OXFORD

Cerebral Cortex, 2023, **33**, 6723–6741
https://doi.org/10.1093/cercor/bhac538
Advance access publication date 20 January 2023
Original Article

Robust associations between white matter microstructure and general intelligence

Christina Stammen[1], Christoph Fraenz[1], Rachael G. Grazioplene[2], Caroline Schlüter[3], Viola Merhof[4], Wendy Johnson[5], Onur Güntürkün[3], Colin G. DeYoung[6], Erhan Genç[1,*]

[1]Department of Psychology and Neuroscience, Leibniz Research Centre for Working Environment and Human Factors (IfADo), 44139 Dortmund, Germany,
[2]Department of Psychology, Yale University, New Haven, CT 06510, United States,
[3]Department of Biopsychology, Institute of Cognitive Neuroscience, Ruhr University Bochum, 44801 Bochum, Germany,
[4]Chair of Research Methods and Psychological Assessment, University of Mannheim, 68161 Mannheim, Germany,
[5]Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom,
[6]Department of Psychology, University of Minnesota, Minneapolis, MN 55455, United States

*Corresponding author: Erhan Genç, Department of Psychology and Neuroscience, Leibniz Research Centre for Working Environment and Human Factors (IfADo), Ardeystraße 67, Dortmund 44139, Germany. Email: genc@ifado.de

Few tract-based spatial statistics (TBSS) studies have investigated the relations between intelligence and white matter microstructure in healthy (young) adults, and those have yielded mixed observations, yet white matter is fundamental for efficient and accurate information transfer throughout the human brain. We used a multicenter approach to identify white matter regions that show replicable structure–function associations, employing data from 4 independent samples comprising over 2000 healthy participants. TBSS indicated 188 voxels exhibited significant positive associations between *g* factor scores and fractional anisotropy (FA) in all 4 data sets. Replicable voxels formed 3 clusters, located around the left-hemispheric forceps minor, superior longitudinal fasciculus, and cingulum–cingulate gyrus with extensions into their surrounding areas (anterior thalamic radiation, inferior fronto-occipital fasciculus). Our results suggested that individual differences in general intelligence are robustly associated with white matter FA in specific fiber bundles distributed across the brain, consistent with the Parieto-Frontal Integration Theory of intelligence. Three possible reasons higher FA values might create links with higher *g* are faster information processing due to greater myelination, more direct information processing due to parallel, homogenous fiber orientation distributions, or more parallel information processing due to greater axon density.

Key words: DWI; general intelligence; multicenter study; TBSS; white matter.

Introduction

People differ in general intelligence, i.e. "[. . .] their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" (Neisser et al. 1996, p. 77). As discovered by Spearman (1904), individuals who do well in one cognitive task tend to perform above average in other cognitive tasks as well. The phenomenon of positively correlated cognitive test scores, which he termed the "positive manifold," led Spearman to declare the existence of "*g*," the general factor of intelligence. Though *g* is actually just a statistical observation, it is an important one because it is relevant to many aspects of everyday life. For example, intelligence is positively correlated with school performance (Neisser et al. 1996; Roth et al. 2015), job performance (Gottfredson 1997; Schmidt and Hunter 2004), socioeconomic success (Strenze 2007), income (Zagorsky 2007), and even physical health, longevity, and ephemerals such as stability of marital relationships (Whalley and Deary 2001; Hemmingsson et al. 2006; Batty et al. 2007; Deary et al. 2010b; Calvin et al. 2011; Calvin et al. 2017; Aspara et al. 2018). Due to the impacts that intelligence or *g* seems to have on life outcomes, it has always been of interest to identify specific structures within the human brain that are associated with its interindividual differences.

While it is one well-replicated observation that bigger brains are weakly to moderately associated with higher intelligence (McDaniel 2005; Pietschnig et al. 2015; Cox et al. 2019), the advent of in vivo neuroimaging techniques has allowed scientists to move from overall brain size to various properties of single brain regions and beyond. Jung and Haier (2007) reviewed 37 neuroimaging studies that aimed to identify intelligence-related brain regions using various intelligence measures and imaging techniques. Based on the commonalities across findings, they proposed the Parieto-Frontal Integration Theory (P-FIT) of intelligence. P-FIT nominates a set of distributed brain regions, mainly located in parietal and frontal areas, whose functional and structural properties are related to interindividual intelligence differences. Each P-FIT area is believed to be involved in the multiple information processing stages used in solving abstract reasoning tasks. Hence, more efficient and flawless information transfer between these regions seems fundamental to intellectual achievements, which in turn indicates roles of brain white matter (Jung and Haier 2007). The brain's white matter mainly consists of myelinated axons that are organized in fiber tracts running from one brain region to another (Filley 2012), which enables thereby the information transfer. The hypothesis that the integrity of certain white matter fiber tracts is crucial for intelligence (Jung and Haier 2007) has been empirically supported by

Received: April 29, 2022. **Revised:** December 21, 2022. **Accepted:** December 22, 2022

Gläscher et al. (2010) who used voxel-based lesion-symptom mapping in a large sample of patients with focal brain damage. Their observations indicated that severe damage to fiber tracts linking P-FIT areas (superior longitudinal fasciculus, arcuate fasciculus, uncinate fasciculus, and inferior fronto-occipital fasciculus) was significantly associated with lower intelligence (Gläscher et al. 2010). Subsequent studies using lesion-symptom mapping were consistent with these observations (Barbey et al. 2012; Barbey et al. 2014; Bowren et al. 2020). Even newer theories based on graph theory, such as Barbey's (2018) Network Neuroscience Theory, which proposes that general intelligence reflects individual differences in whole brain topology's efficiency and in the capacity to dynamically reconfigure brain network states, emphasize the importance of the brain's structural (and functional) organization since it may facilitate or constrain network flexibility. The idea that intelligence relies on a dynamic system comprising interacting subcomponents distributed all over the brain does not contradict previous research reporting that some brain regions or white matter fiber tracts seem to be more commonly implicated in successfully accomplishing cognitive tasks than others (Jung and Haier 2007; Cox et al. 2019). It only shifts the focus so that previously reported, focal differences in brain structure are no longer seen as isolated causes of differences in intelligence, but as traces of employed functional dynamics and architecture enabling easier transition between functional network states.

The advent of diffusion-weighted imaging (DWI) led to an exponential growth of white matter brain imaging studies (Deary et al. 2022). DWI is based on diffusion of water molecules (Le Bihan and Breton 1985; Le Bihan et al. 1986; Le Bihan 2014) and indicates anisotropic, directional diffusion patterns within voxels containing coherently oriented white matter fibers and isotropic, non-directional patterns within voxels containing randomly oriented fibers or fluid-filled spaces such as ventricles (Le Bihan 2003). The most widely used metric to quantify water diffusion's degrees of directionality in a summative manner is fractional anisotropy (FA). Here, higher FA values indicate more parallel diffusion trajectories (Basser and Pierpaoli 1996; Assaf and Pasternak 2008). Although FA is clearly related to white matter microstructure, it may be misleading to use it as a marker of microstructural integrity (Jones et al. 2013). FA is a complex and unspecific measure affected by various physiological factors like axon diameter, fiber density, myelin concentration, or distribution of fiber orientation (Beaulieu 2002; Le Bihan 2003; Jones et al. 2013; Friedrich et al. 2020). These factors make it challenging to disentangle and interpret the actual sources of signal differences (Jones et al. 2013). Nevertheless, FA is a widely used metric and its association with intelligence has been investigated extensively. Studies have analyzed white matter properties by averaging across specific regions of interest (Deary et al. 2006; Tang et al. 2010; Power et al. 2019), extracting them from whole fiber tracts (Yu et al. 2008; Kontis et al. 2009; Penke et al. 2010; Clayden et al. 2012; Penke et al. 2012; Booth et al. 2013; Ferrer et al. 2013; Kievit et al. 2014; Ohtani et al. 2014; Muetzel et al. 2015; Nestor et al. 2015; Urger et al. 2015; Cremers et al. 2016; Kievit et al. 2016; Kievit et al. 2018; Bathelt et al. 2019; Cox et al. 2019; Dubner et al. 2019; Fuhrmann et al. 2020; Góngora et al. 2020; Holleran et al. 2020; Simpson-Kent et al. 2020; Suprano et al. 2020; Kennedy et al. 2021), or by a whole-brain voxel-based approach (Schmithorst et al. 2005; Chiang et al. 2009; Schmithorst 2009; Allin et al. 2011; Navas–Sanchez et al. 2014). As summarized by Genç and Fraenz et al. (2021), the majority of such studies reported positive relations between intelligence and average FA values from many major white matter pathways, mostly representing connections between P-FIT areas. Independent of the specific methods used, similar patterns emerged among different studies. The 4 fiber tracts most commonly associated with intelligence differences are the genu and the splenium of the corpus callosum, the uncinate fasciculus, and the superior longitudinal fasciculus (Genç and Fraenz 2021).

Studies investigating pre-selected brain regions or white matter tracts are prone to miss relevant relations in non-selected areas. Analyses adapting voxel-based methods, such as voxel-based morphometry (Ashburner and Friston 2000), to analyze FA images also have various shortcomings such as alignment inaccuracies (Smith et al. 2006). Tract-based spatial statistics (TBSS) has been introduced as an approach that combines the strengths of tractography-based and voxel-based analyses to overcome the aforementioned limitations (Smith et al. 2006). Although TBSS has advantages, few studies have investigated the relation between FA and intelligence in healthy (young) adults using this method. Dunst et al. (2014) found no significant associations between general intelligence and FA in any white matter voxel, whereas Malpas et al. (2016) reported significant positive relations in 32% of voxels constituting the white matter skeleton (right anterior thalamic radiation, left superior longitudinal fasciculus, left inferior fronto-occipital fasciculus, and left uncinate fasciculus). In line with Dunst et al. (2014), Hidese et al. (2020) found no significant associations between general intelligence and regional white matter FA, despite analyzing a larger sample. Tamnes et al. (2010) employed a sample comprised of 168 participants, aged between 8 and 30 years. They focused their TBSS analyses on verbal and nonverbal reasoning abilities. While 4.6% of voxels in the white matter skeleton showed significant positive associations between FA and verbal abilities (left anterior thalamic radiation, left cingulum–cingulate gyrus, left and right superior longitudinal fasciculus), 1.6% of skeleton voxels (left superior longitudinal fasciculus, forceps major) showed significant positive associations between FA and nonverbal reasoning abilities (Tamnes et al. 2010).

Previous TBSS studies have often had samples small enough that effect size estimates are likely to be highly variable and inaccurate. Furthermore, inconsistencies such as different sample sizes or intelligence measures limited their comparability. In short, they do not allow clear conclusions to be drawn about associations between general intelligence and FA. Some found significant positive relations, whereas others did not. As proposed by Genç and Fraenz et al. (2021), such inconsistent observations may be tackled by following a multicenter approach. To this end, multiple, independent data sets, typically collected by different research groups, are analyzed in the same way. Importantly, only those results that replicate across the majority (or all) of samples are considered robust. We followed this approach as methodologically consistently as possible, searching for replicable observations among 4 independent data sets comprising cross-sectional data from more than 2000 healthy participants. Our group performed whole-brain TBSS analyses to examine the associations between general intelligence, in the form of g factor scores, and FA separately on each sample. Besides the aforementioned advantage of multicenter studies, another reason for choosing this rather conservative approach was that a first-level combination (pooling all) of our 4 data sets with not-identical behavioral measures was not possible since sample mean g levels might differ and because imaging data were obtained on different scanners. However, as g and FA values were available for all samples and relative values between subjects within samples should be comparable, we were able to combine the data sets at a second level with our multicenter approach. Data were collected at Ruhr-University Bochum (RUB) in Germany with $n = 557$

Table 1. Sample characteristics.

Data set	Male/Female	Age range	Age	Handedness
RUB	283/274	18–75	27.3 ± 9.4	73.1 ± 50.7
HCP	490/571	22–37	28.7 ± 3.7	65.9 ± 44.6
UMN	129/122	20–40	26.2 ± 4.9	100.0 ± 0.0
NKI	137/259	6–85	44.4 ± 22.9	65.4 ± 47.1

Age and handedness are depicted as mean ± SD.

(Genç et al. 2021), the Human Connectome Project (HCP) with $n = 1061$ (van Essen et al. 2013), the University of Minnesota (UMN) with $n = 251$ (Grazioplene et al. 2015, 2016), and the Nathan Kline Institute (NKI) with $n = 396$ (Nooner et al. 2012). We compared observations to identify white matter areas that exhibited replicable structure–function associations among data sets. As the overlap among multiple data sets' results will be likely to include fewer areas than a single data set's results, our study might yield relatively circumscribed but robust associations between white matter and *g*. This could give the impression that only focal differences in FA are associated with differences in general intelligence. However, if some white matter fiber tracts seem more commonly implicated in successfully accomplishing cognitive tasks this will not mean other brain white matter areas are irrelevant. Involvement of white matter throughout the brain for information transfer seems relevant for intellectual performance as intelligence is more likely to emerge from a dynamic system comprising interacting subcomponents (Barbey 2018).

