Title:

The neurocognitive architecture of fluid ability in children and adolescents

Authors:

Fuhrmann, D.[1*], Simpson-Kent, I. L.[1], Bathelt, J.[1], the CALM team [1] & Kievit, R. A.[1]

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

[*] corresponding author: Delia Fuhrmann, MRC CBU, 15 Chaucer Road, Cambridge, CB2 7EF, UK, email: delia.fuhrmann@mrc-cbu.cam.ac.uk

Abstract

Fluid ability is the capacity to solve novel problems in the absence of task-specific knowledge, and is highly predictive of outcomes like educational attainment and psychopathology. Here, we modelled the neurocognitive architecture of fluid ability in two cohorts: CALM (*N*=551, aged 5-17) and NKI-RS (*N*=335, aged 6-17). We used multivariate Structural Equation Modelling to test a preregistered 'watershed model' of fluid ability. We found that the watershed model fit the data well for both samples: White matter contributed to working memory and processing speed, which, in turn, contributed to fluid ability (R^2(CALM)=51.2%, R^2(NKI-RS)=78.3%). The relationship between cognitive abilities and white matter differed with age and showed a dip in strength around ages 7-12 years. Speculatively, this age-effect may reflect a reorganization of the neurocognitive architecture around pre- and early puberty. Overall, these findings highlight that fluid ability is part of a complex hierarchical system of partially independent effects.

Introduction

Fluid reasoning or fluid intelligence (g_f) is a core part of human cognition and refers to the capacity to solve novel problems in the absence of task-specific knowledge. It is highly predictive of a number of important life span outcomes, including educational attainment (Primi, Ferrão, & Almeida, 2010; Roth et al., 2015), psychopathology (Gale, Batty, Tynelius, Deary, & Rasmussen, 2010) and mortality (Batty, Deary, & Gottfredson, 2007; Calvin et al., 2011). However, our understanding of the neurocognitive architecture of g_f remains limited. Longstanding debates have considered, for instance, how g_f relates to more fundamental cognitive functions such as working memory and processing speed, and how all of these cognitive functions relate to brain structure and function (Chuderski, 2013; Ferrer et al., 2013; Fry & Hale, 2000; Kyllonen & Christal, 1990).

Working memory refers to the ability to hold and manipulate information in the mind short-term. It has been suggested that working memory is a key determinant of g_f by limiting mental information processing capacity (Chuderski, 2013; Fukuda, Vogel, Mayr, & Awh, 2010). Proponents of this working memory account of g_f cite high correlations between the two domains ranging from 0.5 to 0.9 in meta-analyses (Ackerman, Beier, & Boyle, 2005; Oberauer, Schulze, Wilhelm, & Süß, 2005). Such high correlations have led some to suggest that g_f and working memory are, in fact, isomorphic (Kyllonen & Christal, 1990). However, more recent work has highlighted that this isomorphism only arises under conditions of high time constraints for g_f tasks (Chuderski, 2013). This suggests that g_f and working memory are, in fact, separable constructs and underlines the importance of processing speed for g_f.

Processing speed, the speed of mental computations, is thought to be rate-limiting to g_f and is therefore often proposed to be a particularly good predictor of g_f (Ferrer et al., 2013; Kail,

Lervåg, & Hulme, 2015; Kail & Salthouse, 1994; Salthouse, 1996; Schubert, Hagemann, & Frischkorn, 2017). Proponents of the processing speed account cite moderate but robust correlations between g_f and processing speed of 0.2 in meta-analyses (Sheppard & Vernon, 2008) as well as longitudinal evidence, particularly from ageing populations (Coyle, Pillow, Snyder, & Kochunov, 2011; Finkel, Reynolds, McArdle, & Pedersen, 2005; Kail et al., 2015). Salthouse (1996) argued in the context of cognitive aging, that processing speed determines high-level cognitive performance because slow processing means that relevant sub-operations cannot be completed in a set amount of time or are not available for successful integration. A complementary explanation of individual differences in g_f proposes that processing speed may be a direct reflection of fundamental neuroarchitectonic properties of the brain, such as myelination or white matter microstructure (Chevalier et al., 2015; Lu et al., 2011).

White matter shows protracted development throughout childhood and adolescence, which extends into the third decade of life (Mills et al., 2016). White matter tracts can be characterised *in vivo* using diffusion-tensor imaging (DTI), which is sensitive, but not necessarily specific, to white matter microstructural properties such as myelination or axonal density (Jones, Knösche, & Turner, 2013; Wandell, 2016). Fractional anisotropy (FA) is the most commonly investigated DTI measure and quantifies the directionality of water diffusion in different white matter tracts (Pfefferbaum et al., 2000; Wandell, 2016). Working memory, processing speed and g_f have each been linked to individual differences in FA in several white matter tracts (Bathelt, Johnson, Zhang, & Astle, 2018; Kievit, Davis, Griffiths, Correia, CamCAN, et al., 2016; Vestergaard et al., 2011). Studies of FA have shown that g_f is associated with FA in the cingulum (Bathelt et al., 2018), that working memory is linked to the superior longitudinal fasciculus (Vestergaard et al., 2011) and cingulum (Bathelt et al.,

2018), while processing speed has been associated mainly with the forceps minor (Kievit, Davis, Griffiths, Correia, CamCAN, et al., 2016), the anterior thalamic radiation, and the uncinate fasciculus (MacPherson et al., 2017).

The literature thus shows strong bivariate links between g_f, working memory, processing speed and white matter. There are, however, several methodological and practical issues, which we seek to address here: First, there is limited systematic evidence on the concurrent relationships between g_f, working memory, processing speed and white matter. This leaves the relative contributions of processing speed and working memory to g_f unclear, which, in turn, poses challenges for the design of effective interventions for populations with learning problems and disabilities. Second, studies usually use a single task as an indicator (i.e. observed variable) as a proxy for complex and abstract constructs (i.e. latent variables) such as processing speed, working memory, and g_f. This raises questions about the generalizability of findings (Noack, Lövdén, & Schmiedek, 2014). Third, our understanding of how the relationships between relevant cognitive domains and between brain and cognition change with age remains limited (Achterberg, Peper, van Duijvenvoorde, Mandl, & Crone, 2016; Tamnes, Roalf, Goddings, & Lebel, 2017). Influential work by Garrett (1946) suggested that relationships between *cognitive abilities* may change with age such that cognitive abilities differentiate (i.e. decrease in covariance) over development (Garrett, 1946; Li et al., 2004), or, conversely, de-differentiate (i.e. increase in covariance) with age (Blum & Holling, 2017), while others still found no changes in covariance with age at all (de Mooij, Henson, Waldorp, & Kievit, 2018; Molenaar, Kő, Rózsa, & Mészáros, 2017; Tucker-Drob, 2009). Age-related changes in the relationship between *brain and cognition* is even less well understood. Johnson (Johnson, 2000, 2001, 2011) proposed the interactive specialization theory, which holds that brain maturation involves activity-dependent reorganization and specialization of

networks. Critically, interactive specialization predicts that the same cognitive function may be represented by different structures at different ages. Brain - cognition relationships may therefore change fundamentally and dynamically with age (Johnson, 2011). At present, however, there are very few empirical investigations of potential non-monotonic, dynamic changes in brain-cognition relationship in general and for the neurocognitive architecture of fluid ability in particular (de Mooij et al., 2018).

To address these issues, we here used structural equation modelling (SEM) to investigate the multivariate associations between latent variables of g_f, working memory, processing speed, and observed variables of white matter microstructure (FA) and age in two large samples: the Centre for Attention, Leaning and Memory sample (CALM, N = 551, aged 5 – 17 years (Holmes, Bryant, & Gathercole, 2018)), which consists of children referred to a clinic for having problems which attention, learning and memory, and the Enhanced Nathan Kline Institute – Rockland Sample (NKI-RS, N = 335, aged 6 – 17 years (Nooner et al., 2012)), a community-ascertained sample.

To investigate the neurocognitive architecture of g_f in a principled way, we used a watershed model of individual differences. This model was originally developed for understanding complex psychopathological phenotypes (Cannon & Keller, 2006) and then repurposed for the investigation of individual differences in g_f in an ageing sample (Kievit, Davis, Griffiths, Correia, CamCAN, et al., 2016). Based on the metaphor of a watershed, the model predicts a hierarchical many-to-one mapping of partially independent effects such that upstream tributaries (e.g. brain structure) contribute to intermediate cognitive phenomena (cognitive endophenotypes), which then contribute to downstream, complex cognitive phenomena (Cannon & Keller, 2006; Kievit, Davis, Griffiths, Correia, CamCAN, et al., 2016). Kievit et al.

(2016), for instance, showed that individual differences in white matter anatomy predicted individual differences in processing speed, which in turn predicted fluid intelligence scores and a cohort of participants aged 18 - 87 years. Here, we extend the watershed model to include working memory as well as processing speed as cognitive endophenotypes and tested this model in two heterogeneous developmental cohorts. See Figure 1 for a representation of the model. Using structural equation modeling trees (Brandmaier, von Oertzen, McArdle, & Lindenberger, 2013), a novel, decision-tree-based exploratory technique, we also investigated whether the associations in the watershed model change with age.

