SEM was used to test the hypothesis that both cognitive ability and AA were better modelled through separate latent variables rather than through a single latent measure. In comparing a model with a common cognitive ability factor and a model with a common AA factor with a model where the three cognitive components derived in the PCA and the three academic subject attainment variables were kept separate, we found that the best model kept all the variables separate. First, the fact that it was preferable to keep the cognitive factors as distinct components, provides support for studies that find a multifactor solution for EF variables fits better than a one factor solution in adolescents and adults (Lee et al., 2013; Lehto et al., 2003; Miyake et al., 2000). Second, the fact that despite large correlations between academic subjects it was preferable to keep the attainment measures separate suggests the presence of important unique associations between the individual cognitive ability and subject-specific academic attainment measures.

The *Working Memory* component, which combined WM and sustained attention measures, explained significant unique variance in every subject over and above previous attainment and attainment in the other two subjects, and explained the most unique variance in maths and science improvement between the ages of 11 and 16 years old. This fits with previous research suggesting WM is the best predictor of maths attainment and does so over and above IQ (Alloway & Alloway, 2010; Dumontheil & Klingberg, 2012), while it also predicts reading (Lan et al., 2011) and literacy (Alloway & Alloway, 2010), chemistry (Rhodes et al., 2016) and physics performance (Chen & Whitehead, 2009). Importantly, the current results add to this research by demonstrating that WM continues to explain an equivalent or more variance in English, maths and science than other subject attainment and previous attainment. WM has been proposed to be important across ages for problems requiring procedural knowledge, as interim answers need to be kept in mind and combined to obtain an answer, and for

factual knowledge, as the information, such as the parts of an equation, has to be kept in mind to be memorised (Cragg & Gilmore, 2014). In science, WM has been proposed to support conceptual understanding in chemistry (Rhodes et al., 2016), physics (Chen & Whitehead, 2009) and biology (Rhodes et al., 2014) during adolescence. WM is important for improvements in English possibly due to its positive impact on reading, for example by allowing pupils to keep in mind the context of a text and keep track of long sentences. A study by Stipek and colleagues found a 'fade-out' by adolescence whereby WM and attention ceased to have a role in adolescence attainment (Stipek & Valentino, 2015). However, their EF measures did not increase in difficulty with age and ceiling effects were observed in adolescence, possibly explaining the lack of observed associations. Our findings go against a fade out effect between 11-16 years of age and highlight the continued importance of WM for maths, science and English learning during adolescence.

The *Inhibitory Control* component showed weak correlations with AA at age 16y and in the SEM failed to explain any significant unique variance in change in any of the academic subject outcomes between 11 and 16y. If inhibitory control does reflect some general ability as suggested by Miyake and Friedman (2012), we would not expect it to represent any unique variance, but instead to correlate highly with academic attainment overall. In fact, in the present study, inhibitory control correlated poorly with all other measures. However, two measures of inhibitory control from age 11 were removed from the principal component analysis for double loading, which may be evidence that inhibitory control is a more general ability. Previous research is inconclusive about the role of inhibitory control in attainment. Some studies have found it to explain unique variance in attainment in young children (Blair & Razza, 2007; Espy et al., 2004; Lan et al., 2011) but studies with more complex tasks or older children have not found this association (Bull & Scerif, 2001; Monette, Bigras, & Guay, 2011). It could be that the Stop Signal task, which requires motor inhibition, is less relevant for later academic attainment than a semantic inhibition task (Bialystok & Senman, 2004; Protopapas, Archonti, & Skaloumbakas, 2007; St Clair-Thompson & Gathercole, 2006). Similar to the results of the present study, Christopher et al. (2012) found that latent measures of working memory and processing speed, but not inhibitory control (which included the Stop Signal task) associated with

word reading and comprehension in 11-16 year olds. Note that the Stop Signal task at age 15 had a limitation in that different parameters had been set for different groups of participants. We corrected for this by regressing out the parameters from our scores, however this assumed a linear effect of delay time on accuracy.