Materials and methods

Participants

Data set RUB

The RUB sample encompassed 557 participants (Table 1), mainly university students of different majors, who were either paid for their participation or received course credits. Although the age range was between 18 and 75 years, the data set was predominantly comprised of individuals from young adulthood. Individuals were not admitted to the study if they had insufficient German language skills or reported having undergone any of the employed intelligence tests within the last 5 years. They were also excluded if they or any of their close relatives suffered from neurological and/or mental illnesses, as assessed by a self-report questionnaire. The study protocol was approved by the Local Ethics Committee of the Faculty of Psychology at Ruhr University Bochum (vote Nr. 165). All participants gave written informed consent and were treated according to the Declaration of Helsinki.

Data set HCP

The HCP sample data were provided by the HCP, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657), funded by the 16 US National Institutes of Health (NIH) Institutes, Centers supporting the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. We employed the "1200 Subjects Data Release" (van Essen et al. 2013). It includes behavioral and imaging data from 1206 young adults. To compute a *g* factor, all participants with missing values in one or more of the intelligence measurements listed below had to be excluded, which reduced the sample to $n = 1188$ (mean age: 28.8 years, standard deviation [SD] = 3.7 years, 641 females). Since DWI data were not available for all participants, the final sample for the TBSS analysis was limited to 1061 participants (Table 1). To be included in the data set, participants had to have no significant history of psychiatric disorder, substance abuse, neurological, or cardiological disease and give valid informed consent (van Essen et al. 2012).

Data set UMN

The UMN data set encompassed 335 participants (mean age: 26.3 years, SD = 5.0 years, 164 females) with sufficient data from intelligence testing to compute a general factor *g*. Since DWI data were not available for all participants, the final sample for the TBSS analysis was reduced to 251 participants (Table 1). Individuals who reported a history of neurologic or severe psychiatric disorders, current drug or alcohol problems, or current use of psychotropic medication (antipsychotics, anticonvulsants, and stimulants) were not admitted to the study. The study protocol was approved by the UMN Institutional Review Board and all participants gave written informed consent.

Data set NKI

Data collection for the NKI sample is still ongoing. It is intended to investigate the neurologies of psychiatric disorders (Nooner et al. 2012). The "Enhanced Nathan Kline Institute—Rockland Sample" data set (Nooner et al. 2012) is part of the 1000 Functional Connectomes Project (http://fcon_1000.projects.nitrc.org) and we downloaded it from its official website (http://fcon_1000.projects.nitrc.org/indi/enhanced/). Since our study is focused on healthy participants, we included only individuals who did not report any history of psychiatric illness. Moreover, they also had to have complete intelligence test data. We used these to calculate the *g* factor ($n = 417$, mean age: 43.5 years, SD = 23.5 years, 273 females). For the final sample, usable for TBSS analysis, we had to exclude additional participants due to lack of DWI data ($n = 396$, Table 1). Relative to the other data sets, which mostly consisted of young adults, the NKI sample had a much greater age range and higher mean age (Table 1). However, since exclusion of all participants outside the 20–40 range would have cost 306 participants, we included all participants with suitable data. The study protocol was approved by the Institutional Review Boards at the NKI and Montclair State University. Written informed consent for the study was obtained from all participants or, for children, additionally from a legal guardian (Nooner et al. 2012).

General intelligence factor, g, computation

Research on the psychometric structure of intelligence has modified and extended Spearman's original ideas regarding the existence of *g*. In recent hierarchically organized models, *g* is placed at the apex of a hierarchy with broad cognitive domains at a lower level and narrow cognitive abilities at the basis (Schneider and McGrew 2012; Flanagan and Dixon 2013). There is considerable evidence for the existence of such structures, but their specifics depend on the tests and sample properties. Nonetheless, when ranges of tests included are broad, their *g* factors correlate for all practical purposes completely, e.g. Johnson et al. (2004);

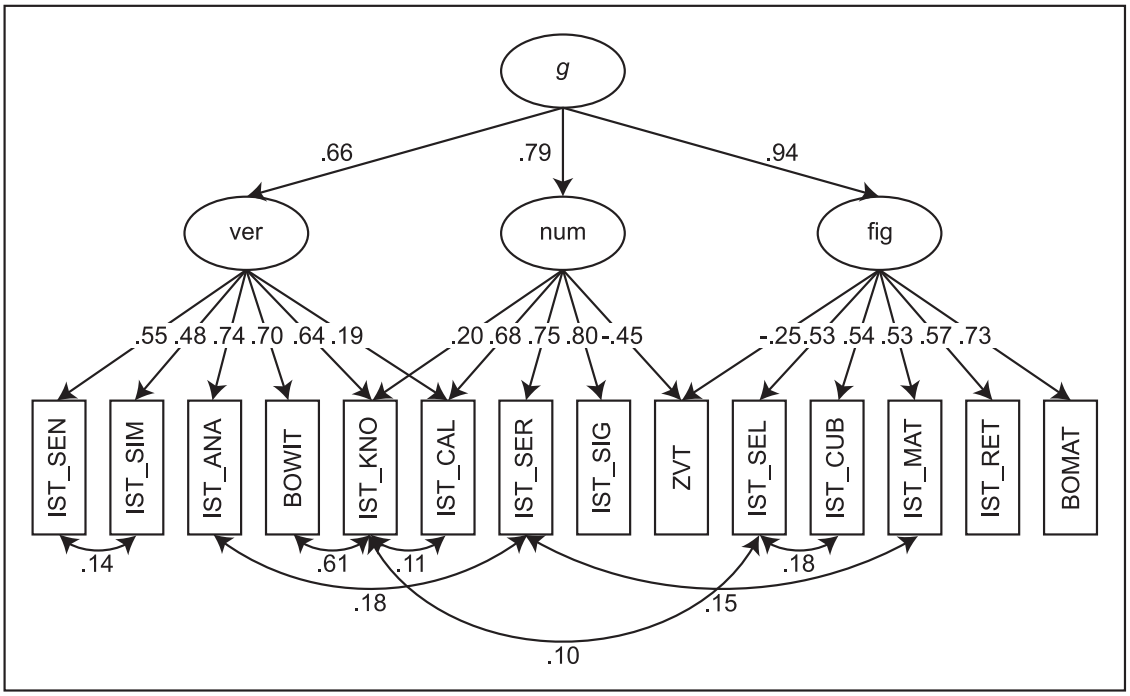

Fig. 1. Confirmatory factor analytic model of the RUB data set. G = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, num = numerical intelligence as broad cognitive domain, fig = figural intelligence as broad cognitive domain, IST_SEN = subtest Sentence Completion of the I-S-T 2000 R, IST_SIM = subtest Similarities of the I-S-T 2000 R, IST_ANA = subtest Analogies of the I-S-T 2000 R, BOWIT = Bochumer Wissenstest, IST_KNO = parameter Knowledge of the I-S-T 2000 R, IST_CAL = subtest Calculations of the I-S-T 2000 R, IST_SER = subtest Number Series of the I-S-T 2000 R, IST_SIG = subtest Numerical Signs of the I-S-T 2000 R, ZVT = Zahlenverbindungstest, IST_SEL = subtest Figure Selection of the I-S-T 2000 R, IST_CUB = subtest Cubes of the I-S-T 2000 R, IST_MAT = subtest Matrices of the I-S-T 2000 R, IST_RET = parameter Retentiveness of the I-S-T 2000 R, BOMAT = Bochumer Matrizentest.

Johnson et al. (2008). Thus, content of *g* is relatively unaffected by the tests from which it was generated, though the level of any one person's factor score certainly could be. Measurement invariance does not hold across *g* ranges. For example, arithmetic tests tend to be processing speed tasks for people with high *g* levels but reasoning tasks for people with low *g* levels. Furthermore, a person with average performance on various intelligence tests may have a standardized *g*-value that is below average in a highly intelligent sample and a *g*-value that is above average in a less intelligent sample. Since sample mean *g* levels might differ and because imaging data were obtained on different scanners (which also affects what is observed), it was not possible to combine the 4 data sets employed in our study.

We used the intelligence test scores of each data set (see section "Description of intelligence tests") to compute *g* factor scores for every participant. To do this, we regressed age, sex, age*sex, age^2, and age^2*sex from the test scores. We added age^2 because we wanted to be sure there were no quadratic relations with age (McGue and Bouchard 1984). We then developed a hierarchical factor model separately for each data set based on the standardized residuals by first using exploratory factor analysis to develop the optimal factor model (results not shown) and then performing confirmatory factor analysis. We assessed model fit using the chi-square (X^2) statistic as well as the fit indices Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), and Tucker–Lewis index (TLI). The chi-square (X^2) statistic tests whether the difference between the model-implied variance–covariance matrix and the empirically observed variance–covariance matrix is zero (Hu and Bentler 1999). Nonsignificance therefore indicates good model fit (Bentler and Bonett 1980), but is essentially never attained in samples of any size, which is why it is important to consider other indices of model fit. Values of RMSEA and SRMR less than 0.05 and values of CFI and TLI greater than 0.97 are considered good (Hu and Bentler 1999). We used these models to calculate regression-based *g*-factor scores for each participant, winsorizing outliers, which is the most robust way to address the potential problems that can create (Wilcox 1997). We examined *g* factor score distributions separately for each sample and limited data points far enough above or below where the data begin to cohere to distort regression lines to those levels. To ensure that we did not alter overall distribution shape unduly, we examined both skew and kurtosis.

Confirmatory factor models.

Figures 1–4 show the postulated confirmatory factor models for the data sets, the z-standardized factor loadings, and the covariances between individual subtests. The chi-square (X^2) statistics and the fit indices to evaluate model fit are listed in Table 2. The confirmatory factor analyses of all data sets yielded quite good (RUB and HCP) to excellent (UMN and NKI) fit. That the chi-square (X^2) statistics of the 2 largest data sets RUB and HCP were significant, does not itself indicate poor model fit because the chi-square (X^2) statistic is a direct function of sample size, which means that the probability of rejecting any model is greater with greater sample size (Jöreskog 1969; Bentler and Bonett 1980). As the RUB data set contained tests intended to tap "general knowledge" (IST_KNO and BOWIT) that are not commonly part of cognitive test batteries, we also calculated an alternative *g* factor without these 2 tests (factor model not shown). It was not possible anymore to get a hierarchical factor model with good model fit. Therefore, the new factor was a nonhierarchical

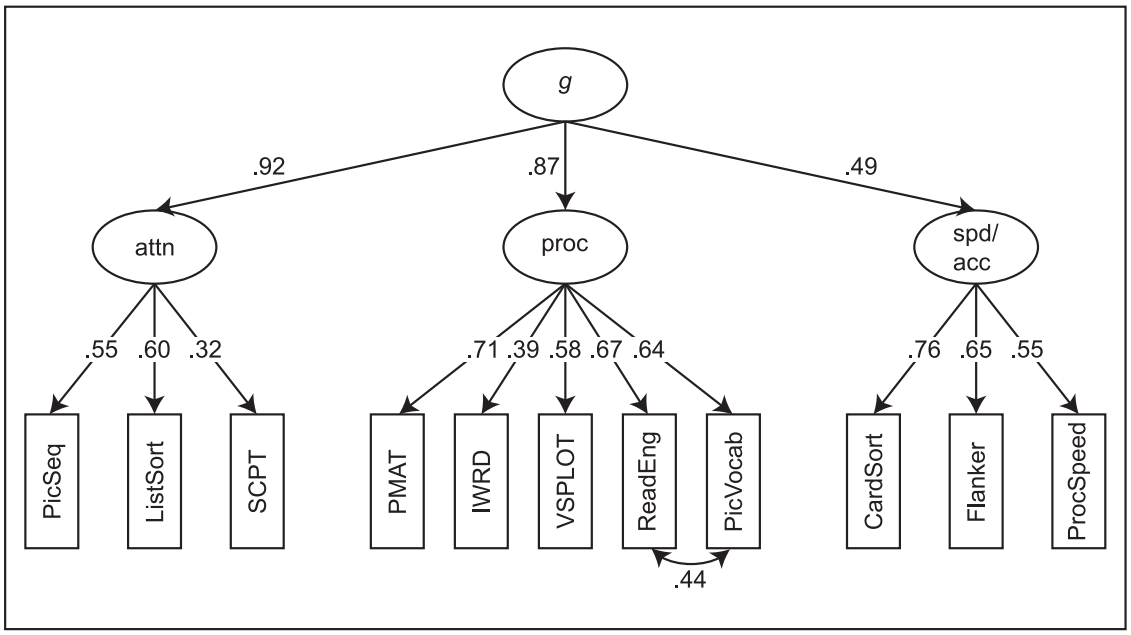

Fig. 2. Confirmatory factor analytic model of the HCP data set. *G* = general factor of intelligence, attn = attention as broad cognitive domain, proc = processing as broad cognitive domain, spd/acc = speed/accuracy as broad cognitive domain, PicSeq = subtest Picture Sequence Memory Test of the NIH toolbox, ListSort = subtest List Sorting Working Memory Test of the NIH toolbox, SCPT = subtest Short Penn Continuous Performance Test of the Penn CNB, PMAT = subtest Penn Matrix Reasoning Task of the Penn CNB, IWRD = subtest Penn Word Memory Test of the Penn CNB, VSPLOT = subtest Variable Short Penn Line Orientation Test of the Penn CNB, ReadEng = subtest Oral Reading Recognition Test of the NIH toolbox, PicVocab = subtest Picture Vocabulary Test of the NIH toolbox, CardSort = subtest Dimensional Change Card Sort Test of the NIH toolbox, Flanker = subtest Flanker Inhibitory Control and Attention Test of the NIH toolbox, ProcSpeed = subtest Pattern Comparison Processing Speed Test of the NIH toolbox.