Based on the watershed model we made the following preregistered predictions (http://aspredicted.org/blind.php?x=u5pf6z):

1. Working memory, g_f and processing speed are separable constructs.
2. Individual differences in g_f are predicted by working memory and processing speed.
3. White matter microstructure is a multi-dimensional construct.
4. There is a hierarchical relationship between white matter microstructure, cognitive endophenotypes (working memory and processing speed) and g_f, such that white matter contributes to working memory and processing speed, which, in turn contribute to g_f.
5. The contribution of working memory and processing speed to g_f changes with age.

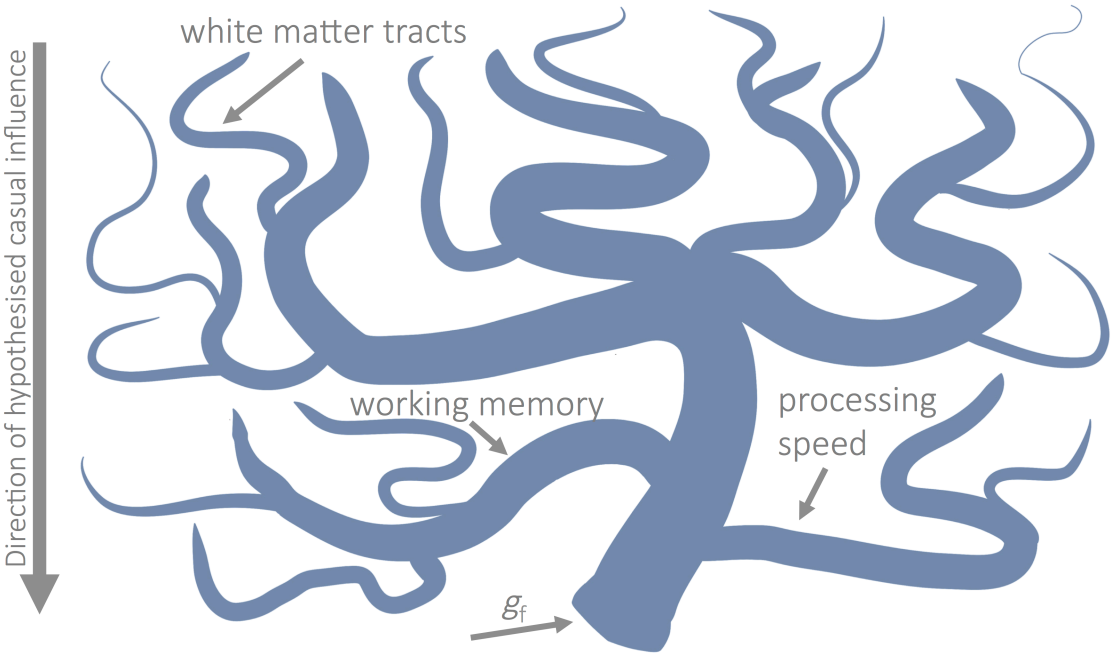

Figure 1. The Watershed Model. Schematic representation of the watershed model developed by Cannon & Keller (2006) and adapted for the present study. Fluid ability is hypothesized to be the downstream product of working memory and processing speed, which are, in turn, the product of white matter contributions. Figure adapted from Kievit et al. (2016).

Results

Measurement Models of Cognition and White Matter

Cognition. To examine the neurocognitive architecture of g_f, we built up the watershed model in steps, starting with the measurement model of g_f, working memory and processing speed. We first examined whether these three cognitive skills are best described as three separate, two separate or a single cognitive factor (Figure 2).

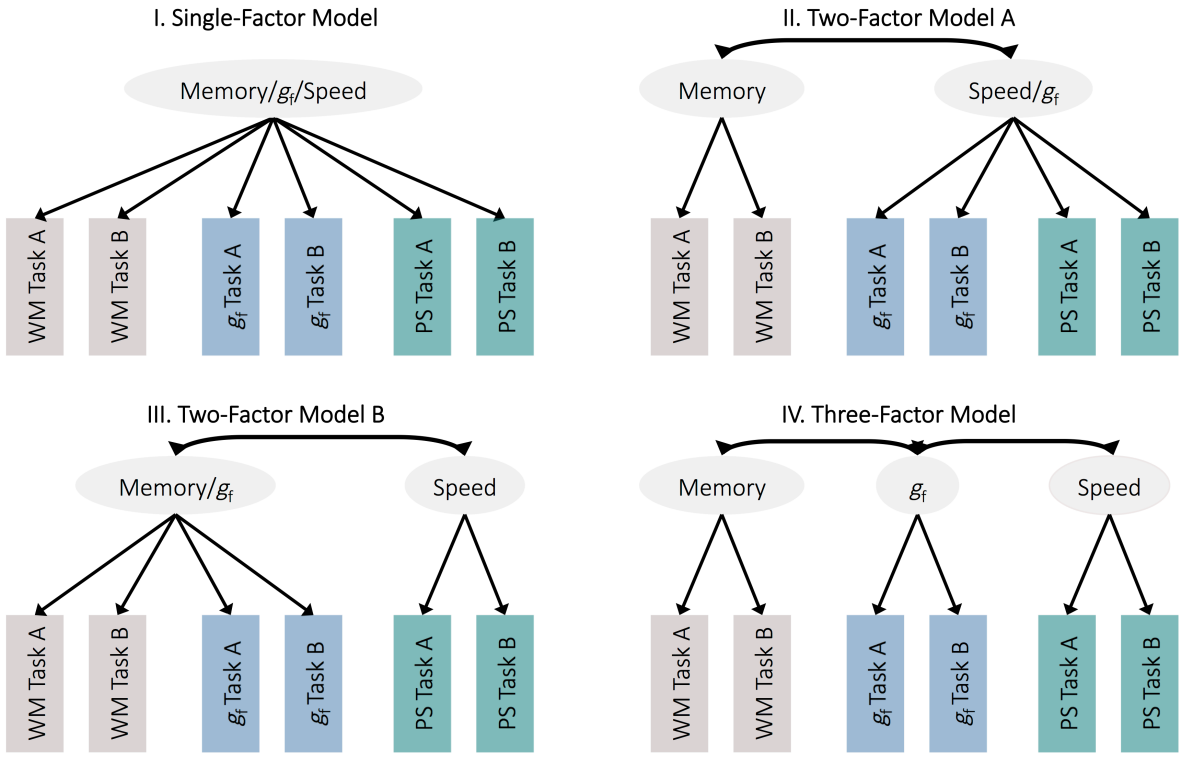

Figure 2. Different Measurement Models of Cognition. Abbreviations: WM: working memory, PS: processing speed

The Single-Factor Model (Figure 2) showed acceptable absolute fit (Table 1) but poor comparative fit: The Akaike weight (AIC_{weight}) of the Single-Factor Model was approximately 0% for both samples, indicating an extremely low probability that the Single-Factor Model was the data-generating model compared to more complex alternatives. Two-Factor Model A (Figure 2, treating processing speed and g_f as a unitary factor) showed good absolute fit (Table 1) but also approximately 0% probability of being data-generating model. Two-Factor Model B (Figure 2, treating working memory and g_f as a unitary factor) fit the data well (Table 1). Two-Factor Model B showed low probability of being the data-generating model for NKI-RS (AIC_{weight} = 3.40%) but a high probability of being the data-generating model for CALM (AIC_{weight} = 72.85%). The Three-Factor Model (Figure 2) showed 96.60% probability of being data-generating model for NKI-RS and 27.15% probability for CALM.

Table 1. Model Fit of Competing Measurement Models

Single-Factor Model	Two-Factor Model A	Two-Factor Model B	Three-Factor Model
		CALM	
$\chi^2(20) = 70.28$, $p < .001$	$\chi^2(19) = 67.99$, $p < .001$	$\chi^2(19) = 41.66$, $p = .002$	$\chi^2(18) = 41.74$, $p = .001$
RMSEA = .068	RMSEA = .068	RMSEA = .047	RMSEA = .049
[.051-.085]	[.052 - .086]	[.027 - .066]	[.030 - .068]
CFI = .963	CFI = .964	CFI = .983	CFI = .983
SRMR = .047	SRMR = .043	SRMR = .032	SRMR = .032
AIC = 9697.18	AIC = 9696.44	AIC = 9668.58	AIC = 9670.55
BIC = 9800.66	BIC = 9804.24	BIC = 9776.37	BIC = 9782.66
		NKI-R	
$\chi^2(35) = 109.96$, $p < .001$	$\chi^2(34) = 108.15$, $p < .001$	$\chi^2(34) = 64.85$, $p = .001$	$\chi^2(32) = 54.15$, $p = .009$
RMSEA = .080	RMSEA = .081	RMSEA = .052	RMSEA = .045
[.064 - .097]	[.064 - .098]	[.033 - .071]	[.024 - .065]
CFI = .936	CFI = .936	CFI = .974	CFI = .981
SRMR = .045	SRMR = .044	SRMR = .035	SRMR = .030
AIC = 7155.64	AIC = 7155.74	AIC = 7109.43	AIC = 7102.74
BIC = 7270.07	BIC = 7273.98	BIC = 7227.67	BIC = 7228.60

Note. See Figure 2 for the configuration of different models. Abbreviations: root mean square error of approximation (RMSEA), comparative fit index (CFI), standardized root mean square residual (SRMR), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC)

Even though working memory, processing speed and g_f were highly correlated in both samples (Table 2), this pattern of results indicates that processing speed formed a clearly separable factor from working memory and g_f in both samples. Working memory and g_f were also clearly separable in NKI-RS, but not in CALM. To facilitate comparison across samples and in accordance with our preregistered analysis plan we used the Three-Factor Model (Table 2, Supplementary Table 1) in all subsequent analyses.