The final component from the PCA was a processing speed measure. Processing speed is highly correlated with white matter integrity (Kievit et al., 2016), has a strong developmental trajectory and appears to moderate fluid intelligence or working memory (Fry & Hale, 2002; Huizinga et al., 2006; Kievit et al., 2016; Lee et al., 2013; McAuley & White, 2011) perhaps by assisting in goal maintenance (Iveson, Della Sala, Anderson, & MacPherson, 2017). Processing speed was found to have a different association with progression in each of the three subjects. Faster processing speed was beneficial for improvements in English, possibly due to the importance of processing speed in reading (Kail & Hall, 1994). With maths it appeared beneficial to have a slower processing speed. Further work will need to replicate this finding, which may reflect an advantage of a slower more deliberative approach to complex maths problems. Consistent with these results, a previous study in adolescents found that the ability to withhold a response on an inhibitory control Go/No-Go task was associated with longer RTs, suggesting more reflection, on science and math problems requiring counterintuitive reasoning (Brookman-Byrne, Mareschal, Tolmie, & Dumontheil, 2018). In the present study however science was not associated with processing speed.

As with previous research correlations between academic attainment and vocabulary, our measure of verbal IQ, were found to be between 0.47 – 0.60, lower correlations were observed for matrix reasoning, our measure of non-verbal IQ (0.23 – 0.36) and associations increased with age (Roth et al., 2015). Vocabulary was associated with significant unique variance in all three subjects and explained the most amount of variance in improvement in English across adolescence. It also explained more unique variance in science than previous science attainment. Weaker associations were observed for maths, a pattern in line with the finding by Deary and colleagues (2007) that beyond a common *g* factor, a verbal residual factor explained some variance in English and science

(in particular biology) but not in maths. Matrix reasoning, our measure of non-verbal IQ, did not explain any unique variance in English improvement, but explained a small amount in maths and science. Comparing these results to those of Alloway & Alloway (2010), who found that verbal IQ and non-verbal IQ at age 5 predicted literacy and numeracy at age 11 respectively, suggests that verbal IQ may become more important for maths in secondary school than in it in primary school.

As expected, SES was correlated with AA and these correlations increased with age from .30 – .33 at age 11 to .41 – .43 at age 16, in line with previous research showing an increase in association between AA and SES during adolescence (Caro, 2009). SES explained a significant proportion of variance in AA, with 10 – 13% of unique variance explained across subjects between the ages of 11 and 16. This provides evidence that SES may influence progresses in academic attainment during early and mid-adolescence, beyond effects observed in the early years, which has been the main focus of research and policy (Hackman, Farah, & Meaney, 2010). There is some suggestion that this increased influence of SES on attainment over time could be due to the hierarchical nature of learning, compounding effects over time (Caro, 2009). However it is interesting that despite controlling for cognitive ability, previous attainment and general attainment, SES still predicted change in attainment during secondary school, suggesting other factors may come into play. Further work would be needed to identify what mediates these specific associations. For example the presence of books in the house may mediate associations with English (Evans, Kelley, & Sikora, 2014; Sikora, Evans, & Kelley, 2018), while parents' numeracy may mediate associations with maths (Segers, Kleemans, & Verhoeven, 2015). Subject interest has been shown to interact with SES in predicting science attainment at age 15 across 57 countries (Drob, Cheung, & Briley, 2014) and it is also possible that specific subject interest plays an increasing role in attainment in other subjects over adolescence. Correlations in other studies between SES and individual academic subjects tend to be higher than with overall attainment, suggesting again there may be specific routes through which SES affects individual subject attainment at different points over development (Sirin, 2005).

Although the ALSPAC sample is a longitudinal sample, different tasks were performed at different ages and therefore it was not possible to chart longitudinal changes in specific cognitive abilities over time. The approach taken here was to look at associations across a broad adolescence phase to allow for the identification of latent variables. Averaging cognitive measures across ages may have hidden more specific relationships that exist at certain times in development. Indeed performance on EF tasks improves over time, with adult performance generally achieved around 15 to 17 years (Best et al., 2011). However, there is evidence that the relative strengths of EFs within individuals are stable across adolescence (Friedman et al., 2016). An alternative approach could have been to map tasks purporting to be measuring similar constructs. However, considering that it is extremely controversial what each task is measuring (Burgess, 1997), it was felt that this would require a large number of additional assumptions. The fact that the WM and PS components were populated by variables from across different measures and ages, coupled with the good fit of the SEM, suggests that the two latent variables were measuring the underlying cognitive construct (i.e. WM or PS) across the period of adolescence, rather than simply age specific or task-specific commonalities. These results are in line with those of Christopher et al. (2012) who found that EF tasks measured similar constructs over the ages of 8-10 and 11-16 years old. However, it is worth being more cautious of the IC measure, which had only variables from the age 15 stop signal task, and may have been less reliable.