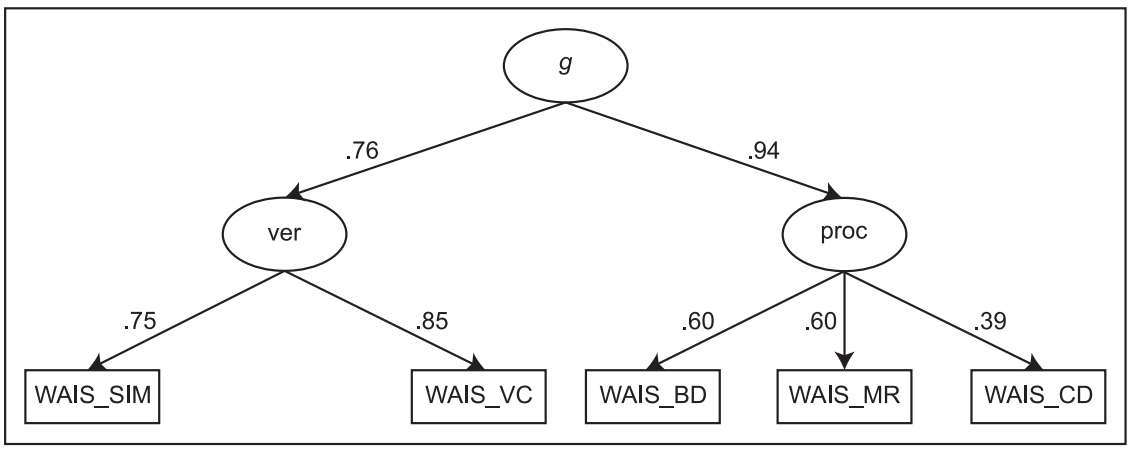

Fig. 3. Confirmatory factor analytic model of the UMN data set. *G* = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, proc = processing as broad cognitive domain, WAIS_SIM = subtest Similarities of the WAIS-IV, WAIS_VC = subtest Vocabulary of the WAIS-IV, WAIS_BD = subtest Block Design of the WAIS-IV, WAIS_MR = subtest Matrix Reasoning of the WAIS-IV, WAIS_CD = subtest Coding of the WAIS-IV.

single-factor solution. It correlated at $r = 0.976$ with the one shown in Fig. 1. Since there was no substantial difference, we decided to use the hierarchical factor model (Fig. 1) and include all available intelligence measures.

Description of intelligence tests

Data set RUB

I-S-T 2000 R

The Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R; Liepmann et al. 2007) is a broadly applicable, well-established German intelligence test battery that takes about 2.5 hours to complete. It measures multiple intelligence facets as well as general intelligence (Table 3). Most included cognitive tasks are presented in multiple-choice format. Reliability estimates (Cronbach's α) are between 0.88 and 0.96 for subtests and composite scores (Liepmann et al. 2007).

BOMAT-Advanced Short

The Bochumer Matrizentest (BOMAT; Hossiep et al. 2001) is a nonverbal German intelligence test (Table 3) whose structure is comparable to the well-established Raven's Advanced Progressive Matrices (Raven et al. 1990). For the study at hand, we used the advanced short version, which is widely used in neuroscientific research and known to have high discriminatory power in samples with generally high intellectual abilities, thus avoiding possible ceiling effects (Hossiep et al. 2001; Jaeggi et al. 2008; Oelhafen et al. 2013; Genç et al. 2018; Genç et al. 2019; Fraenz et al. 2021). Split-half reliability of the BOMAT is 0.89 and Cronbach's α is 0.92 (Hossiep et al. 2001).

BOWIT

The Bochumer Wissenstest (BOWIT; Hossiep and Schulte 2008) is a German "general knowledge" questionnaire. It is available in 2

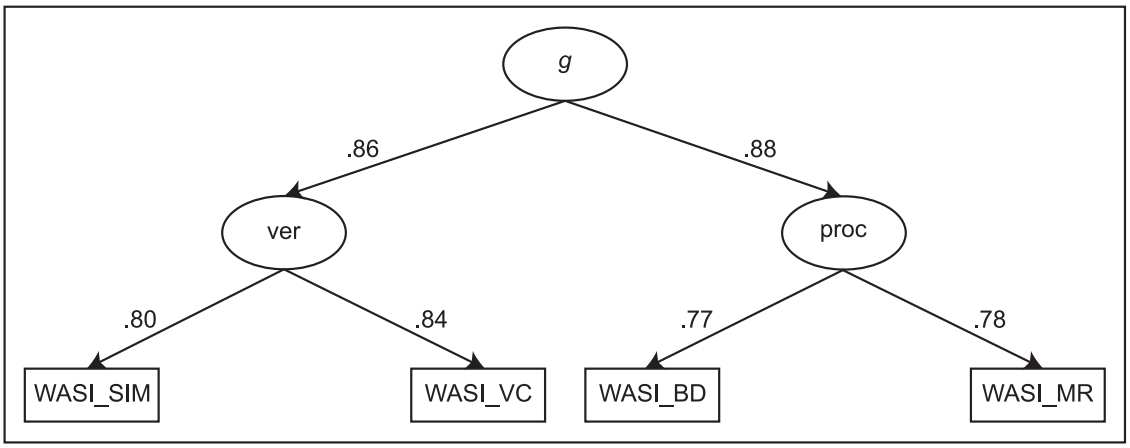

Fig. 4. Confirmatory factor analytic model of the NKI data set. G = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, proc = processing as broad cognitive domain, WASI_SIM = subtest Similarities of the WASI-II, WASI_VC = subtest Vocabulary of the WASI-II, WASI_BD = subtest Block Design of the WASI-II, WASI_MR = subtest Matrix Reasoning of the WASI-II.

Table 2. Fit indices of the confirmatory factor analyses.

Model	X^2	*df*	RMSEA	SRMR	CFI	TLI	R^2
Data set RUB	127.97***	64	.042	.033	.979	.969	.39
Data set HCP	118.35***	40	.041	.028	.973	.963	.32
Data set UMN	2.51	4	.000	.013	1.000	1.012	.47
Data set NKI	0.13	1	.000	.002	1.000	1.009	.65

RMSEA = Root Mean Square Error of Approximation. SRMR = Standardized Root Mean Square Residual. CFI = Comparative Fit Index. TLI = Tucker-Lewis-Index. R^2 = amount of variance of the data sets' subtests explained by *g* (calculated via an one-factor-model). *** $P < 0.001$.

Table 3. Cognitive tests used to estimate *g* in the RUB sample.

Intelligence test	Task description	No. of items	Construct measured
I-S-T 2000 R			
1. IST_SEN	Complete sentences	20	Verbal intelligence
2. IST_ANA	Find analogies	20	
3. IST_SIM	Recognize similarities	20	
4. IST_CAL	Solve arithmetic calculations	20	Numerical intelligence
5. IST_SER	Complete number series	20	
6. IST_SIG	Add arithmetic signs to mathematical equations	20	
7. IST_SEL	Select and reassemble parts of a cut-up figure	20	Figural intelligence
8. IST_CUB	Mentally rotate and match 3-dimensional objects	20	
9. IST_MAT	Solve matrix-reasoning problems	20	
10. IST_RET	Memorize series of words or figure pairs	23	Retention
11. IST_KNO	Multiple-choice questions on 6 knowledge facets: art/literature, economy, geography/history, mathematics, science, and daily life	84	General knowledge
BOMAT	Solve matrix-reasoning problems (5 × 3 matrices)	29	Non-verbal reasoning
BOWIT	Multiple-choice questions on 11 different knowledge facets: biology/chemistry, mathematics/physics, nutrition/exercise/health, technology/electronics, arts/architecture, civics/politics, economics/laws, geography/logistics, history/archeology, language/literature, and philosophy/religion	308	General knowledge
ZVT	Connect numbers from 1 to 90 based on a specific rule as fast as possible	4	Processing speed

I-S-T 2000 R = Intelligenz-Struktur-Test 2000 R. IST_SEN = subtest Sentence Completion of the I-S-T 2000 R. IST_ANA = subtest Analogies of the I-S-T 2000 R. IST_SIM = subtest Similarities of the I-S-T 2000 R. IST_CAL = subtest Calculations of the I-S-T 2000 R. IST_SER = subtest Number Series of the I-S-T 2000 R. IST_SIG = subtest Numerical Signs of the I-S-T 2000 R. IST_SEL = subtest Figure Selection of the I-S-T 2000 R. IST_CUB = subtest Cubes of the I-S-T 2000 R. IST_MAT = subtest Matrices of the I-S-T 2000 R. IST_RET = subtest Retentiveness of the I-S-T 2000 R. IST_KNO = subtest Knowledge of the I-S-T 2000 R. BOMAT = Bochumer Matrizentest. BOWIT = Bochumer Wissenstest. ZVT = Zahlenverbindungstest.

parallel test forms, in which each knowledge facet is represented by 14 multiple-choice questions (Table 3). All participants completed both test forms, resulting in 308 items. In the BOWIT's manual split-half reliability is reported as 0.96, Cronbach's α 0.95, test–retest reliability 0.96, and parallel-form reliability 0.91 (Hossiep and Schulte 2008).

ZVT

The Zahlenverbindungstest (ZVT; Oswald and Roth 1987) is a trail-making test to assess cognitive processing speed in both children and adults. The test consists of 2 short sample tasks and 4 assessed tasks (Table 3). The reliability across the 4 tasks is reported as 0.95 in adults. The six-month retest-reliability

Table 4. Cognitive tests used to estimate *g* in the HCP sample.

Intelligence test	Task description	No. of items	Construct measured
Penn CNB			
1. PMAT	Solve matrix-reasoning problems (2 × 2, 3 × 3, or 1 × 5 matrices)	24	Non-verbal reasoning
2. SCPT	Indicate when lines (presented for 300 milliseconds) form a number or a letter	180	Visual attention
3. VSPLOT	Rotate one line on a computer screen so that it is parallel to another line	24	Visual–spatial processing
4. IWRD	Memorize 20 words and recognize them afterwards within 40 words including 20 distractors matched for length, imageability, and concreteness	Form A	Verbal episodic memory
NIH Toolbox			
1. Flanker	Indicate the direction of a central arrow, flanked by arrows pointing in the same or the opposite direction as the target	40	Executive function (attention)
2. CardSort	Assign pictures that vary along 2 dimensions (e.g. shape and color) to 1 of 2 target pictures so that the pictures match either in shape or in color (the criterion is displayed and varies without a predictable pattern)	40	Executive function (cognitive flexibility)
3. ListSort	Repeat stimuli, beforehand presented as a series, in order of size (first condition: all stimuli come from the same category; second condition: stimuli belong to 2 categories and must be repeated in order of size as well as category-specific)	Stop criterion: failure in 2 trials of the same length	Working memory capacity
4. PicSeq	Arrange pictures according to a previously seen spatial arrangement	3	Episodic memory
5. ReadEng	Pronounce letters and words as correctly as possible	30–40 depending on performance	Reading decoding skill
6. PicVocab	Choose out of 4 images the one that matches to a spoken word	25	Vocabulary knowledge
7. ProcSpeed	Identify as many image pairs as possible, displayed side-by-side, as identical or not	130 image pairs (time limit: 90 seconds)	Processing speed

Penn CNB = University of Pennsylvania Computerized Neurocognitive Battery. PMAT = subtest Penn Matrix Reasoning Task of the Penn CNB. SCPT = subtest Short Penn Continuous Performance Test of the Penn CNB. VSPLOT = subtest Variable Short Penn Line Orientation Test of the Penn CNB. IWRD = subtest Penn Word Memory Test of the Penn CNB. NIH Toolbox = NIH Toolbox for the Assessment of Neurological and Behavioral Function. Flanker = subtest Flanker Inhibitory Control and Attention Test of the NIH Toolbox. CardSort = subtest Dimensional Change Card Sort Test of the NIH Toolbox. ListSort = subtest List Sorting Working Memory Test of the NIH Toolbox. PicSeq = subtest Picture Sequence Memory Test of the NIH Toolbox. ReadEng = subtest Oral Reading Recognition Test of the NIH Toolbox. PicVocab = subtest Picture Vocabulary Test of the NIH Toolbox. ProcSpeed = subtest Pattern Comparison Processing Speed Test of the NIH Toolbox.

is reported to be between 0.84 and 0.90 (Oswald and Roth 1987).

Data set HCP

Penn CNB

Four subtests from the University of Pennsylvania Computerized Neurocognitive Battery (Penn CNB; Gur et al. 2001; Gur et al. 2010; Moore et al. 2015) were used to assess intelligence (Table 4). These included the Penn Matrix Reasoning Task (PMAT), the Short Penn Continuous Performance Test (SCPT), the Variable Short Penn Line Orientation Test (VSPLOT), and the Penn Word Memory Test (IWRD). The reliability estimates (Cronbach's α) for all subtests of the Penn CNB are reported to be between 0.55 and 0.98 (Gur et al. 2010). Internal consistency was reported in a Dutch study to have a median Cronbach's α of 0.86 across all Penn CNB subtests (Swagerman et al. 2016).