Table 2. Covariance between Cognitive Measures in the Three-Factor Model

Sample	Path	Standardized Estimate
CALM	g_f <-> memory	0.71, $z = 28.42$, $p < .001$
	g_f <-> speed	0.55, $z = 12.20$, $p < .001$
	memory <-> speed	0.79, $z = 19.35$, $p < .001$
NKI-RS	g_f <-> memory	0.91, $z = 19.51$, $p < .001$
	g_f <-> speed	0.81, $z = 24.73$, $p < .001$
	memory <-> speed	0.87, $z = 17.43$, $p < .001$

Note. See Supplementary Table 1 for factor loadings.

White matter. We examined whether white matter microstructure could be adequately captured by a single factor by examining model fit. The single-factor model of white matter microstructure did not fit the data well (CALM: $\chi^2(35) = 124.63$, $p < .001$; RMSEA = .125 [.103 - .147]; CFI = .933; SRMR = .039; NKI-RS: $\chi^2(35) = 132.33$, $p < .001$; RMSEA = .204 [.167 - .242]; CFI = .885; SRMR = .023). This indicates that white matter microstructure could not be reduced to a single 'global FA' dimension in our samples, in line with Lövdén et al. (2013) and Kievit et al. (2016). We therefore modelled each of the ten white matter tracts separately in all subsequent models.

The Relationship between Working Memory, Processing Speed and g_f

Using Multiple Indicators, Multiple Causes (MIMIC) models (Jöreskog & Goldberger, 1975; Kievit et al., 2012), we examined the relationship between proposed cognitive endophenotypes (processing speed and working memory) and g_f. The MIMIC model fit well for both samples (CALM: $\chi^2(18) = 41.74$, $p = .001$; RMSEA = .049 [.030 - .068]; CFI = .983; SRMR = .032, NKI-RS: $\chi^2(32) = 54.15$, $p = .009$; RMSEA = .045 [.024 - .065]; CFI = .981; SRMR = .030). To understand whether working memory and processing speed made significant joint and independent contributions to g_f, we compared a freely-estimated MIMIC model to one in which regression paths from processing speed and working memory to g_f were constrained to zero. The freely-estimated MIMIC model fit better than a model in which these regression paths were constrained to zero (CALM: $\Delta\chi^2(2) = 250.20$, $p < .001$; NKI-RS: $\Delta\chi^2(2) = 199.67$, $p < .001$). This indicates that working memory and processing speed jointly contributed to g_f. In CALM ($\Delta\chi^2(1) = 15.53$, $p < .001$), but not NKI-RS ($\Delta\chi^2(1) = 3.25$, $p = .072$), the freely-estimated MIMIC model also fit better than a model in which paths from processing speed and working memory to g_f were constrained to be equal. Working memory showed a greater effect (as

indicated by standardized path estimates) than processing speed in both samples, but the difference was more pronounced in CALM (Table 3). Contrary to our prediction, therefore, processing speed seemed to make a negligible contribution to g_f above and beyond working memory.

Table 3. Regression Paths in the MIMIC Model.

Sample	Path	Standardized Estimate
CALM	speed -> g_f	-0.01, z = -0.16, p = .876
	memory -> g_f	0.72, z = 7.65, p < .001
NKI-RS	speed -> g_f	0.06, z = 0.21, p = .208
	memory -> g_f	0.86, z = 1.81, p = .070

The Watershed Model: Relationships between Cognition and White Matter

Next, we built the full watershed model including white matter, working memory, processing speed and g_f. We examined path estimates and alternative configurations of the model to assess whether the main prediction from the watershed model holds, namely, that white matter contributes to working memory capacity and processing speed, which, in turn, contribute to g_f.

We found largely converging results across samples. The watershed model showed good fit in CALM ($\chi^2(78)$ = 107.78, p = .014; RMSEA = .026 [.012 - .038]; CFI = .981; SRMR = .043) and acceptable fit in NKI-RS ($\chi^2(112)$ = 219.22, p < .001; RMSEA = .053 [.043 - .064]; CFI = .928; SRMR = .088). White matter explained large amounts of variance in working memory (R^2_{CALM} = 32.3%; $R^2_{NKI\text{-}RS}$ = 46.1%) and processing speed (R^2_{CALM} = 38.2%; $R^2_{NKI\text{-}RS}$ = 54.4%), which, in turn, explained even more variance in g_f (R^2_{CALM} = 51.2%; $R^2_{NKI\text{-}RS}$ = 78.3%).

Comparing the freely estimated watershed model to alternative, constrained, models showed that white matter contributed significantly to memory and processing speed.

Specifically, a model in which paths from white matter to processing speed were constrained to zero fit worse than the freely-estimated model (CALM: $\Delta\chi^2(10) = 50.26$, $p < .001$; NKI-RS: $\Delta\chi^2(10) = 27.19$, $p = .002$), as did a model in which paths from white matter to working memory were constrained to zero (CALM: $\Delta\chi^2(10) = 52.26$, p < .001; NKI-RS: $\Delta\chi^2(10) = 25.85$, $p = .004$).

White matter tracts differed in their relationship to cognition. A model in which paths from white matter to cognition were constrained to be equal, fit worse than the freely-estimated model for CALM ($\Delta\chi^2(18) = 47.76$, $p < .001$) and NKI-RS ($\Delta\chi^2(18) = 30.42$, $p = .034$), indicating that the role of white matter microstructure in supporting working memory and processing speed differed across tracts. For CALM, only the anterior thalamic radiation contributed significantly to processing speed while the superior longitudinal fasciculus, forceps major and cingulum were significantly and positively related to working memory (Figure 3). For NKI-RS, the superior longitudinal fasciculus was significantly and positively related to processing speed and working memory (Figure 4). Two tracts showed an unexpected, strongly negative (< -1), relationship: the forceps minor for CALM and the inferior fronto-occipital fasciculus for NKI-RS. We found that these negative estimates occurred only when all other brain to cognition pathways were also estimated: When estimated on their own, path estimates became positive (forceps minor to working memory: standardized estimate = 0.36, $z = 4.05$, $p < .001$; inferior fronto-occipital fasciculus to working memory: standardized estimate = 0.14, $z = 0.859$, $p = .390$; inferior fronto-occipital fasciculus to processing speed: standardized estimate = 0.26, $z = 1.41$, $p = .158$). This pattern of results suggests that the negative pathways were potentially due to modelling several, highly-correlated paths at the same time (Jöreskog, 1999).

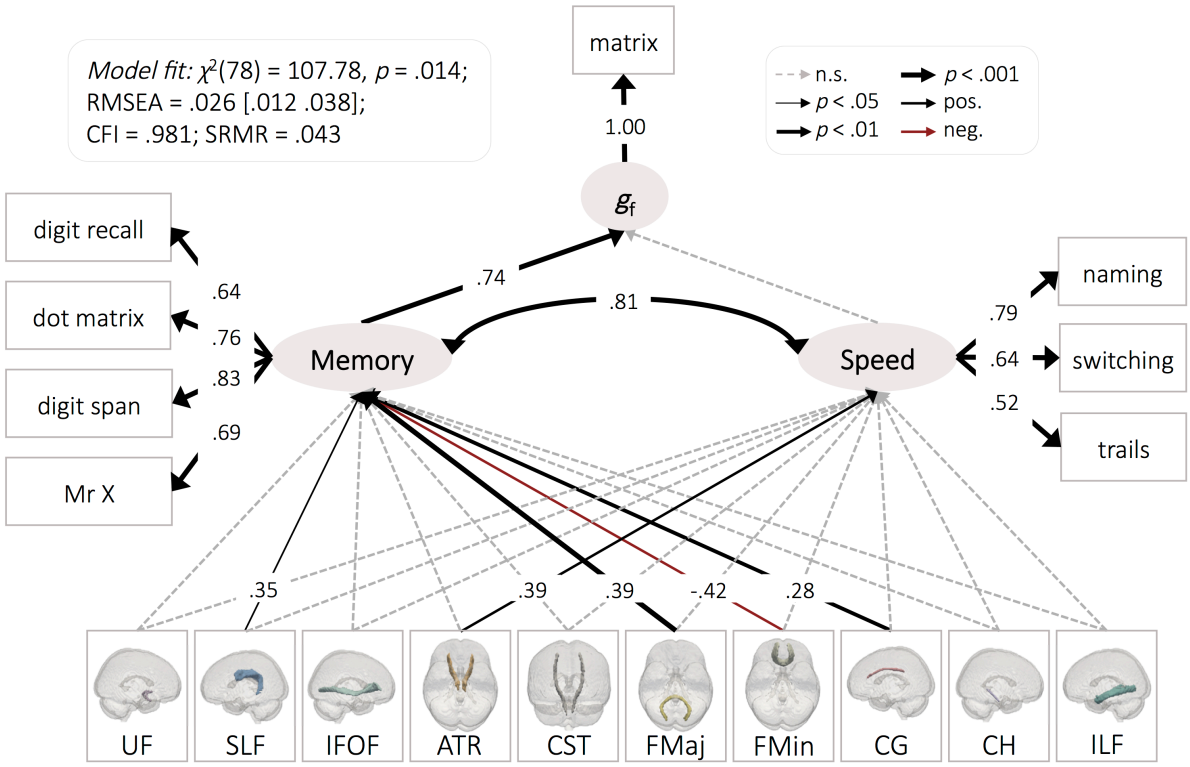

Figure 3. The Watershed Model in CALM. See Supplementary Table 3 for regression estimates. Residual covariances between white matter tracts were allowed but are not shown for simplicity. Abbreviations: uncinate fasciculus (UNC), superior longitudinal fasciculus (SLF), inferior fronto-occipital fasciculus (IFOF), anterior thalamic radiations (ATR), cerebrospinal tract (CST), forceps major (FMaj), forceps minor (FMin), dorsal cingulate gyrus (CG), ventral cingulate gyrus (CH), inferior longitudinal fasciculus (ILF).