4. Conclusion

The present study sought to characterise cognitive abilities and their differential associations with attainment in English, maths and science during adolescence. We found that vocabulary, reasoning, working memory and processing speed measures all had some specificity in terms of variance explained in different subjects and that they contributed to change in attainment over adolescence. The only measure that did not contribute was inhibitory control. SES was found to also contribute to explain unique variance in progress in maths, science and English in secondary school. These findings highlight the need to maintain specificity when investigating adolescent cognitive predictors of academic attainment, rather than using broad measures. The study also provides novel information regarding cognitive predictors of attainment in science, an under-studied academic subject which

shares some features with English in the shared importance of vocabulary, and with maths in the role of working memory and reasoning. Finally, the study shows the continuing importance of cognitive ability and SES in change in attainment over adolescence.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting the participants and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionist and nurses. The UK Medical Research Council and Wellcome (Grant ref:102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and they will serve as guarantors for the contents of this paper.

References

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *8*(2), 71–82.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126.

Beasley, T. M., Erickson, S., & Allison, D. B. (2009). Rank-based inverse normal transformations are increasingly used, but are they merited? *Behavior Genetics*, *39*(5), 580–595.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660.

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180–200.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336.

Bialystok, E., & Senman, L. (2004). Executive processes in appearance–reality tasks: The role of inhibition of attention and symbolic representation. *Child Development*, *75*(2), 562–579.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663.

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., … (2013). Cohort profile: The 'Children of the 90s'; the index offspring of The Avon Longitudinal Study of Parents and Children (ALSPAC). *International Journal of Epidemiology*, (42), 111–127.

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, *24*(3), 337–349.

Brookman-Byrne, A., Mareschal, D., Tolmie, A. K., & Dumontheil, I. (2018). Inhibitory control and counterintuitive science and maths reasoning in adolescence. *PLOS ONE*, *13*(6), e0198973.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273–293.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbit, (Ed.) *Methodology of frontal and executive function* (pp. 81–116). Psychology Press.

Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1-67.

Caro, D. H. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education*, *3*(32), 558–590.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386–404.

Chen, W.-C., & Whitehead, R. (2009). Understanding physics in relation to working memory. *Research in Science & Technological Education*, *27*(2), 151–160.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., … (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology: General, 141*(3), 470-488.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63–68.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21.

Delis, D. C., Kaplan, E., & Kramer, J. (2001). Delis-Kaplan Executive Function System®(D-KEFS®): Examiner's Manual: Flexibility of Thinking, Concept Formation, Problem Solving, Planning, Creativity, Impluse Control, Inhibition. Pearson.

Drob, E. M. T., Cheung, A. K., & Briley, D. A. (2014). National GDP, science interest and science achievement: A person-by-nation interaction. *Psychological Science*, *25*(11), 2047–2057.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, *22*(5), 1078–1085.

Erickson, K., Drevets, W. C., Clark, L., Cannon, D. M., Bain, E. E., Zarate Jr, C. A., … (2005). Mood-congruent bias in affective go/no-go performance of unmedicated patients with major depressive disorder. *American Journal of Psychiatry*, *162*(11), 2171–2173.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–486.

Evans, M. D. R., Kelley, J., & Sikora, J. (2014). Scholarly culture and academic performance in 42 nations. *Social Forces*, *92*(4), 1573–1605.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. Sage Publications, Inc.

Fougnie, D. (2008). The relationship between attention and working memory. In N. B. Johansen (Ed.), *New research on short-term memory* (p. 1-45). New York: Nova Biomedical Books.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204.

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, *52*(2), 326–340.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179.

Fry, A., & Hale, S. (2002). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*(1), 1–34.