NIH Toolbox

Seven subtests from the NIH Toolbox for the Assessment of Neurological and Behavioral Function (http://www.nihtoolbox.org; Gershon et al. 2013; Weintraub et al. 2013; Heaton et al. 2014) were selected to assess intelligence (Table 4). These were the Flanker Inhibitory Control and Attention Test (Flanker), the Dimensional Change Card Sort Test (CardSort), the List Sorting Working Memory Test (ListSort), the Picture Sequence Memory Test (PicSeq), the Oral Reading Recognition Test (ReadEng), the Picture Vocabulary Test (PicVocab), and the Pattern Comparison Processing Speed Test (ProcSpeed). The NIH Toolbox has been validated with several American samples (Heaton et al. 2014; Weintraub et al. 2013). For the subtests, Weintraub et al. (2013) reported test–retest reliabilities (intraclass correlation coefficients) between $r = 0.78$ and 0.99. Heaton et al. (2014) built and analyzed composite scores and found acceptable internal consistency (Cronbach's α between 0.77 and 0.84) as well as excellent test–retest reliabilities between $r = 0.86$ and 0.92.

Data set UMN

WAIS-IV

Intelligence was assessed using 5 subtests (Table 5) of the Wechsler Adult Intelligence Scale, fourth edition (WAIS-IV; Wechsler 2008): Block Design (WAIS_BD), Matrix Reasoning (WAIS_MR), Similarities (WAIS_SIM), Vocabulary (WAIS_VC), and Coding (WAIS_CD). The WAIS-IV subtests' Cronbach's αs have been reported to be between 0.84 and 0.94 and test–retest reliabilities to range between $r = 0.69$ and 0.91 (Wechsler 2008).

Data set NKI

WASI-II

The Wechsler Abbreviated Scale of Intelligence, second edition (WASI-II; Wechsler 2011), measured intelligence. The inventory has 4 subtests: Block Design (WASI_BD, 13 items), Matrix Reasoning (WASI_MR, 30 items), Similarities (WASI_SIM, 24 items), and Vocabulary (WASI_VC, 31 items), which are comparable to the subtests from the WAIS-IV (Table 5). The WASI-II can be administered in about 30 minutes and is considered to be the measure of choice for brief intelligence assessments. Split-half reliabilities of the subtests varied between $r = 0.87$ and 0.91 in the child norming sample (6–16 years) and between $r = 0.90$ and 0.92

Table 5. Cognitive tests used to estimate *g* in the UMN sample.

Intelligence test	Task description	No. of items	Construct measured
WAIS-IV			
1. WAIS_BD	Reproduce a shown 2-dimensional pattern with several 3-dimensional building blocks	14	Perceptual reasoning
2. WAIS_MR	Solve matrix-reasoning problems	26	Perceptual reasoning
3. WAIS_SIM	Describe the qualitative similarity between 2 words	18	Verbal comprehension
4. WAIS_VC	Define or describe words or concepts	30	Verbal comprehension
5. WAIS_CD	Add corresponding abstract symbols to as many numbers of a given sequence as possible within a time limit	135	Processing speed

WAIS-IV = Wechsler Adult Intelligence Scale, fourth edition. WAIS_BD = subtest Block Design of the WAIS-IV. WAIS_MR = subtest Matrix Reasoning of the WAIS-IV. WAIS_SIM = subtest Similarities of the WAIS-IV. WAIS_VC = subtest Vocabulary of the WAIS-IV. WAIS_CD = subtest Coding of the WAIS-IV.

in the adult norming sample (17–90 years). Test–retest reliability was $r = 0.79$ in the child sample and 0.94 in the adult sample. The interrater reliabilities of the 4 subtests were between $r = 0.94$ and 0.99, considered exceptionally high (McCrimmon and Smith 2012).

Distribution of intelligence scores

As outlined above, average *g* levels in the samples might vary, indicating different degrees of population representation, cohort differences, and/or test coverage. Because tests differed, we could not compare intelligence levels among our samples or link *g* to the intelligence quotient (IQ) scale. Nevertheless, we tried to estimate the ranges of intelligence the various samples covered. For the RUB data set, we used the norming data of the 11 subtests of the I-S-T 2000 R to estimate IQ scores. The sample's mean IQ was 115 (SD = 13.0), one standard deviation above average. The range of intelligence scores in the HCP data set also seemed to lie at the higher end of the distribution. Dubois et al. (2018) used published norming data from the NIH toolbox subtests, reporting that the sample's means on all tests were significantly higher than the means in the full population. We could generate IQ scores in the UMN and NKI data sets by applying the standard Wechsler formulae. While the mean (114.1; SD = 15.0) was almost one standard deviation above average in the UMN data set, it was about average (101.9; SD = 13.1) in the NKI data set. So, 3 of our 4 samples leaned heavily toward the higher end of the distribution. This may have impacted which brain region associations we observed. For example, basic arithmetic tests are basically speed and accuracy tests for well-educated, high-IQ people (who access automatized information to do them), but are reasoning tests for less educated, lower-IQ people (who must think them through).

Acquisition of DWI data

Data set RUB

All images were collected on a Philips 3 T Achieva scanner at Bergmannsheil Hospital in Bochum, Germany, using a 32-channel head coil. Diffusion-weighted images were acquired using echo planar imaging (Table 6). Diffusion weighting was uniformly distributed along 60 directions using a *b*-value of 1000 s/mm^2. Additionally, 6 volumes with no diffusion weighting ($b = 0$ s/mm^2) were acquired as an anatomical reference for motion correction. To increase the signal-to-noise ratio of diffusion-weighted images, we acquired 3 consecutive scans that were subsequently averaged (Genç et al. 2011a; Genç et al. 2011b). The total acquisition time was 30 minutes.

Data set HCP

All images were collected on a customized Siemens 3 T Connectome Skyra scanner housed at Washington University in St. Louis, using a standard 32-channel Siemens head coil. Diffusion-weighted images were acquired using echo planar imaging (Table 6; Feinberg et al. 2010; Moeller et al. 2010; Setsompop et al. 2012; Xu et al. 2012). The complete DWI session was divided into 6 runs, each lasting approximately 9 minutes and 50 seconds (total acquisition time of about 1 hour). The 6 runs represented 3 different gradient tables, once acquired in the right-to-left and in the left-to-right phase-encoding direction. Each gradient table comprised 90 diffusion weighting directions as well as 6 acquisitions with $b = 0$ s/mm^2 interspersed throughout each run. Diffusion weighting was based on a multi-shell scheme consisting of equally distributed diffusion-weighted images for *b*-values of 1000, 2000, and 3000 s/mm^2.

Data set UMN

All images were collected on a 3 T Siemens Trio scanner at the Center for Magnetic Resonance Research at the UMN in Minneapolis, using a 12-channel head coil. Diffusion-weighted images were acquired using echo planar imaging (Table 6). Diffusion weighting was uniformly distributed along 71 directions. Nine measurements with a *b*-value of 1000 s/mm^2 were conducted. The total acquisition time was 12 minutes, 34 seconds.

Data set NKI

All images were collected on a Siemens Magnetom TrioTim syngo MR B17 scanner at the NKI in Orangeburg, New York. Diffusion-weighted images were acquired using echo planar imaging (Table 6). Diffusion weighting was uniformly distributed along 128 directions using a *b*-value of 1500 s/mm^2. In addition, 9 volumes without diffusion weighting ($b = 0$ s/mm^2) were obtained. The total acquisition time was 5 minutes, 58 seconds.

Image processing and analysis

We processed and analyzed all data sets in the same manner. Since FA is one of the most commonly derived measures from diffusion data (Smith et al. 2006) and has been observed to be associated with intelligence in many studies (Genç and Fraenz 2021), we focused on FA. We used voxel-based statistical analysis of the FA data based on TBSS (Smith et al. 2006), which is part of Oxford Centre for Functional Magnetic Resonance Imaging of the Brain's (FMRIB) Software Library (FSL), version 5.0.9 (Smith et al. 2004). First, DWI images were subjected to brain extraction using Brain Extraction Tool (Smith 2002). Then, FA images were created by fitting tensor models to the raw diffusion data using FMRIB's Diffusion Toolbox. We transformed the resulting FA images into a common space via FMRIB's Nonlinear Image Registration Tool (Andersson et al. 2007a, 2007b), which uses b-spline representations of the registration warp fields (Rueckert et al. 1999). For this purpose, we chose the DTI template FSL_HCP1065_FA_1mm within FSL, which is based on 1065 participants from the HCP and

Table 6. Imaging parameters.

Data set	TR (in ms)	TE (in ms)	Flip angle	Number of slices	Matrix size	Voxel size (in mm)
RUB	7652	87	90°	60	112 × 112	2 × 2 × 2
HCP	5520	89.5	78°	111	145 × 174	1.25 × 1.25 × 1.25
UMN	7900	86	90°	64	128 × 128	2 × 2 × 2
NKI	2400	85	90°	64	106 × 90	2 × 2 × 2

is available in Montreal Neurologic Institute 152 standard space (1 × 1 × 1 mm). Next, we created and thinned mean FA images to generate mean FA skeletons representing the centers of all tracts common to the sample. We set the FA threshold at 0.20 to include only major white matter tracts and exclude peripheral tracts, which are more vulnerable to intra- and intersubject variability. Each participant's aligned FA image was projected onto the skeleton by filling each skeleton voxel with the FA value of the nearest tract center. We used the resulting data to compute voxel-based statistics.

Statistical analysis

We used permutation-based inference (Nichols and Holmes 2002) to analyze voxel-based statistics. To this end, we used the FSL tool "randomise" (Winkler et al. 2014) with 5000 permutations for each analysis. Within the white matter skeleton of each data set, we used a general linear model to identify positive and negative associations between *g* and FA while controlling age, sex, age*sex, age^2, and age^2*sex. We treated them as nuisance variables since they explain relatively little (~10%) of the total variance in whole-brain average FA (Kochunov et al. 2015), to be consistent (same control variables as for computing the *g* factors), and we were not interested in possible age and sex differences.

We used threshold-free clustering enhancement (Smith and Nichols 2009) to avoid arbitrarily specifying a cluster-forming threshold a priori. We adjusted the resulting statistical parametric maps for multiple comparisons by the family-wise error rate thresholded at $P < 0.05$. We binarized them via the FSL tool "fslmaths," so that voxels exhibiting a significant relation between *g* and FA were assigned 1 and all remaining voxels 0. We carried out each step separately in each data set.

As the focal final step, we compared our observations from the individual data sets to identify white matter areas exhibiting replicable structure–function associations. For this purpose, we used the FSL tool "fslmaths" to compute the sums of the 4 binarized maps depicting positive contrasts and the 4 binarized maps depicting negative contrasts (Fig. 5). This resulted in 2 statistical parametric maps with values between 0 (no positive/negative associations between *g* und FA in any data set) and 4 (positive/negative associations in all data sets). We thresholded these maps once again to generate conservative maps only showing those voxels that exhibited significant associations across all 4 data sets (100% consensus). We multiplied those conservative maps with thresholded (value 10) fiber tracts of the Johns Hopkins University White Matter Tractography Atlas, implemented in FSL, to determine the anatomical location of the voxels (Mori et al. 2005; Wakana et al. 2007; Hua et al. 2008). We averaged the FA values of all significant voxels within a voxel cluster for each participant. These mean FA values were related to *g* by calculating partial correlation with age, sex, age*sex, age^2, and age^2*sex as controls. We did this separately for each data set and results were visualized using scatter plots.

Additional exploratory analyses

We also took an exploratory and more liberal approach by creating brain maps including all voxels that exhibited significant associations in 3 out of 4 data sets (75% consensus).

Beyond that, we conducted further explorative analyses. These were based on previous studies' reports that made different observations for broad, first-order intelligence factors such as verbal and nonverbal reasoning abilities (Tamnes et al. 2010). First, we used each of the first-order intelligence factors from each data set (Figs 1–4) as regressors on FA while adding age, sex, age*sex, age^2, age^2*sex, and the remaining first-order intelligence factors for each data set as nuisance factors. For example, the association between verbal intelligence and FA in the RUB data set was analyzed with age, sex, age*sex, age^2, age^2*sex, numerical intelligence, and figural intelligence serving as nuisance factors. Second, we removed the effects of *g* from all first-order intelligence factors and used these variables as regressors on FA, along with age, sex, age*sex, age^2, and age^2*sex as nuisance variables.

We also tried to compare the first-order intelligence factors by binarizing, adding, and thresholding their statistical parametric maps as described above for *g* to test whether there were robust observations among our 4 data sets below *g*. Since the factor models of our data sets had different first-order factors, it was not possible to compare them directly in all data sets. One example is the HCP data set, which does not have a first-order intelligence factor related to only verbal abilities (Fig. 2). Nonetheless, we still tried to include this sample in our comparison of first-order intelligence factors. Hereby, we tested whether there was a robust relation between FA and verbal abilities by combining the results of the first-order intelligence factors ver (RUB, UMN, and NKI) and proc (HCP) (Figs. 1–4). For processing abilities, we combined the first-order intelligence factors fig (RUB) and proc (HCP, UMN, and NKI).

Results

Relations between *g* and FA

Main analysis with 100% consensus.