Testing alternative hierarchies in the model showed that the watershed model had a high probability of being the data-generating model for CALM (AIC_{weight} = 49.64%) but not NKI-RS (AIC_{weight} = 0.10%). For NKI-RS, a model with an additional level, such that g_f was regressed on working memory, which was regressed on processing speed, which was then regressed on white matter showed the highest probability (AIC_{weight} = 96.95%). This model showed a low probability for CALM, however (AIC_{weight} = 0.19%). For CALM, reversing the hierarchy of levels of the watershed model such that white matter was regressed on processing speed and working memory, which are then regressed on g_f, showed as similarly high probability as the original watershed model (AIC_{weight} = 49.64%). This was not the case for NKI-RS (AIC_{weight} = 0.10%). A model, in which the hierarchy between cognitive endophenotypes and g_f was

inverted, showed comparatively poor fit for CALM (AIC_{weight} = 0.00%) and NKI-RS (AIC_{weight} = 2.85%). Similarly, a model in which g_f was directly regressed on white matter, working memory and processing speed, showed no clear advantage over the watershed model for CALM (AIC_{weight} = 0.53%) or NKI-RS (AIC_{weight} = 0.00%).

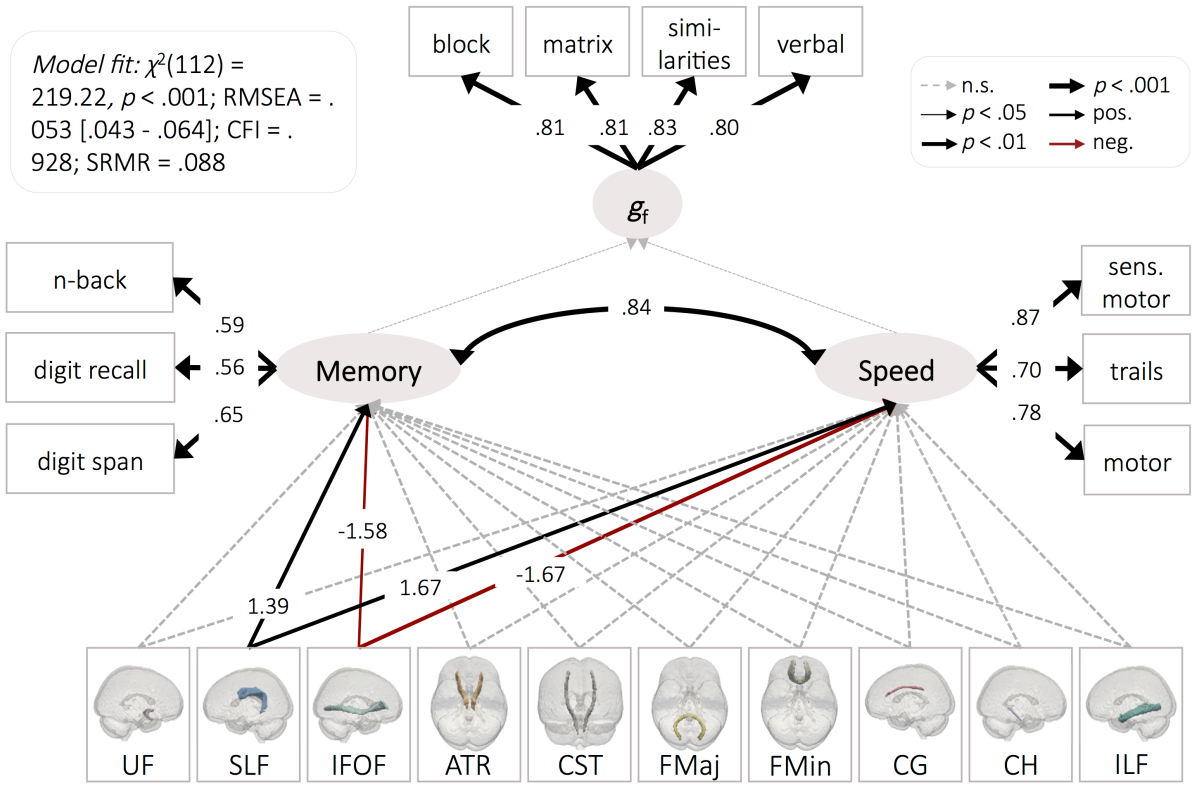

Figure 4. The Watershed Model in NKI-RS. See Supplementary Table 4 for regression estimates. Residual covariances between white matter tracts were allowed but are not shown for simplicity.

We carried out a series of supplementary and exploratory analyses to examine whether possible confounders influenced our models. These analyses showed that our findings were robust to the inclusion of covariates such as scanner motion or socio-economic status. They were also robust across genders and participants taking or not taking medication. There were no differences between participants with and without diagnosed disorders for CALM, while potential differences cannot be ruled out for NKI-RS, likely due to the low number of diagnosed participants (N = 106) (see Supplementary Analyses).

In summary, we found that the watershed model performed well for both cohorts and explained large amounts of variance in g_f. As predicted by the watershed model, white matter contributed to working memory and processing speed, which, in turn, contributed to g_f.

Age-Related Differences in the Neurocognitive Architecture of g_f

We investigated whether there were age-related differences in the neurocognitive architecture of g_f. An examination of cross-sectional age differences in g_f, working memory and processing speed showed that factor scores increased with age for all three cognitive phenotypes (Figure 5).

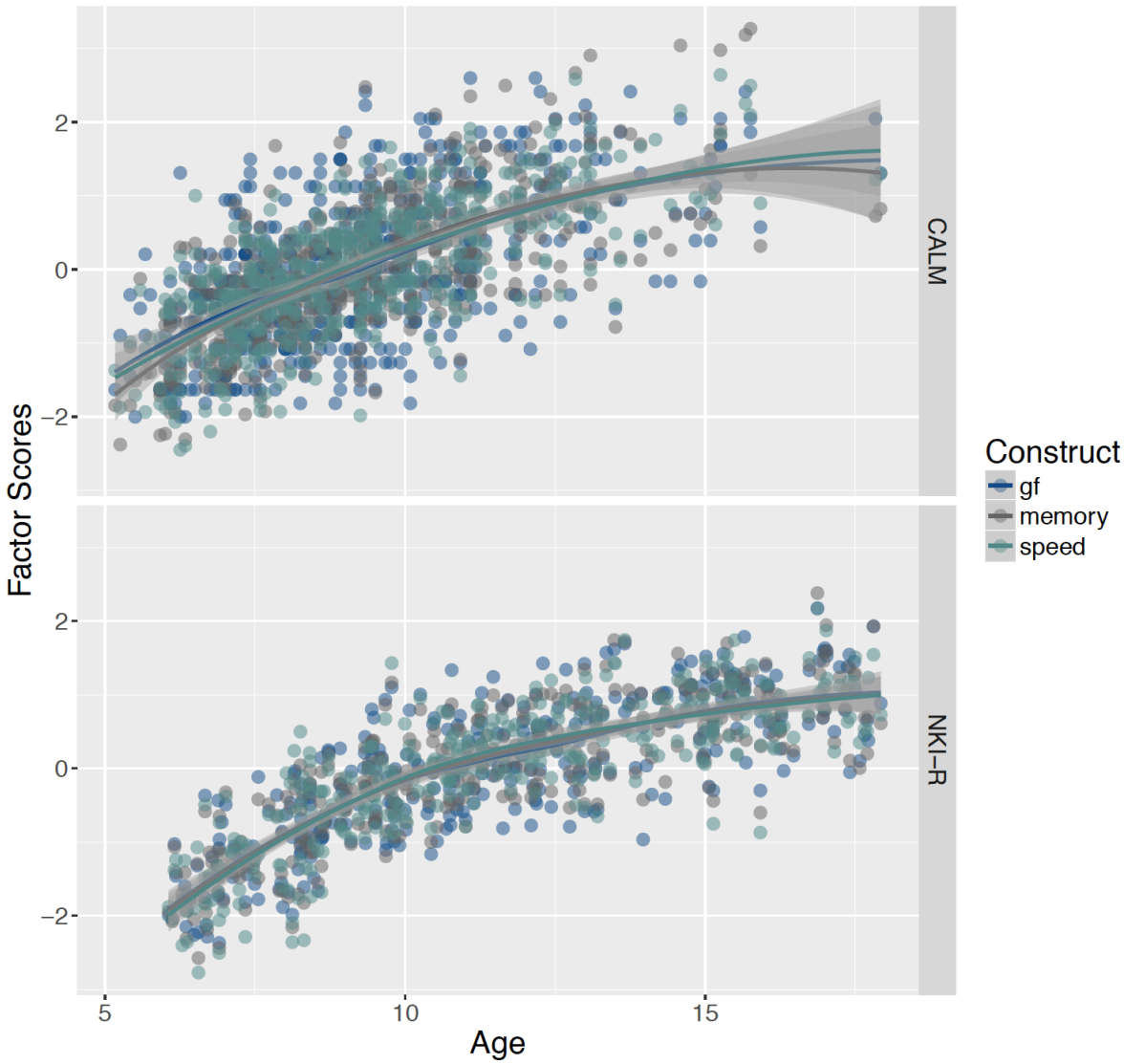

Figure 5. Cognitive Factor Scores by Age.