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., … (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLOS ONE*, *8*(6), e67374.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036.

Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology*, *38*(1), 51–61.

Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology*, *30*(6), 949–954.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169–183.

Kieffer, M. J. (2012). Before and after third grade: Longitudinal evidence for the shifting role of socioeconomic status in reading growth. *Reading and Writing*, *25*(7), 1725–1746.

Kievit, R. A., Davis, S. W., Griffiths, J., Correia, M. M., Cam-CAN, & Henson, R. N. (2016). A watershed model of individual differences in fluid intelligence. *Neuropsychologia*, *91*(Supplement C), 186–198.

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*(3), 441–451.

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677–692.

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology*, *32*(5), 455–462.

Lawson, G. M., Hook, C. J., Hackman, D. A., & Farah, M. J., (2016). Socioeconomic status and neurocognitive development: Executive function. In Griffin, J. A., McCardle, P., & Freund, L. S. (Eds.) *Executive Function in Preschool Children: Integrating Measurement, Neurodevelopment, and Translational Research.* Washington: American Psychological Association.

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933–1953.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295.

Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., & Chahal, R. (2015). An integrative model of the maturation of cognitive control. *Annual Review of Neuroscience*, *38*(1), 151–170.

Lynn, R., & Mikk, J. (2007). National differences in intelligence and educational attainment. *Intelligence*, *35*(2), 115–121.

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., & Robertson, I. H. (2001). The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry*, *42*(8), 1065–1081.

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, *29*, 43–55.

Mazzocco, M. M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology*, *13*(1), 18–45.

McAuley, T., & White, D. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology*, *108*(3), 453–468.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Monette, S., Bigras, M., & Guay, M.-C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*(2), 158–173.

OPCS. (1990). Standard Occupational Classification. London: Her Majesty's Stationary Office.

Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology*, *54*(3), 251–282.

R Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither Opportunity*, 91–116.

Rhodes, S. M., Booth, J. N., Campbell, L. E., Blythe, R. A., Wheate, N. J., & Delibegovic, M. (2014). Evidence for a role of executive functions in learning biology: Executive functions and science. *Infant and Child Development*, *23*(1), 67–83.

Rhodes, S. M., Booth, J. N., Palmer, L. E., Blythe, R. A., Delibegovic, M., & Wheate, N. J. (2016). Executive functions predict conceptual learning of science. *British Journal of Developmental Psychology*, *34*(2), 261–275.

Robertson, I., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society*, *2*(6), 525–534.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *2*(48), 1–36.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323–338.

Segers, E., Kleemans, T., & Verhoeven, L. (2015). Role of parent literacy and numeracy expectations and activities in predicting early numeracy skills. *Mathematical Thinking and Learning*, *17*(2–3), 219–236.

Shimi, A., & Scerif, G. (2017). Towards an integrative model of visual short-term memory maintenance: Evidence from the effects of attentional control, load, decay, and their interactions in childhood. *Cognition*, *169*, 61–83.

Sikora, J., Evans, M. D. R., & Kelley, J. (in press). Scholarly culture: How books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Social Science Research*. doi: 10.1016/j.ssresearch.2018.10.003

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453.

Soares, D. L., Lemos, G. C., Primi, R., & Almeida, L. S. (2015). The relationship between intelligence and academic achievement throughout middle school: The role of students' prior academic performance. *Learning and Individual Differences*, *41*(Supplement C), 73–78.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745–759.

Stipek, D., & Valentino, R. (2015). Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology*, *107*(3), 71.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. Oxford University Press.

Tolmie, A. K., Ghazali, Z., & Morris, S. (2016). Children's science learning: A core skills approach. *British Journal of Educational Psychology*, *86*(3), 481–497.

Verbruggen, F., & Logan, G. D. (2009). Proactive adjustments of response strategies in the stop-signal paradigm. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(3), 835–854.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence.* The Psychological Corporation: Harcourt Brace & Company. New York, NY.

Wendelken, C., Baym, C. L., Gazzaley, A., & Bunge, S. A. (2011). Neural indices of improved attentional modulation over middle childhood. *Developmental Cognitive Neuroscience*, *1*(2), 175–186.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399.