No voxels exhibited significant negative associations between *g* and FA in any of the 4 data sets. In total 188 individual voxels, 0.12% of the white matter skeleton, exhibited significant positive associations between *g* and FA in all 4 data sets, controlling age, sex, age*sex, age^2, and age^2*sex (for the results of the single data sets, see Supplementary Fig. 1). These voxels could be pooled into 3 contiguous clusters. Cluster "Forceps minor" was the largest and comprised 97 voxels. It overlapped completely with parts of the forceps minor as well as with crossing extensions of the anterior thalamic radiation, the cingulum–cingulate gyrus, and the inferior fronto-occipital fasciculus in the left hemisphere. Scatter plots illustrating the associations between this cluster's mean FA and *g* are shown in Fig. 6 (RUB: $r = 0.15$; HCP: $r = 0.14$; UMN: $r = 0.13$; NKI: $r = 0.16$). The second cluster "SLF" comprised 79 voxels and was located around the superior longitudinal fasciculus in the

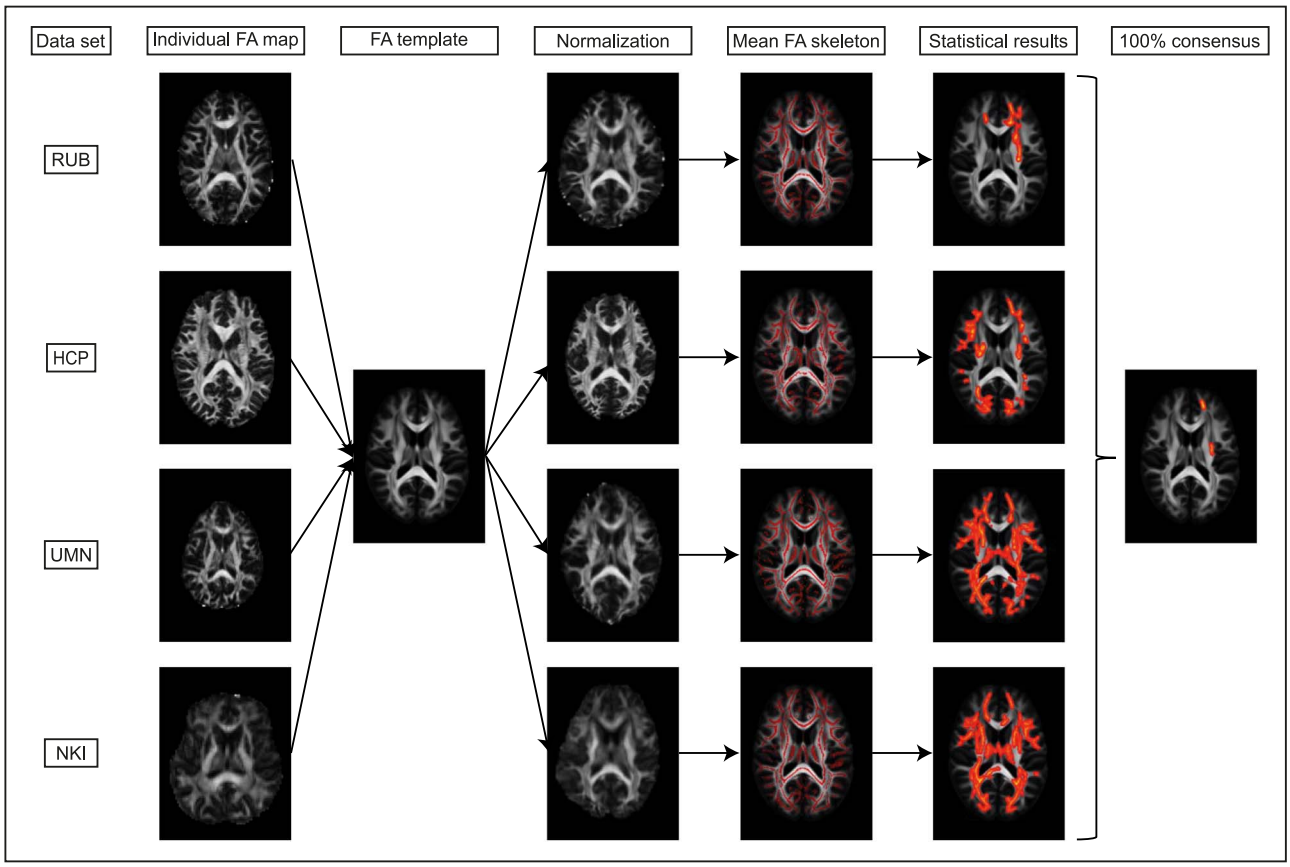

Fig. 5. Methodological sequence depicting the different steps of the image analysis and statistical analysis. The TBSS approach was carried out for each data set separately. We used nonlinear registration to transform individual FA images to a common stereotactic space. By averaging all aligned images, we obtained mean FA maps (not shown). Next, we thinned these to generate white matter skeletons only including voxels at the center of fiber tracts common to all participants. We projected each participant's aligned FA map onto a skeleton by filling the skeleton voxels with FA values from the nearest relevant tract center (not shown). We used the skeletonized FA maps to compute voxel-based cross-subject statistical comparisons. The second last column depicts statistical maps showing voxels that exhibited a significant positive relation between *g* and FA (controlled for age, sex, age*sex, age^2, and age^2*sex). The last image on the right shows voxels that matched across all 4 data sets.

left hemisphere. Figure 7 shows the 4 scatter plots illustrating the associations between this cluster's mean FA and *g* (RUB: $r = 0.18$; HCP: $r = 0.14$; UMN: $r = 0.22$; NKI: $r = 0.12$). The third cluster "Cingulum" was rather small and comprised 12 voxels. Since this cluster did not overlap with any of the thresholded fiber tracts, we used their unthresholded versions to assign the voxels to the fiber tracts. We observed matching voxels with fading extensions of the cingulum–cingulate gyrus, the inferior fronto-occipital fasciculus, and the anterior thalamic radiation in the left hemisphere. The 4 scatter plots illustrating the associations between this cluster's mean FA and *g* are shown in Fig. 8 (RUB: $r = 0.14$; HCP: $r = 0.12$; UMN: $r = 0.13$; NKI: $r = 0.13$).

Exploratory approach with 75% consensus.

The more liberal approach, requiring results to replicate in 3 of the 4 data sets, yielded 8364 voxels, 5.5% of the white matter skeleton, with significant positive associations between *g* and FA, controlling age, sex, age*sex, age^2, and age^2*sex. As depicted in Supplementary Fig. 2, these voxels were widely scattered across the skeleton. Supplementary Table 1 shows the distribution of significant voxels in relation to various major white matter fiber tracts.

Exploratory approach for first-order intelligent factors below g

As mentioned above, we also tested whether there were robust associations below the level of *g*. The different analyses focused on first-order intelligence factors did not yield consistent results for 100% consensus, 75% consensus, or 50% consensus. Hence, we do not present our observations of single data sets.

Discussion

Previous research focused on the relations between general intelligence and white matter microstructure in healthy participants has yielded mixed results. Hence, the primary goal of this study was to find replicable structure–function associations between general intelligence and white matter FA. Indeed, our analyses, involving a TBSS approach across 4 independent, cross-sectional samples, led to the conclusion that such replicable associations exist. We were able to identify a total of 188 voxels, 0.12% of the white matter skeleton, that exhibited significant positive relations between *g* and FA across all 4 data sets, controlling age, sex, age*sex, age^2, and age^2*sex. These voxels formed 3 contiguous clusters. The first was located around the forceps minor, crossing with extensions of the anterior thalamic radiation, the cingulum–cingulate gyrus, and the inferior fronto-occipital fasciculus in the left hemisphere. The second was located around the left-hemispheric superior longitudinal fasciculus. The third was located around the left-hemispheric cingulum–cingulate gyrus, crossing with extensions of the anterior thalamic radiation and the inferior fronto-occipital fasciculus.

There were no voxels exhibiting significant negative associations between *g* and FA in any of the 4 data sets. This was

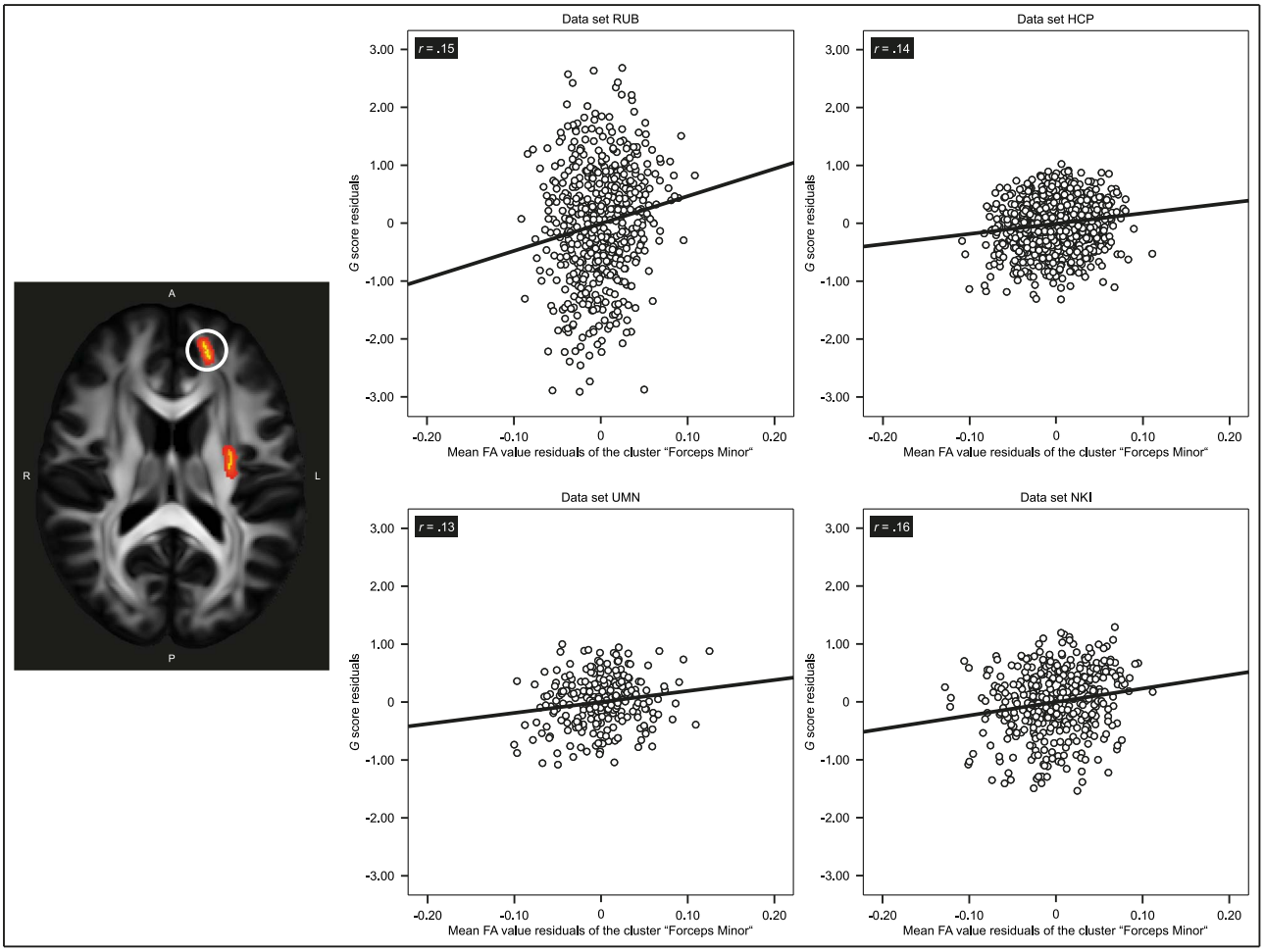

Fig. 6. Associations between *g* and mean FA values from the cluster "Forceps minor." the image on the left side shows the voxel cluster named "Forceps minor" (encircled). The FA values of these voxels were significantly positively associated with *g* in all 4 data sets (independent of effects of age, sex, age*sex, age^2, and age^2*sex). The voxels completely overlapped with parts of the forceps minor as well as with crossing extensions of the anterior thalamic radiation, the cingulum–cingulate gyrus, and the inferior fronto-occipital fasciculus in the left hemisphere. The right side of the figure shows 4 scatter plots, one for each data set. Here, mean FA values from cluster "Forceps minor" are plotted against standardized *g* values. Age, sex, age*sex, age^2, and age^2*sex were used as controlling variables. Reporting partial correlation coefficients is not common. We did so only to convey a general sense of the correlation levels.

consistent with previous research. Multiple studies have examined the associations between various measures of intelligence and FA using various approaches including ROI-based, tract-based, whole-brain-based, and TBSS-based analyses. Despite these differences in design, these studies almost exclusively reported positive associations (Genç and Fraenz 2021). This suggests that individuals with higher intelligence scores tend to have white matter with stronger anisotropic diffusion patterns. However, as FA is a metric aggregating many tissue properties (Beaulieu 2002; Le Bihan 2003; Jones et al. 2013; Friedrich et al. 2020), the exact neurobiological underpinnings driving FA signal differences remain unclear. We can thus only speculate about how higher FA values link to higher *g*. Causal implications could not be drawn from our analyses. Previous studies examining healthy older people suggested that information processing efficiency might mediate the association between FA values and *g* (Deary et al. 2006; Penke et al. 2010). Whether this finding extends to other age groups remains to be seen, but it provides first indications that higher *g* values might emerge from faster, more direct, or more parallel information processing. As summarized by Friedrich et al. (2020), myelination and fiber density have been considered 2 likely neurobiological contributors to FA. Higher FA values might create links with higher mental speed via greater underlying myelination enabling faster conduction velocity (Nave 2010). More direct information transfer throughout the brain might rely on higher FA values that emerge from more parallel, homogeneous fiber orientation distributions. Voxels without complex fiber architectures such as multiple fiber populations, bending fibers, or crossing fibers run directly from one brain region to another, thereby enabling efficient and direct network communication. Greater axon density underlying higher FA might also lead to higher intelligence by providing more pathways to think through various solutions to given problems relatively simultaneously. Future studies are needed to examine intelligence-related differences in such factors (axon diameter, fiber density, myelin concentration, and distribution of fiber orientation) affecting FA values.