Using SEM trees, we investigated whether the associations among cognitive and neural dimensions varied with age. As per our preregistration, we investigated potential age-differences in the relationships between g_f, working memory and processing speed. In additional, exploratory, analyses we investigated potential age differences in significant watershed paths between white matter and cognitive endophenotypes (Figure 3 and 4).

SEM trees combine SEMs with decision tree methods, whereby separating a dataset into subgroups (in this case age groups) if SEM parameter estimates of interest differ sufficiently (Brandmaier et al., 2013). SEM trees allowed us to investigate age as a potential moderator without imposing a-priori categorical age splits. It is an inherently data-driven technique and the results obtained from SEM trees depend on parameters passed to the algorithm such as the minimum number of individuals allowed per group (here N = 50) or the number of levels per branch of the tree allowed, which determines the maximum number of possible subgroups. To counterbalance the inherently exploratory nature of SEM Trees and maximise the robustness of our results, we implemented cross-validated model selection and a conservative, Bonferroni-corrected alpha level of .001. We initially allowed for a maximum depth of one level per tree branch, which could result in no more than two age groups. This yielded inconsistent results for CALM and NKI-RS (Supplementary Table 4). To test whether these inconsistencies were an artefact of allowing for only two groups, we repeated our analysis and allowed for a maximum depth of two levels per branch, which could result in up to four age groups. This analysis yielded consistent results between CALM and NKI-RS (Table 4). This pattern of results indicates that the initial parameters of our analysis caused us to miss relevant age differences.

Table 4. SEM Tree Results for the Watershed Model. The table shows differences in parameter estimates for paths of interest (as shown in Figure 3 and 4) depending on participants' age in years.

Path	Est. Before	Age Split 1	Est. Betw.	Age Split 2	Est. Betw.	Age Split 3	Est. After
			CALM				
memory <–> speed	0.85	**8.46**	0.97	**9.46**	0.74	-	-
memory –> g_f	0.83	**9.38**	0.42	**10.04**	1.14	**10.88**	0.94
speed –> g_f	0.04	**6.88**	-0.19	**11.21**	0.17	-	-
SLF –> memory	0.67	**7.21**	0.18	**11.21**	0.76	-	-
FMaj –> memory	0.59	**7.71**	0.14	**9.29**	0.33	**11.13**	0.74
CG –> memory [1]	0.64	**6.96**	0.09	**11.04**	0.70	-	-
ATR –> speed	0.96	**7.13**	0.68	**7.96**	0.17	**11.96**	0.65
			NKI-RS				
memory <–> speed	0.90	**9.82**	0.48	**14.72**	1.11	-	-
memory –> g_f	1.10	**8.59**	0.59	**12.67**	1.03	-	-
speed –> g_f	0.53	**8.59**	-0.12	**12.96**	0.52	-	-
SLF –> memory	2.15	**8.30**	1.47	**12.15**	1.93	-	-
SLF –> speed	3.12	**8.63**	1.83	**15.09**	2.31	-	-

As shown in Table 4, we observed age-related shifts in brain-behaviour associations throughout childhood and adolescence. For both samples and all but one path, there was an initially strong relationship between variables, then a dip around ages 7 - 9 for CALM and age 8 for NKI-RS, followed by an increase in path strength around ages 11 to 12. Speculatively, this pattern of results is consistent with an interpretation of a reorganization of neurocognitive faculties in late childhood, followed by a consolidation of neurocognitive pathways around the onset of adolescence.

Discussion

Here we used multivariate statistical techniques to investigate the neurocognitive architecture of g_f in two large developmental cohorts and, for the first time, investigated how the neurocognitive architecture of g_f changes dynamically with age. We tested a preregistered watershed model of g_f, which predicts a hierarchy of partially independent effects. We found that the watershed model performed well for both the atypical (CALM) and typical cohort (NKI-RS): White matter contributed to working memory and processing speed, which, in turn, contributed to g_f and explained 51% of variance therein for the CALM sample and 78% of variance for the NKI-RS. The watershed model performed well against most alternative models tested. Models were robust across genders, participants taking or not taking medication and when controlling for socio-economic status and scanner motion. Investigations of age effects showed that the relationship between cognitive abilities and white matter dipped in strength around ages 7 - 12. Speculatively, this age-effect may reflect a reorganization of the neurocognitive architecture during pre-puberty and early puberty (Byrne et al., 2017). These findings have implications for understanding and targeting cognitive impairments in populations with learning difficulties.

Our watershed model consisted of three levels: g_f was conceptualized as the most down-stream point, with working memory and processing speed as intermediate tributaries, and white matter microstructural tracts as upstream sources. Previous studies suggested that matter microstructure is best characterised by a single, 'global FA' factor (Penke et al., 2010) while others have contended that association patterns among different white matter tracts are more complex (Kievit, Davis, Griffiths, Correia, CamCAN, et al., 2016; Lövdén et al., 2013). Here we found strong evidence for a multifactorial view of white matter tracts – for both

samples, a unidimensional model of white matter fit poorly, and for CALM multiple tracts also showed partially independent contributions to distal cognitive outcomes. In line with previous research (Bathelt et al., 2018; Kievit, Davis, Griffiths, Correia, Cam-CAN, et al., 2016; MacPherson et al., 2017), we found that the anterior thalamic radiation was related to processing speed, as were the forceps major, forceps minor and the cingulum to working memory. Of note, the superior longitudinal fasciculus was consistently associated with working memory in both samples. For NKI-RS, the superior longitudinal fasciculus was also associated with processing speed. The superior longitudinal fasciculus is a large, bilateral association fibre connecting temporal, occipital, parietal and frontal regions (Kamali, Flanders, Brody, Hunter, & Hasan, 2014). It may therefore be particularly suited to contributing to cognitive processes that rely on general multiple-demand systems such as g_f (Parlatini et al., 2017).

The cognitive levels of the watershed model highlighted a close relationship between working memory and g_f. Previous studies had variably suggested that g_f and working memory (Fukuda et al., 2010; Kyllonen & Christal, 1990), or g_f and processing speed (Ferrer et al., 2013; Kail et al., 2015; Kail & Salthouse, 1994; Salthouse, 1996; Schubert et al., 2017) may be most closely related (but see Chuderski, 2013). We found that all three cognitive factors were highly correlated for both samples. Nonetheless, processing speed formed a cognitive factor clearly separable from working memory and g_f. Working memory and g_f, in turn, were separable in NKI-RS but not in the CALM sample. This close relationship between g_f and working memory was also evident in other models where processing speed and working memory were used as predictors of g_f. Working memory showed a stronger relationship to g_f than processing speed. Contrary to the predictions of the watershed model, processing speed became non-significant after controlling for working memory for CALM. This indicates

that working memory was more strongly related to g_f than processing speed and supports the notion that mental information processing capacity is a key determinant of g_f (Fukuda et al., 2010; Kyllonen & Christal, 1990).

The associations in the watershed model differed between ages in a complex, non-monotonic fashion. Previous research had suggested either a decrease in covariance among cognitive domains with age (age differentiation) (Garrett, 1946; Li et al., 2004), an increase in covariance with age (age de-differentiation) (Blum & Holling, 2017), or no changes with age (de Mooij et al., 2018; Molenaar et al., 2017; Tucker-Drob, 2009). These investigations have traditionally focussed on relations between cognitive domains, however, not on relationships between brain and cognition (although see de Mooij et al., 2018 for a recent example in an aging cohort). Possible linear and non-linear changes in brain-behaviour mapping with age have remained mostly unexplored (Tamnes et al., 2017). Using structural equation modelling trees, a novel decision-tree-based exploratory technique, we here found evidence of complex developmental differences that were consistent across samples and relationships in the watershed model: Initially strong path estimates showed a pronounced decrease in strength around ages 7 - 9, followed by a renewed increase in the strength, even surpassing initial levels, around ages 10 - 15. In line with the interactive specialization theory (Johnson, 2000, 2001, 2011), we speculate that this dip may reflect a remapping of the relationships between brain substrates and cognitive abilities during late childhood and early adolescence, manifesting in weaker covariance between cognitive domains and between brain and cognition.

On a physiological level, this effect could be driven by neuroendocrine changes during pre- and early puberty. Puberty is driven by a complex and only partially understood set of hormonal events including gonadarche and andrenarche (Blakemore, Burnett, S., & Dahl, R.

382 E., 2010; Sisk & Zehr, 2005). Gonadarche begins with the secretion of gonadotropin-releasing
383 hormone from the hypothalamus around ages 10 - 11 and closely tracks the overt bodily
384 changes of puberty (Dorn, 2006). Andrenarche, beginning with the maturation of the
385 andrenal gland, starts as early as six years of age, and is increasingly recognized as a
386 complimentary driver of puberty and brain development (Byrne et al., 2017). It is possible
387 that the hormonal changes of andrenarche and early gonadarche may lead to a level of
388 neural reorganization, which may initially weaken relationships in the watershed model. The
389 sweeping bodily, social and cognitive changes happening in early adolescence may then drive
390 a consolidation of the neurocognitive architecture of g_f.

391 On a more general level, this result highlights the existence of potential non-linear changes in
392 brain-behaviour mapping during childhood and adolescence and underlines the value of
393 modern statistical approaches, such as SEM Trees, for the study of age-related differences.
394 Nonetheless, it is worth noting that this exploratory result will need to be replicated in future
395 confirmatory studies with fine-grained data on puberty and larger sample sizes. The latter will
396 also allow for detailed investigations of potential gender differences.

397 Testing our model in two different samples allowed us to address several critical questions:
398 First, participants from both samples completed a small set of common and a larger set of
399 different cognitive tasks. Therefore, the results obtained here are likely not task-specific, but
400 rather can be expected to generalize to the domains of working memory, processing speed
401 and g_f (Noack et al., 2014). Second, by comparing an atypical (CALM) and a typical (NKI-RS)
402 sample, we demonstrated that the watershed model performed well for different
403 populations. We are not able to make general claims about potential differences between
404 typical and atypical populations, however. CALM and NKI-RS were collected in countries with

different socio-economic conditions (United Kingdom and United States of America) and were of different sample sizes. It will therefore be necessary to replicate these findings in large typical and atypical cohorts collected in the same setting. Another limitation of our study is that our samples were cross-sectional, and not longitudinal. This means that were able to investigate individual differences in g_f, but we could not assess intra-individual changes during childhood and adolescence. Although the relatively narrow age range makes large cohort effects unlikely, it may still be that there were differences in recruitment and selection that varied across the age range. As such, the cross-sectional nature of our samples limits our ability to make inferences about developmental dynamics and causality. The watershed model predicts that white matter changes would precede changes in cognitive endophenotypes, which then precede changes in g_f, but this remains to be tested by future longitudinal studies such as ABCD (Volkow et al., 2018).

The findings from our study have implications understanding and targeting cognitive impairments in populations with learning difficulties. First, the close relationship between working memory and g_f found here, and in other studies (Fukuda et al., 2010; Kyllonen & Christal, 1990), indicates that children and adolescents struggling with reasoning tasks in school are likely to also struggle in terms of working memory. Either reducing working memory load or training working memory and fluid ability capacity in such populations may therefore be promising lines of inquiry for intervention studies. It is worth highlighting, however, that cognitive training studies have so far shown little generalizability: Training abstract reasoning, a common measure of g_f, has not resulted in robust increases in working memory (Knoll et al., 2016) and working memory training has not been shown to transfer to reasoning skills or school performance (Dunning, Holmes, & Gathercole, 2013; Schwaighofer, Fischer, & Bühner, 2015). Similarly, transfer from processing speed to reasoning seems to be

429 limited (Mackey, Hill, Stone, & Bunge, 2011). The results obtained here suggest that
430 interventions may increase their chance of success by implementing programs of sufficient
431 complexity to affect the entire neurocognitive architecture of effects (see also Kievit et al.,
432 2016). The level of intensity required to produce sustained benefits may also need to be
433 similarly demanding as education itself, which shows robust effects in increasing general
434 cognitive abilities over time (Ritchie & Tucker-Drob, 2018). This work and work by others
435 (Noack et al., 2014) also highlights the value of assessing, modeling and potentially
436 intervening on multiple tasks, rather than relying on a single task to capture complex
437 cognitive domains such as g_f. Finally, the age-related differences in the relationships of the
438 watershed model observed here also highlight that some interventions may work best at
439 particular developmental phases.

440 # Materials and Methods

441 ## Samples

442 We analysed data from the CALM and NKI-RS sample, as described in detail by (Holmes et al.,
443 2018) and (Nooner et al., 2012) respectively. We had also preregistered to analyse data from
444 the ABCD cohort (Volkow et al., 2018). The latter cohort contains only data for 9 - and 10 -
445 year olds at present, however, which limits comparability to CALM and NKI-RS, and makes it
446 unsuitable for investigations of developmental differences. We therefore opted to not
447 analyse ABCD data here and instead recommend a replication of the analyses presented here
448 in ABCD once longitudinal data is available. The CALM sample consists of children and
449 adolescents referred by health and educational professionals as having difficulties in
450 attention, learning and/or memory. The NKI-RS is a community-ascertained, lifespan sample,
451 and representative of the general population of Rockland, New York, and the United States as

452 a whole, in terms of ethnicity, socioeconomic status etc. For NKI-RS, we included data for
453 participants under the age of 18 only to match the age range of CALM and excluded data that
454 were completed more than half a year after enrolment. The latter was implemented to
455 ensure that age at assessment did not differ substantively between cognitive measures. The
456 final samples included 551 participants from CALM (30.85% female, aged 5.17 – 17.92,
457 $N_{\text{Neuroimaging}}$ = 165) and 335 participants from NKI-RS (43.48% female, aged 6.06 – 17.92,
458 $N_{\text{Neuroimaging}}$ = 67). See Table 5 for prevalence of relevant disorders and problems in the
459 samples.

460 *Table 5. Prevalence of Relevant Disorders and Problems*

Variable	Percentage CALM	Percentage NKI-RS
ADHD	31.94	17.01
Dyslexia	5.81	5.67
Autism	6.72	0.60
Mood disorder	0.54	0.90
Anxiety disorder	2.36	18.21
Medicated[1]	10.53	17.01
Speech/language problems	38.11	19.40

461 *Note.* [1] unspecified medication for NKI-RS, ADHD-medication for CALM

462 ## Cognitive Tasks

463 We included cognitive tasks measuring the domains of g_f, working memory or processing
464 speed for CALM and NKI-RS. See Table 6 for the complete list of tasks used and the
465 Supplementary Material for detailed task descriptions. See Supplementary Figure 1 and 2 for
466 raw scores on all tasks. The tasks modelled here were preregistered for CALM but not NKI-RS.

467 ## White Matter Microstructure

468 We modelled mean FA for ten tracts of the Johns Hopkins University (JHU) white matter
469 tractography atlas (Hua et al., 2008) averaged over the hemispheres (Figure 5). See
470 Supplementary Methods for details of MRI data acquisition, quality control and DTI data
471 processing.

472 *Table 6. Cognitive Tasks Modelled*

	CALM	NKI-RS
Working memory	AWMA Digit Recall (forward digit span)	WISC-R Forward Digit Span
	AWMA Backward Digit Span	WISC-R Backward Digit Span
	AWMA Dot Matrix	-
	AWMA Mr X	-
	-	CNB N-back task
g_f	WASI-II Matrix Reasoning	WASI-II Matrix Reasoning
	-	WASI-II Block Design
	-	WASI-II Similarities
	-	CNB Verbal Reasoning
Processing speed	DKEFS Trail-Making	DKEFS Trail-Making
	PhAB Rapid Naming	-
	TEA-Ch RBBS	-
	-	CNB Motor Speed
	-	CNB Sensory Motor Speed

473 *Note.* Abbreviations: AWMA - Automated Working Memory Assessment (Alloway, 2007), CNB - Computerized
474 Neurocognitive Battery (Gur et al., 2001), DKEF - Delis-Kaplan Executive Functioning System (Delis, Kramer,
475 Kaplan, & Holdnack, 2004), PhAB – Phonological Assessment Battery (Gallagher & Frederickson, 1995), TEA-Ch
476 RBBS - Test of Everyday Attention for Children, Red & Blues, Bags & Shoes subscale (Manly et al., 2001), WASI -
477 Wechsler Abbreviated Scale of Intelligence - Second Edition(Wechsler, 2011), WISC-R - Wechsler Intelligence
478 Scale for Children - Revised (Kaufman, 1975).

479 Data Processing and Structural Equation Modelling

480 Covariance matrices and scripts replicating key analyses can be obtained from:
481 https://github.com/df1234/gf_development. We modelled raw scores for g_f and working
482 memory tasks, as preregistered. Raw scores on processing speed tasks were transformed,
483 which was not preregistered. First, we inverted response time scores to obtain more intuitive
484 measures of speed for all but the CNB Motor Speed task, for which raw scores were already a
485 measure of speed. To satisfy SEM assumptions of normality, we additionally applied a log-

transformation to reaction time tasks. For the CNB Motor Speed task only, we removed values ± 2 *SD* of the mean (*N* = 6) because of the presence of outliers.