Not only were our observations generally consistent with previous research in direction of correlations, but the loci of voxels we identified were similarly consistent. Relevant voxels were situated in regions of the forceps minor, anterior thalamic radiation, cingulum–cingulate gyrus, inferior fronto-occipital fasciculus, and superior longitudinal fasciculus in the left hemisphere. All these

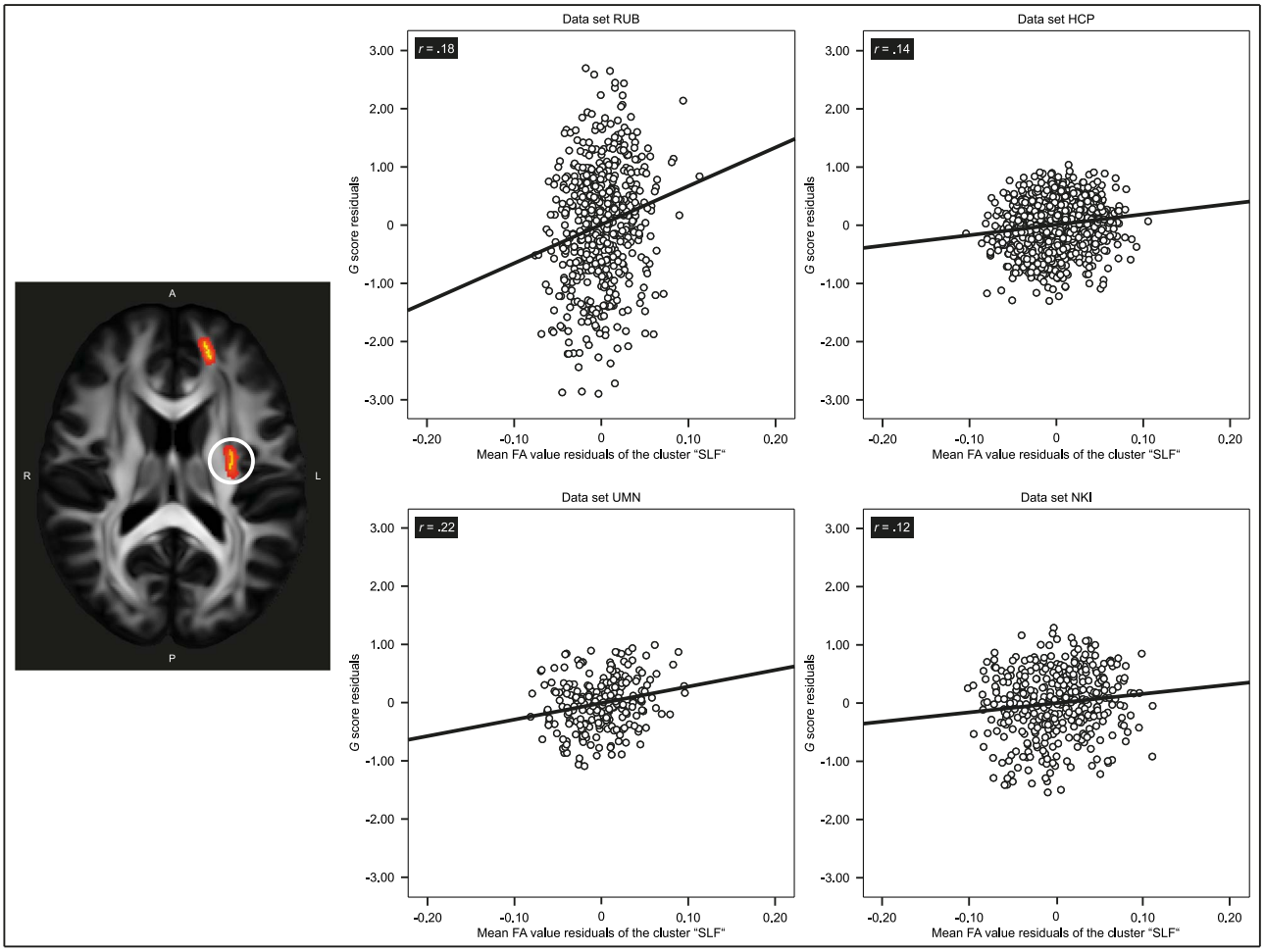

Fig. 7. Associations between *g* and mean FA values from the cluster "SLF." The image on the left side shows the voxel cluster named "SLF" (encircled). The FA values of these voxels were significantly positively associated with *g* in all 4 data sets (independent of the effects of age, sex, age*sex, age^2, and age^2*sex). The voxels were located around the superior longitudinal fasciculus in the left hemisphere. The right side of the figure shows 4 scatter plots, one for each data set. Here, mean FA values from cluster "SLF" are plotted against standardized *g* values. Age, sex, age*sex, age^2, and age^2*sex were used as controlling variables. Reporting partial correlation coefficients is not common. We did so only to convey a general sense of the correlation levels.

fiber tracts have been reported in previous TBSS-studies (Tamnes et al. 2010; Dunst et al. 2014; Malpas et al. 2016).

Fibers running through the genu, i.e. the anterior part of the corpus callosum, form the forceps minor (Catani and Thiebaut de Schotten 2008). As summarized by Genç and Fraenz et al. (2021), the genu of the corpus callosum is the brain region in which FA is most often associated with interindividual differences in intelligence. The corpus callosum is the largest commissural fiber bundle in the brain and consists of approximately 200 million axons (Aboitiz et al. 1992). It connects the left and the right hemispheres and is thus crucial for interhemispheric transfer and integration (van der Knaap and van der Ham 2011). As functional lateralization is a prominent feature of the human (and other mammalian) brain(s) (Kolb and Whishaw 2015; Karolis et al. 2019) and the 2 hemispheres play different roles in inferential reasoning in particular (Marinsek et al. 2014), it seems essential to have recourse to both hemispheres' specializations for intelligent behavior. Fibers of the genu especially link the 2 hemispheres' prefrontal cortices across the hemispheres (Catani and Thiebaut de Schotten 2008). Macrostructural and functional properties of the prefrontal cortex have been repeatedly associated with intelligence (Jung and Haier 2007; Deary et al. 2010a; Basten et al. 2015). In general, the prefrontal cortex is highly relevant for higher cognitive skills such as abstract reasoning, problem solving, memory retrieval, attention, working memory, social interactions, language, and planning (Cabeza and Nyberg 2000; Wood and Grafman 2003).

The anterior thalamic radiation is a projection tract that connects the thalamus to the frontal lobe (Mori et al. 2002; Mori et al. 2005). Of all subcortical structures, thalamus volume seems to be most strongly associated with interindividual differences in intelligence (Bohlken et al. 2014; Cox et al. 2019). In addition, the thalamus has a complex connectivity profile, and its various nuclei establish connections to many areas of the brain (Behrens et al. 2003; Aggleton et al. 2010). Although the thalamus has traditionally been considered to serve merely as a relay station for cortical inputs, more recent observations suggest that its role in cognition could be much broader. It is conceivable that the thalamus also performs dynamic computations that take contextual information into account and reconfigure cortical representations (Rikhye et al. 2018; Dehghani and Wimmer 2019).

The cingulum is a medial associative fiber bundle that runs within the cingulated gyrus from the orbital frontal regions along the dorsal surface of the corpus callosum down toward the temporal lobe (Catani and Thiebaut de Schotten 2008; Bubb et al. 2018). Its fibers form intracortical connections between the

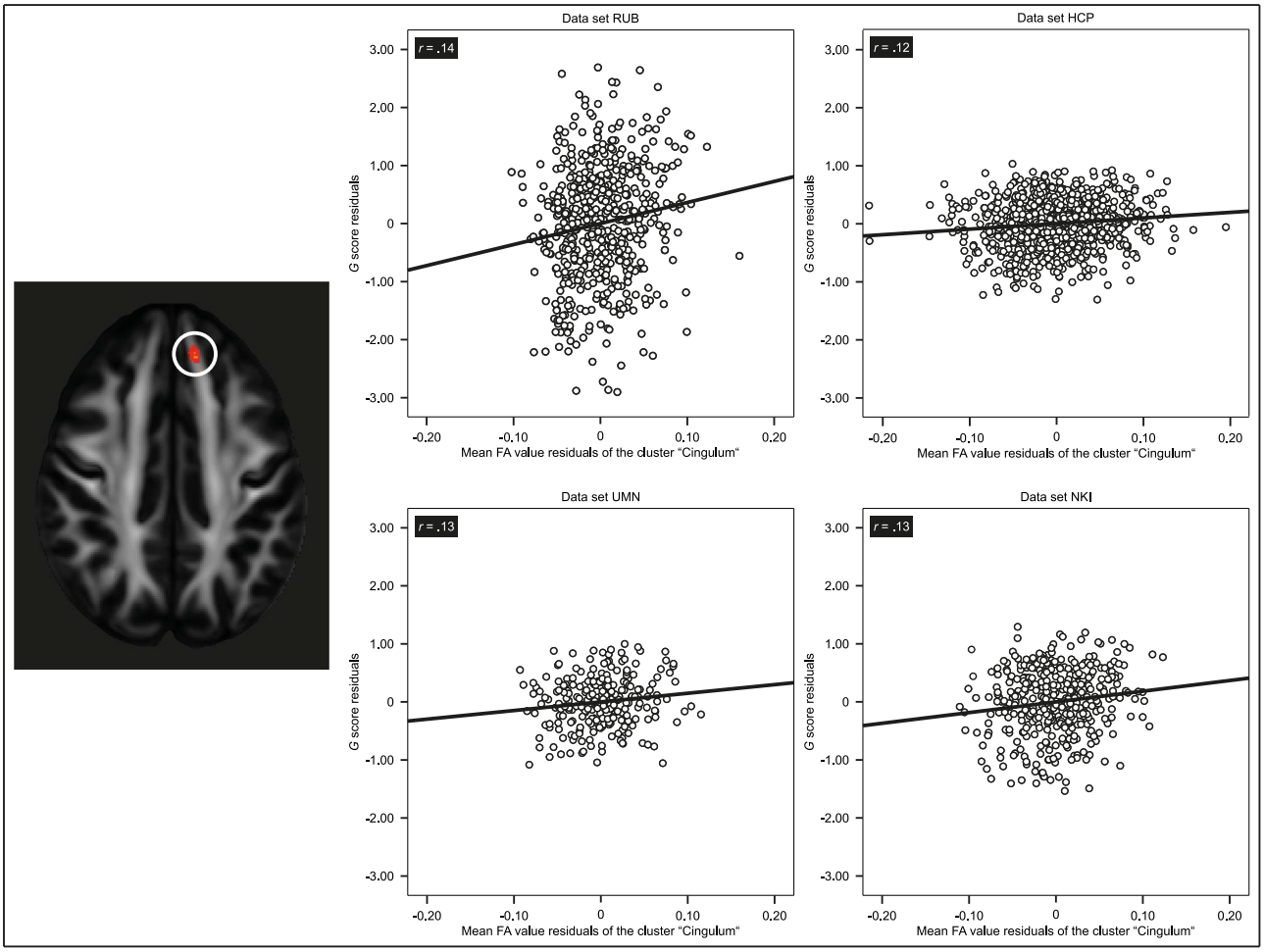

Fig. 8. Associations between *g* and mean FA values from the cluster "Cingulum." The image on the left side shows the voxel cluster named "Cingulum" (encircled). The FA values of these voxels were significantly positively associated with *g* in all 4 data sets (independent of the effects of age, sex, age*sex, age^2, and age^2*sex). The voxels overlapped with fading extensions of the unthresholded fiber tracts cingulum–cingulate gyrus, inferior fronto-occipital fasciculus, and anterior thalamic radiation in the left hemisphere. The right side of the figure shows 4 scatter plots, one for each data set. Here, mean FA values from cluster "Cingulum" are plotted against standardized *g* values. Age, sex, age*sex, age^2, and age^2*sex were used as controlling variables. Reporting partial correlation coefficients is not common. We did so only to convey a general sense of the correlation levels.

medial frontal, parietal, occipital, and temporal lobes as well as different portions of the cingulated cortex. The fiber bundle is also part of the limbic system and one component of the Papez circuit (Papez 1937) constituting connections among the anterior thalamic nuclei, the parahippocampal region, and the cingulate cortex (Catani and Thiebaut de Schotten 2008; Buyanova and Arsalidou 2021). The cingulum appears to be involved in various cognitive domains such as cognitive control, attention, executive functions, memory, language, and visual–spatial functions (Takahashi et al. 2010; Kantarci et al. 2011; Bettcher et al. 2016; Bubb et al. 2018; Buyanova and Arsalidou 2021).