We modelled the associations between cognition and white matter microstructure using Confirmatory Factor Analysis and SEM in R (R core team, 2015) using the package lavaan (Rosseel, 2012). All models were fit using maximum likelihood estimation with robust Huber-White standard errors and a scaled test statistic. Missing data was addressed using full information maximum likelihood estimation. Model fit was inspected using the chi-square test, RMSEA, CFI and SRMR. Good fit was defined as approximately RMSEA < 0.05, CFI > 0.97 and SRMR < 0.05; acceptable fit as approximately RMSEA = 0.08 - 0.05, CFI = 0. 95 - 0.97, SRMR = 0.05 - 0.10 (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Comparative model fit for nested models was assessed using the chi-square test. Non-nested models were compared using the AIC_{weight}, which indicates the probability of a model being the data-generating model compared to all other models tested (Wagenmakers & Farrell, 2004). Effect sizes were evaluated by inspecting R^2 and standardized parameter estimates manifest variables. Absolute standardized parameter estimates above 0.10 were defined as small effects, 0.20 as typical and 0.30 as large (Gignac & Szodorai, 2016).

We used SEM Tress to investigate whether the associations among cognitive and neural measures differed with age. SEM Trees use decision tree methods to hierarchically split a dataset into subgroups if parameter estimates differ significantly based on a covariate of interest - in this case age (Brandmaier et al., 2013). We first ran the watershed model in OpenMx (Boker et al., 2011) and then passed this model object to semtree to compute the SEM Trees. We ran one SEM Tree for each parameter of interest (e.g. the covariance between working memory and processing speed). All other parameters in each semtree

object were set to be invariant across groups to ensure that splits were specific to the parameter of interest. We used a 10-fold cross-validation estimation method as recommend by Brandmaier et al. (2013). For the path from the cingulate to working memory only we used 5-fold cross-validation because the model did not converge using 10-fold cross-validation. Minimum sample size in age group was set to $N = 50$ to ensure reliable estimation of standard errors. Note that this choice effectively limited search space for potential splits to ages 6.58 - 12.42 years for CALM and 8.08 - 15.49 years for NKI-RS

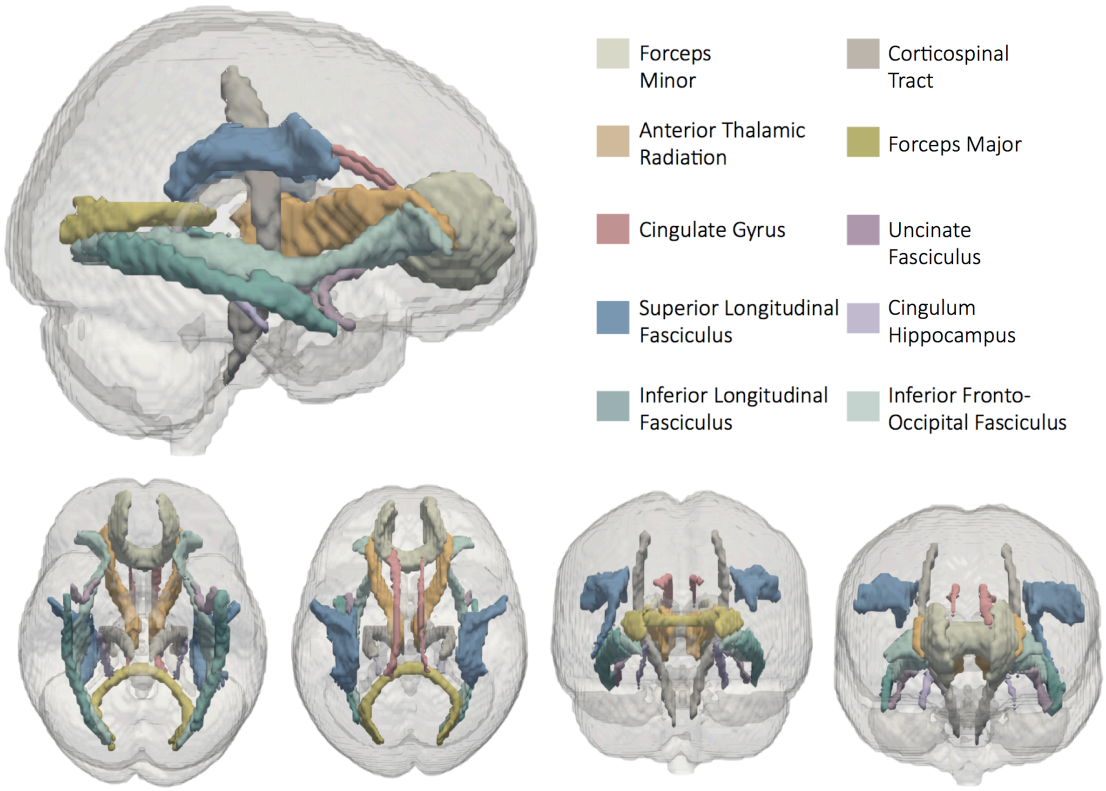

Figure 5. White Matter Tracts Modelled in the Analyses.

Acknowledgements

The Centre for Attention Learning and Memory (CALM) research clinic is based at and supported by funding from the MRC Cognition and Brain Sciences Unit, University of Cambridge. The Principal Investigators are Joni Holmes (Head of CALM), Susan Gathercole

(Chair of CALM Management Committee), Duncan Astle, Tom Manly and Rogier Kievit. Data collection is assisted by a team of researchers and PhD students at the CBU that includes Sarah Bishop, Annie Bryant, Sally Butterfield, Fanchea Daily, Laura Forde, Erin Hawkins, Sinead O'Brien, Cliodhna O'Leary, Joseph Rennie, and Mengya Zhang. The authors wish to thank the many professionals working in children's services in the South-East and East of England for their support, and to the children and their families for giving up their time to visit the clinic. We would also like to thank all NKI-RS participants and researchers. We are grateful to Amber Ruigrok for helpful suggestions regarding pubertal development.

Conflicting Interests

The authors report no conflict of interest.

References

Achterberg, M., Peper, J. S., van Duijvenvoorde, A. C. K., Mandl, R. C. W., & Crone, E. A. (2016). Frontostriatal white matter integrity predicts development of delay of gratification: A longitudinal study. *The Journal of Neuroscience*, *36*(6), 1954. https://doi.org/10.1523/JNEUROSCI.3459-15.2016

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*(1), 30–60. https://doi.org/10.1037/0033-2909.131.1.30

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London, UK: The Psychological Corporation.

Bathelt, J., Johnson, A., Zhang, M., & Astle, D. E. (2018). Data-driven brain-types and their cognitive consequences. *BioRxiv*. https://doi.org/10.1101/237859

Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and later mortality risk: Systematic review. *Annals of Epidemiology*, *17*(4), 278–288. https://doi.org/10.1016/j.annepidem.2006.07.010

Blakemore, S. J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping*, *31*(6), 926–933. https://doi.org/10.1002/hbm.21052

Blum, D., & Holling, H. (2017). Spearman's law of diminishing returns. A meta-analysis. *Intelligence*, *65*, 60–66. https://doi.org/10.1016/j.intell.2017.07.004

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., … Fox, J. (2011). OpenMx: An open source extended Structural Equation Modeling framework. *Psychometrika*, *76*(2), 306–317. https://doi.org/10.1007/s11336-010-9200-6

Brandmaier, A. M., von Oertzen, T., McArdle, J. J., & Lindenberger, U. (2013). Structural Equation Model Trees. *Psychological Methods*, *18*(1), 71–86. https://doi.org/10.1037/a0030001

Byrne, M. L., Whittle, S., Vijayakumar, N., Dennison, M., Simmons, J. G., & Allen, N. B. (2017). A systematic review of adrenarche as a sensitive period in neurobiological development and mental health. *Sensitive Periods across Development*, *25*, 12–28. https://doi.org/10.1016/j.dcn.2016.12.004

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2011). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*, *40*(3), 626–644. https://doi.org/10.1093/ije/dyq190

Cannon, T. D., & Keller, M. C. (2006). Endophenotypes in the genetic analyses of mental disorders. *Annual Review of Clinical Psychology*, *2*(1), 267–290. https://doi.org/10.1146/annurev.clinpsy.2.022305.095232

Chevalier, N., Kurth, S., Doucette, M. R., Wiseheart, M., Deoni, S. C. L., Dean, D. C., III, … LeBourgeois, M. K. (2015). Myelination is associated with processing speed in early childhood: Preliminary insights. *PLOS ONE*, *10*(10), e0139897. https://doi.org/10.1371/journal.pone.0139897

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence*, *41*(4), 244–262. https://doi.org/10.1016/j.intell.2013.04.003

Coyle, T. R., Pillow, D. R., Snyder, A. C., & Kochunov, P. (2011). Processing speed mediates the development of general intelligence (g) in adolescence. *Psychological Science*, *22*(10), 1265–1269. https://doi.org/10.1177/0956797611418243

de Mooij, S. M. M., Henson, R. N. A., Waldorp, L. J., & Kievit, R. A. (2018). Age differentiation within grey matter, white matter and between memory and white matter in an adult lifespan cohort. *The Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.1627-17.2018

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: An update. *Journal of the International Neuropsychological Society*, *10*(2), 301–303. https://doi.org/10.1017/S1355617704102191

Dorn, L. D. (2006). Measuring puberty. *Journal of Adolescent Health*, *39*(5), 625–626. https://doi.org/10.1016/j.jadohealth.2006.05.014

Dunning, D. L., Holmes, J., & Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working memory? A randomized

controlled trial. *Developmental Science*, *16*(6), 915–925. https://doi.org/10.1111/desc.12068

Ferrer, E., Whitaker, K., Steele, J., Green, C., Wendelken, C., & Bunge, S. (2013). White matter maturation supports the development of reasoning ability through its influence on processing speed. *Developmental Science*, *16*(6), 941–951. https://doi.org/10.1111/desc.12088

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. *Behavior Genetics*, *35*(5), 535–549. https://doi.org/10.1007/s10519-005-3281-5

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*(1), 1–34. https://doi.org/10.1016/S0301-0511(00)00051-X

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Gale, C. R., Batty, G. D., Tynelius, P., Deary, I. J., & Rasmussen, F. (2010). Intelligence in early adulthood and subsequent hospitalisation and admission rates for the whole range of mental disorders: Longitudinal study of 1,049,663 men. *Epidemiology (Cambridge, Mass.)*, *21*(1), 70–77. https://doi.org/10.1097/EDE.0b013e3181c17da8

Gallagher, A., & Frederickson, N. (1995). The Phonological Assessment Battery (PhAB): An initial assessment of its theoretical and practical utility. *Educational and Child Psychology*.