The inferior fronto-occipital fasciculus forms a major association fiber bundle linking the orbitofrontal cortex with the ventral occipital lobe (Catani and Thiebaut de Schotten 2008). Studies suggest that the inferior fronto-occipital fasciculus participates in semantic and visual processing as well as attention (Catani and Thiebaut de Schotten 2008; Leng et al. 2016; Buyanova and Arsalidou 2021).

The superior longitudinal fasciculus is a major white matter tract that connects frontal and opercular areas with the temporoparietal junction and parietal regions (Buyanova and Arsalidou 2021), allowing widespread intracortical information exchange. It is a matter of debate whether the arcuate fasciculus, which connects brain areas relevant for language processing (Broca's and Wernicke's area), can be considered part of the superior longitudinal fasciculus or is merely adjacent to it (Dick and Tremblay 2012; Kamali et al. 2014; Cox et al. 2019). Buyanova and Arsalidou (2021) noted that the right superior longitudinal fasciculus has been associated with cognitive functions such as attention (Frye et al. 2010) and visuospatial abilities (Hoeft et al. 2007), whereas the left superior longitudinal fasciculus has been observed to be crucial for language (Dick and Tremblay 2012) and reading skills (Frye et al. 2010). Buyanova and Arsalidou (2021) further stated that the arcuate fasciculus has been related to reasoning abilities and language processing (Lebel and Beaulieu 2009; Zemmoura et al. 2015). Therefore, both fiber tracts seem to be crucial for higher-order language functions (Friederici 2009). Language, in turn, is viewed as an important cognitive tool for problem solving since the lexicon symbols encapsulate abstract notions, making them more readily manipulable (Varley 2007). Grammatical mechanisms have similar roles in articulating relations among entities. Hence, language in the form of inner speech may allow tasks to be broken into finite series of sub-steps that guide reasoning processes (Varley 2007). Based on this inference, it is not surprising that the superior longitudinal fasciculus is one of the 4 fiber tracts being

most often associated in the kinds of tasks used in intelligence tests (Genç and Fraenz 2021), especially given the constraints (e.g. many, extremely finite, rigidly structured items, administration under tight time and space conditions) involved in attempting to measure intelligence.

Our observations suggest that these brain regions play vital roles in intelligence test performance via white matter tract integrity and coherently anisotropic organization, which is supported by previous research. Jung and Haier (2007) also posited these fiber tracts' relevance in their P-FIT model. They proposed that working on intelligence test reasoning tasks involves multiple processing stages and harmonic interplay of the brain regions constituting their "P-FIT" network. More precisely, they suggested that brain regions in the temporal and occipital lobes are crucial in successfully recognizing and initially processing sensory information. Subsequently, they presumed that the parietal cortex is essential for the interpretation, abstraction, and elaboration of the information's symbolic content. The parietal cortex is believed to interact with frontal regions, which are thought to orchestrate generation and testing of potential solutions to given problems. Once a solution has been selected, it is thought that the anterior cingulate cortex chooses an appropriate reaction and inhibits alternative responses.

Based on this, Jung and Haier (2007) proposed that the rapid and error-free transfer of information from posterior to frontal brain areas depends on underlying white matter integrity. They also emphasized the importance of information exchange between parietal and frontal association areas, which would highlight a role for the superior longitudinal fasciculus (Jung and Haier 2007). Therefore, our observations relating the superior longitudinal fasciculus to general intelligence supported the P-FIT model. Our cingulum observations fit within the P-FIT network. As noted by Fraenz et al. (2021), the P-FIT network is not organized exclusively intrahemispherically. Hence, interhemispheric information transfer between prefrontal areas, e.g. via the forceps minor, seems to be consistent as well. The P-FIT model does not propose direct connections between occipital and (orbito-)frontal areas. However, our observations, highlighting the importance of the inferior fronto-occipital fasciculus, did not necessarily contradict the model, given that this fiber tract also connects distal cortical regions of the P-FIT network. Instead, additional connections offer the possibility of more parallel flows of information. Since individuals who score identically in an intelligent test may use different cognitive strategies as well as different brain structures to reach their performance level (Deary et al. 2010a), there may be more than one adequate solution path and overall good brain function may be more important for general intelligence than using any specific parts well.

Jung and Haier (2007) assumed that brain regions beyond the cerebral cortex, such as thalamus, hippocampus, and cerebellum, are involved only in rather basic functions. Hence, they believed that they would not contribute to interindividual intelligence differences significantly. However, more recent studies indicate that the thalamus and the hippocampus as well as their connections could play more important roles in reasoning than originally thought (Bohlken et al. 2014; Rikhye et al. 2018; Cox et al. 2019; Dehghani and Wimmer 2019; Deary et al. 2022). Our observations, involving the anterior thalamic radiation, supported these studies in suggesting that the P-FIT model (Jung and Haier 2007) needs some updating, which is only to be expected after 15 years more research.

We initially took a rather conservative analytical approach. To be considered for discussion, voxels had to exhibit significant associations between *g* and FA across all 4 data sets (100% consensus). A more liberal threshold (75% consensus) yielded about 44 times more voxels. This was simply because more data sets inevitably vary in more ways. Moreover, as illustrated in Supplementary Fig. 2, significant voxel clusters were no longer exclusively located in the left hemisphere. However, Supplementary Table 1 indicates that more significant voxels could be assigned to fiber tracts in the left hemisphere (59.3%, out of 9 fiber tracts with left–right symmetry 7 had more voxels in the left hemisphere). As the left and the right hemisphere differ in their specialized functions (Marinsek et al. 2014; Kolb and Whishaw 2015; Karolis et al. 2019), both hemispheres and their functional interaction are relevant for intelligent behavior.

The additional exploratory analyses of different first-order intelligence factors did not lead to any overlapping results in even 2 data sets. Our observations were not consistent with Tamnes et al. (2010), who reported significant positive associations between FA and verbal/nonverbal reasoning abilities. This could be because the first-order intelligence factors differed among samples (Figs. 1–4). As they include much less information than *g*, differences in the specific tasks might have impacted these factors' contents more than they did *g*. But our results could also differ from Tamnes et al'.s because our analyses of these narrower intelligence factors controlled *g* itself, which theirs did not, so we examined only factor-specific variance. *g* explains about 40% of total variance in typical test batteries (Deary et al. 2010a), in our cases 32–65% (Table 2). To resolve such inconsistencies, future studies should also focus on specific intelligence factors, although keeping in mind that no factor identified in this manner actually "carves nature at its joints". They all vary considerably depending on specific test battery content and sampling.

Limitations

Making use of multiple samples, as we did is more likely to yield replicable observations. However, the question arises why particular observations in one sample failed to replicate in other data sets (Supplementary Fig. 1). This might be because there is no robust association between *g* and FA, but it might also be that differences among data sets hindered cross-sample replication. The 4 data sets included in our study used different intelligence tests, had different sample sizes, sex ratios, age distributions, and image acquisition protocols. As can be seen in Supplementary Fig. 1, the RUB data set was the most common exception to 100% overlap. This data set differed from the other, more similar 3 in several aspects: the sample had been collected in Germany and therefore influenced by German pedagogies (vs. USA), magnetic resonance imaging (MRI) measurements were obtained on a Philips scanner (vs. Siemens scanners), and its *g*-factor residuals had greater variance despite the sample's high indicated mean IQ (Figs. 6–8). As 2 other (HCP and UMN) of our 4 samples leaned heavily toward the higher end of the intelligence distribution, population representativeness was limited in these data sets. This may have heavily impacted which brain region associations we observed since, for example, basic arithmetic tests are basically speed and accuracy tests for well-educated, high-IQ people but reasoning tests for less educated, lower-IQ people. Outlined at the discussion's beginning, 2 possibilities for why higher FA values might show links with higher *g* are faster or more direct information processing due to greater myelination and more parallel, homogenous distributions of fiber orientation. The RUB data set's

intelligence test battery included more verbal tasks with no time limit (e.g. BOWIT), whereas all the other data sets' *g* factors did not rely so heavily on such tasks and instead included more nonverbal tasks. This difference could explain why the latter generated more associations with FA. Furthermore, the RUB sample mainly consisted of German university students who are not representative of the European population in age, educational background, or ethnic composition. As our samples came from different populations, represented to different degrees, one should not draw conclusions about humans in general based on our results. We attempted to minimize the effects of these differences by calculating *g* factor scores, standardizing data processing for all data sets, and statistically controlling age, sex, age*sex, age^2, and age^2*sex. Nevertheless, these differences might have hindered detection of potential associations and/or distorted those we did observe. In general, use of complementary methods, including fine-grained cortical parcellation schemes in combination with DWI and graph theory, may lead to new insights and are highly encouraged.

Conclusion

In conclusion, we reported replicable associations between general intelligence and FA among 4 different cross-sectional data sets. By analyzing data from more than 2000 healthy participants, we were able to observe a total of 188 voxels with significant positive associations between *g* and FA in all 4 data sets, controlling age, sex, age*sex, age^2, and age^2*sex. These voxels were located around the forceps minor, crossing with extensions of the anterior thalamic radiation, the cingulum–cingulate gyrus, and the inferior fronto-occipital fasciculus in the left hemisphere, around the left-hemispheric superior longitudinal fasciculus, and around the left-hemispheric cingulum–cingulate gyrus, crossing with extensions of the anterior thalamic radiation and the inferior fronto-occipital fasciculus. Our observations do not imply that other brain's white matter areas not observed are irrelevant for intellectual performance, but only that the mentioned fiber tracts appear to be more commonly or intensely relevant to carrying out cognitive tasks than others. For the most part, our observations were consistent with previous research on the associations between white matter correlates and intelligence differences. We hope that future studies will make use of multiple samples because it is more likely to avoid false positive observations and could ultimately yield truly robust findings.

Acknowledgments

The authors thank all research assistants for their support during the behavioral measurements. Further, the authors thank PHILIPS Germany (Burkhard Mädler) for the scientific support with the MRI measurements as well as Tobias Otto for technical assistance.

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Data availability

The data from Ruhr University Bochum (RUB) and University of Minnesota (UMN) that support the findings of this study are available from the corresponding author upon reasonable request. The final result map with 75% consensus can be downloaded from an Open Science Framework repository (https://osf.io/km5g3/). The data used for the Human Connectome Project (HCP) sample are part of the S1200 release provided by the Human Connectome Project and can be accessed via its ConnectomeDB platform (https://db.humanconnectome.org/). The data used for the Nathan Kline Institute (NKI) sample are part of the Nathan Kline Institute Rockland Sample release and can be accessed via this link (http://fcon_1000.projects.nitrc.org/indi/enhanced/).

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (GU 227/16-1).

Conflict of interest statement: None declared.

References

Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. *Brain Res*. 1992:598:143–153.

Aggleton JP, O'Mara SM, Vann SD, Wright NF, Tsanov M, Erichsen JT. Hippocampal-anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *Eur J Neurosci*. 2010:31(12):2292–2307.

Allin MP, Kontis D, Walshe M, Wyatt J, Barker GJ, Kanaan RA, McGuire P, Rifkin L, Murray RM, Nosarti C. White matter and cognition in adults who were born preterm. *PLoS One*. 2011:6(10):e24525.

Andersson JLR, Jenkinson M, Smith SM. *Non-linear optimisation*. FMRIB technical report TR07JA1, Oxford: University of Oxford FMRIB Centre,; 2007a.

Andersson JLR, Jenkinson M, Smith SM. *Non-linear registration aka spatial normalisation*. FMRIB technical report TR07JA2, Oxford: University of Oxford FMRIB Centre; 2007b.

Ashburner J, Friston KJ. Voxel-based morphometry — the methods. *NeuroImage*. 2000:11(6):805–821.

Aspara J, Wittkowski K, Luo X. Types of intelligence predict likelihood to get married and stay married: large-scale empirical evidence for evolutionary theory. *Pers Individ Differ*. 2018:122:1–6.

Assaf Y, Pasternak O. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *J Mol Neurosci*. 2008:34(1):51–61.

Barbey AK. Network neuroscience theory of human intelligence. *Trends Cogn Sci*. 2018:22(1):8–20.

Barbey AK, Colom R, Solomon J, Krueger F, Forbes C, Grafman J. An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain*. 2012:135(4):1154–1164.

Barbey AK, Colom R, Paul EJ, Grafman J. Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Struct Funct*. 2014:219(2):485–494.

Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson Series B*. 1996:111(3):209–219.

Basten U, Hilger K, Fiebach CJ. Where smart brains are different: a quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence*. 2015:51:10–27.

Bathelt J, Johnson A, Zhang M, Astle DE. The cingulum as a marker of individual differences in neurocognitive development. *Sci Rep*. 2019:9(1):2281.

Batty GD, Deary IJ, Gottfredson LS. Premorbid (early life) IQ and later mortality risk: systematic review. *Ann Epidemiol*. 2007:17(4):278–288.

Beaulieu C. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed*. 2002:15(7–8):435–455.