Garrett, H. E. (1946). A developmental theory of intelligence. *American Psychologist*, *1*(9), 372–378. https://doi.org/10.1037/h0056380

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., … Gur, R. E. (2001). Computerized Neurocognitive Scanning: I. methodology and validation in healthy people. *Neuropsychopharmacology*, *25*, 766.

Holmes, J., Bryant, A., & Gathercole, S. E. (2018). A transdiagnostic study of children with problems of attention, learning and memory (CALM). *BioRxiv*. https://doi.org/10.1101/303826

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., … Mori, S. (2008). Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, *39*(1), 336–347. https://doi.org/10.1016/j.neuroimage.2007.07.053

Johnson, M. H. (2000). Functional brain development in infants: Elements of an interactive specialization framework. *Child Development*, *71*(1), 75–81.

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, *2*, 475.

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–254. https://doi.org/10.1016/j.neuroimage.2012.06.081

Jöreskog, K. G. (1999). *How large can a standardized coefficient be?* Retrieved from http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*(351), 631–639. https://doi.org/10.2307/2285946

Kail, R. V., Lervåg, A., & Hulme, C. (2015). Longitudinal evidence linking processing speed to the development of reasoning. *Developmental Science*, *19*(6), 1067–1074. https://doi.org/10.1111/desc.12352

Kail, R. V., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*(2), 199–225. https://doi.org/10.1016/0001-6918(94)90003-5

Kamali, A., Flanders, A. E., Brody, J., Hunter, J. V., & Hasan, K. M. (2014). Tracing superior longitudinal fasciculus connectivity in the human brain using high resolution diffusion tensor tractography. *Brain Structure & Function*, *219*(1), 10.1007/s00429-012-0498-y. https://doi.org/10.1007/s00429-012-0498-y

Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 61/2 and 161/2 years. *Journal of Consulting and Clinical Psychology*, *43*(2), 135–147. https://doi.org/10.1037/h0076502

Kievit, R. A., Davis, S. W., Griffiths, J., Correia, M. M., CamCAN, & Henson, R. N. (2016). A watershed model of individual differences in fluid intelligence. *Neuropsychologia*, *91*, 186–198. https://doi.org/10.1016/j.neuropsychologia.2016.08.008

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*, 186–198. https://doi.org/10.1016/j.neuropsychologia.2016.08.008

Kievit, R. A., van Rooijen, H., Wicherts, J. M., Waldorp, L. J., Kan, K. J., Scholte, H. S., & Borsboom, D. (2012). Intelligence and the brain: A model-based approach. *Cognitive Neuroscience*, *3*(2), 89–97. https://doi.org/10.1080/17588928.2011.628383

Knoll, L. J., Fuhrmann, D., Sakhardande, A. L., Stamp, F., Speekenbrink, M., & Blakemore, S. J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, *27*(12), 1620–1631. https://doi.org/10.1177/0956797616671327

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1

Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x

Lövdén, M., Laukka, E. J., Rieckmann, A., Kalpouzos, G., Li, T., Jonsson, T., … Bäckman, L. (2013). The dimensionality of between-person differences in white matter microstructure in old age. *Human Brain Mapping*, *34*(6), 1386–1398.

Lu, P. H., Lee, G. J., Raven, E. P., Tingus, K., Khoo, T., Thompson, P. M., & Bartzokis, G. (2011). Age-related slowing in cognitive processing speed is associated with myelin integrity in a very healthy elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *33*(10), 1059–1068. https://doi.org/10.1080/13803395.2011.595397

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science*, *14*(3), 582–590. https://doi.org/10.1111/j.1467-7687.2010.01005.x

MacPherson, S. E., Cox, S. R., Dickie, D. A., Karama, S., Starr, J. M., Evans, A. C., … Deary, I. J. (2017). Processing speed and the relationship between Trail Making Test-B performance, cortical thinning and white matter microstructure in older adults. *Cortex*, *95*, 92–103. https://doi.org/10.1016/j.cortex.2017.07.021

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. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(8), 1065–1081. https://doi.org/10.1017/S0021963001007909

Mills, K. L., Goddings, A.-L., Herting, M. M., Meuwese, R., Blakemore, S.-J., Crone, E. A., … Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044

Molenaar, D., Kő, N., Rózsa, S., & Mészáros, A. (2017). Differentiation of cognitive abilities in the WAIS-IV at the item level. *Intelligence*, *65*, 48–59. https://doi.org/10.1016/j.intell.2017.10.004

Noack, H., Lövdén, M., & Schmiedek, F. (2014). On the validity and generality of transfer effects in cognitive training research. *Psychological Research*, *78*(6), 773–789. https://doi.org/10.1007/s00426-014-0564-6

Nooner, K., Colcombe, S., Tobe, R., Mennes, M., Benedict, M., Moreno, A., … Milham, M. (2012). The NKI-Rockland Sample: A model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*, *6*, 152. https://doi.org/10.3389/fnins.2012.00152

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence - Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61–65. https://doi.org/10.1037/0033-2909.131.1.61

Parlatini, V., Radua, J., Dell'Acqua, F., Leslie, A., Simmons, A., Murphy, D. G., … Thiebaut de Schotten, M. (2017). Functional segregation and integration within fronto-parietal networks. *NeuroImage*, *146*, 367–375. https://doi.org/10.1016/j.neuroimage.2016.08.031

Penke, L., Maniega, S. M., Murray, C., Gow, A. J., Valdés Hernández, M. C., Clayden, J. D., … Deary, I. J. (2010). A general factor of brain white matter integrity predicts information processing speed in healthy older people. *The Journal of Neuroscience*, *30*(22), 7569. https://doi.org/10.1523/JNEUROSCI.1553-10.2010

Pfefferbaum, A., Sullivan, E. V., Hedehus, M., Lim, K. O., Adalsteinsson, E., & Moseley, M. (2000). Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. *Magnetic Resonance in Medicine*, *44*(2), 259–268.

Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences*, *20*(5), 446–451. https://doi.org/10.1016/j.lindif.2010.05.001

R core team, the. (2015). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org/

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science*, 0956797618774253. https://doi.org/10.1177/0956797618774253

Rosseel, Y. (2012). lavaan: An R package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 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. https://doi.org/10.1016/j.intell.2015.09.002

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of Structural Equation Models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*(2), 23–74.

Schubert, A.-L., Hagemann, D., & Frischkorn, G. T. (2017). Is general intelligence little more than the speed of higher-order processing? *Journal of Experimental Psychology: General*, *146*(10), 1498–1512. https://doi.org/10.1037/xge0000325

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*(2), 138–166.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*(3), 535–551. https://doi.org/10.1016/j.paid.2007.09.015

Sisk, C. L., & Zehr, J. L. (2005). Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinologgy*, *26*(3–4), 163–174. https://doi.org/10.1016/j.yfrne.2005.10.003

Tamnes, C. K., Roalf, D. R., Goddings, A.-L., & Lebel, C. (2017). Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Developmental Cognitive Neuroscience*. https://doi.org/10.1016/j.dcn.2017.12.002

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the lifespan. *Developmental Psychology*, *45*(4), 1097–1118. https://doi.org/10.1037/a0015864

Vestergaard, M., Madsen, K. S., Baaré, W. F., Skimminge, A., Ejersbo, L. R., Ramsøy, T. Z., … Jernigan, T. L. (2011). White matter microstructure in superior longitudinal fasciculus associated with spatial working memory performance in children. *Journal of Cognitive Neuroscience*, *23*(9), 2135–2146.

Volkow, N. D., Koob, G. F., Croyle, R. T., Bianchi, D. W., Gordon, J. A., Koroshetz, W. J., … Weiss, S. R. B. (2018). The conception of the ABCD study: From substance use to a broad NIH collaboration. *The Adolescent Brain Cognitive Development (ABCD) Consortium: Rationale, Aims, and Assessment Strategy*, *32*, 4–7. https://doi.org/10.1016/j.dcn.2017.10.002

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192–196. https://doi.org/10.3758/BF03206482

Wandell, B. A. (2016). Clarifying human white matter. *Annual Review of Neuroscience*, *39*(1), 103–128. https://doi.org/10.1146/annurev-neuro-070815-013815

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence -Second Edition*. San Antonio, TX: Pearson.