Behrens TE, Johansen-Berg H, Woolrich MW, Smith SM, Wheeler-Kingshott CAM, Boulby PA, Barker GJ, Sillery EL, Sheehan K, Ciccarelli O, et al. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat Neurosci*. 2003:6(7):750–757.

Bentler PM, Bonett G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull*. 1980:88(3):588–606.

Bettcher BM, Mungas D, Patel N, Elofson J, Dutt S, Wynn M, Watson CL, Stephens M, Walsh CM, Kramer JH. Neuroanatomical substrates of executive functions: beyond prefrontal structures. *Neuropsychologia*. 2016:85:100–109.

Bohlken MM, Brouwer RM, Mandl RC, van Haren NE, Brans RG, van Baal GC, de Geus EJ, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genes contributing to subcortical volumes and intellectual ability implicate the thalamus. *Hum Brain Mapp*. 2014:35(6):2632–2642.

Booth T, Bastin ME, Penke L, Maniega SM, Murray C, Royle NA, Gow AJ, Corley J, Henderson RD, Hernandez Mdel C, et al. Brain white matter tract integrity and cognitive abilities in community-dwelling older people: the Lothian Birth Cohort, 1936. *Neuropsychology*. 2013:27(5):595–607.

Bowren M Jr, Adolphs R, Bruss J, Manzel K, Corbetta M, Tranel D, Boes AD. Multivariate lesion-behavior mapping of general cognitive ability and its psychometric constituents. *J Neurosci*. 2020:40(46):8924–8937.

Bubb EJ, Metzler-Baddeley C, Aggleton JP. The cingulum bundle: anatomy, function, and dysfunction. *Neurosci Biobehav Rev*. 2018:92:104–127.

Buyanova IS, Arsalidou M. Cerebral white matter myelination and relations to age, gender, and cognition: A selective review. *Front Hum Neurosci*. 2021:15:662031.

Cabeza R, Nyberg L. Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J Cogn Neurosci*. 2000:12(1):1–47.

Calvin CM, Deary IJ, Fenton C, Roberts BA, Der G, Leckenby N, Batty GD. Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. *Int J Epidemiol*. 2011:40(3):626–644.

Calvin CM, Batty GD, Der G, Brett CE, Taylor A, Pattie A, Cukic I, Deary IJ. Childhood intelligence in relation to major causes of death in 68 year follow-up: prospective population study. *BMJ*. 2017:357:j2708.

Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*. 2008:44(8):1105–1132.

Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C, Klunder AD, Toga AW, McMahon KL, de Zubicaray GI, et al. Genetics of brain fiber architecture and intellectual performance. *J Neurosci*. 2009:29(7):2212–2224.

Clayden JD, Jentschke S, Munoz M, Cooper JM, Chadwick MJ, Banks T, Clark CA, Vargha-Khadem F. Normative development of white matter tracts: similarities and differences in relation to age, gender, and intelligence. *Cereb Cortex*. 2012:22(8):1738–1747.

Cox SR, Ritchie SJ, Fawns-Ritchie C, Tucker-Drob EM, Deary IJ. Structural brain imaging correlates of general intelligence in UK Biobank. *Intelligence*. 2019:76:101376.

Cremers LG, de Groot M, Hofman A, Krestin GP, van der Lugt A, Niessen WJ, Vernooij MW, Ikram MA. Altered tract-specific white matter microstructure is related to poorer cognitive performance: the Rotterdam study. *Neurobiol Aging*. 2016:39:108–117.

Deary IJ, Bastin ME, Pattie A, Clayden JD, Whalley LJ, Starr JM, Wardlaw JM. White matter integrity and cognition in childhood and old age. *Neurology*. 2006:66(4):505–512.

Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev Neurosci*. 2010a:11(3):201–211.

Deary IJ, Weiss A, Batty GD. Intelligence and personality as predictors of illness and death: how researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychol Sci Public Interest*. 2010b:11(2):53–79.

Deary IJ, Cox SR, Hill WD. Genetic variation, brain, and intelligence differences. *Mol Psychiatry*. 2022:27(1):335–353.

Dehghani N, Wimmer RD. A computational perspective of the role of the thalamus in cognition. *Neural Comput*. 2019:31(7):1380–1418.

Dick AS, Tremblay P. Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. *Brain*. 2012:135(12):3529–3550.

Dubner SE, Dodson CK, Marchman VA, Ben-Shachar M, Feldman HM, Travis KE. White matter microstructure and cognitive outcomes in relation to neonatal inflammation in 6-year-old children born preterm. *NeuroImage Clin*. 2019:23:101832.

Dubois J, Galdi P, Paul LK, Adolphs R. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos Trans R Soc Lond Ser B Biol Sci*. 2018:373(1756):20170284.

Dunst B, Benedek M, Koschutnig K, Jauk E, Neubauer AC. Sex differences in the IQ-white matter microstructure relationship: a DTI study. *Brain Cogn*. 2014:91:71–78.

van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, et al. The Human Connectome Project: a data acquisition perspective. *NeuroImage*. 2012:62(4):2222–2231.

van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, WU-MH C. The WU-Minn Human Connectome Project: an overview. *NeuroImage*. 2013:80:62–79.

Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Gunther M, Glasser MF, Miller KL, Ugurbil K, Yacoub E. Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS One*. 2010:5(12):e15710.

Ferrer E, Whitaker KJ, Steele JS, Green CT, Wendelken C, Bunge SA. White matter maturation supports the development of reasoning ability through its influence on processing speed. *Dev Sci*. 2013:16(6):941–951.

Filley C. *The behavioral neurology of white matter*. New York: Oxford University Press; 2012

Flanagan DP, Dixon SG. The Cattell-Horn-Carroll theory of cognitive abilities. In: Reynolds CR, Vannest KJ, Fletcher-Janzen E, editors. *Encyclopedia of special education*. Hoboken, New Jersey: John Wiley & Sons, Inc; 2013

Fraenz C, Schlüter C, Friedrich P, Jung RE, Güntürkün O, Genç E. Interindividual differences in matrix reasoning are linked to functional connectivity between brain regions nominated by Parieto-Frontal Integration Theory. *Intelligence*. 2021:87:101545.

Friederici AD. Pathways to language: fiber tracts in the human brain. *Trends Cogn Sci*. 2009:13(4):175–181.

Friedrich P, Fraenz C, Schlüter C, Ocklenburg S, Madler B, Güntürkün O, Genç E. The relationship between axon density, myelination, and fractional anisotropy in the human corpus callosum. *Cereb Cortex*. 2020:30(4):2042–2056.

Frye RE, Hasan K, Malmberg B, Desouza L, Swank P, Smith K, Landry S. Superior longitudinal fasciculus and cognitive dysfunction in adolescents born preterm and at term. *Dev Med Child Neurol*. 2010:52(8):760–766.

Fuhrmann D, Simpson-Kent IL, Bathelt J, Team C, Kievit RA. A hierarchical watershed model of fluid intelligence in childhood and adolescence. *Cereb Cortex*. 2020:30(1):339–352.

Genç E, Fraenz C. Diffusion-weighted imaging of intelligence. In: Barbey AK, Karama S, Haier RJ, editors. *The Cambridge handbook of*

intelligence and cognitive neuroscience. 1st ed. New York: Cambridge University Press; 2021. pp. 191–209.

Genç E, Bergmann J, Singer W, Kohler A. Interhemispheric connections shape subjective experience of bistable motion. *Curr Biol*. 2011a:21(17):1494–1499.

Genç E, Bergmann J, Tong F, Blake R, Singer W, Kohler A. Callosal connections of primary visual cortex predict the spatial spreading of binocular rivalry across the visual hemifields. *Front Hum Neurosci*. 2011b:5:161.

Genç E, Fraenz C, Schlüter C, Friedrich P, Hossiep R, Voelkle MC, Ling JM, Güntürkün O, Jung RE. Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nat Commun*. 2018:9(1):1905.

Genç E, Fraenz C, Schlüter C, Friedrich P, Voelkle MC, Hossiep R, Güntürkün O. The neural architecture of general knowledge. *EJP*. 2019:33(5):589–605.

Genç E, Schlüter C, Fraenz C, Arning L, Metzen D, Nguyen HP, Voelkle MC, Streit F, Güntürkün O, Kumsta R, et al. Polygenic scores for cognitive abilities and their association with different aspects of general intelligence—a deep phenotyping approach. *Mol Neurobiol*. 2021:58(8):4145–4156.

Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ. NIH toolbox for the assessment of neurological and behavioral function. *Neurology*. 2013:80(Suppl. 3):S2–S6.

Gläscher J, Rudrauf D, Colom R, Paul LK, Tranel D, Damasio H, Adolphs R. Distributed neural system for general intelligence revealed by lesion mapping. *PNAS USA*. 2010:107(10):4705–4709.

Góngora D, Vega-Hernández M, Jahanshahi M, Valdés-Sosa PA, Bringas-Vega ML, CHBMP. Crystallized and fluid intelligence are predicted by microstructure of specific white-matter tracts. *Hum Brain Mapp*. 2020:41(4):906–916.

Gottfredson LS. Why g matters: the complexity of everyday life. *Intelligence*. 1997:24(1):79–132.

Grazioplene RG, Ryman SG, Gray JR, Rustichini A, Jung RE, CG DY. Subcortical intelligence: caudate volume predicts IQ in healthy adults. *Hum Brain Mapp*. 2015:36(4):1407–1416.

Grazioplene RG, Chavez RS, Rustichini A, DeYoung CG. White matter correlates of psychosis-linked traits support continuity between personality and psychopathology. *J Abnorm Psychol*. 2016:125(8):1135–1145.

Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C, Siegel SJ, Gur RE. Computerized neurocognitive scanning: I. methodology and validation in healthy people. *NPP*. 2001:25(5):766–776.

Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, Brensinger C, Gur RE. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods*. 2010:187(2):254–262.

Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, Beaumont J, Casaletto KB, Conway K, Slotkin J, et al. Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *J Int Neuropsychol Soc*. 2014:20(6):588–598.

Hemmingsson T, Melin B, Allebeck P, Lundberg I. The association between cognitive ability measured at ages 18-20 and mortality during 30 years of follow-up–a prospective observational study among Swedish males born 1949-51. *Int J Epidemiol*. 2006:35(3):665–670.

Hidese S, Ota M, Matsuo J, Ishida I, Hiraishi M, Yokota Y, Hattori K, Yomogida Y, Kunugi H. Correlation between the Wechsler adult intelligence scale- 3rd edition metrics and brain structure in healthy individuals: a whole-brain magnetic resonance imaging study. *Front Hum Neurosci*. 2020:14:211.

Hoeft F, Barnea-Goraly N, Haas BW, Golarai G, Ng D, Mills D, Korenberg J, Bellugi U, Galaburda A, Reiss AL. More is not always better: increased fractional anisotropy of superior longitudinal fasciculus associated with poor visuospatial abilities in Williams syndrome. *J Neurosci*. 2007:27(44):11960–11965.

Holleran L, Kelly S, Alloza C, Agartz I, Andreassen OA, Arango C, Banaj N, Calhoun V, Cannon D, Carr V, et al. The relationship between white matter microstructure and general cognitive ability in patients with schizophrenia and healthy participants in the ENIGMA consortium. *Am J Psychiatry*. 2020:177(6):537–547.

Hossiep R, Schulte M. *BOWIT: Bochumer Wissenstest*. Göttingen (Germany): Hogrefe; 2008.

Hossiep R, Hasella M, Turck D. *BOMAT-advanced-short version: Bochumer Matrizentest*. Göttingen (Germany): Hogrefe; 2001.

Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, Calabresi PA, Pekar JJ, van Zijl PC, Mori S. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *NeuroImage*. 2008:39(1):336–347.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci U S A*. 2008:105(19):6829–6833.

Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II. Just one g: consistent results from three test batteries. *Intelligence*. 2004:32(1):95–107.

Johnson W, Jt N, Bouchard TJ. Still just 1 g: consistent results from five test batteries. *Intelligence*. 2008:36(1):81–95.

Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *NeuroImage*. 2013:73:239–254.

Jöreskog KG. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*. 1969:34(2):183–202.

Jung RE, Haier RJ. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci*. 2007:30(2):135–154.

Kamali A, Flanders AE, Brody J, Hunter JV, Hasan KM. Tracing superior longitudinal fasciculus connectivity in the human brain using high resolution diffusion tensor tractography. *Brain Struct Funct*. 2014:219(1):269–281.

Kantarci K, Senjem ML, Avula R, Zhang B, Samikoglu AR, Weigand SD, Przybelski SA, Edmonson HA, Vemuri P, Knopman DS, et al. Diffusion tensor imaging and cognitive function in older adults with no dementia. *Neurology*. 2011:77(1):26–34.

Karolis VR, Corbetta M, Thiebaut de Schotten M. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nat Commun*. 2019:10(1):1417.

Kennedy E, Poppe T, Tottman A, Harding J. Neurodevelopmental impairment is associated with altered white matter development in a cohort of school-aged children born very preterm. *NeuroImage Clin*. 2021:31:102730.

Kievit RA, Davis SW, Mitchell DJ, Taylor JR, Duncan J, Cam CANRT, Henson RN, Cam CANRT. Distinct aspects of frontal lobe structure mediate age-related differences in fluid intelligence and multitasking. *Nat Commun*. 2014:5(1):5658.

Kievit RA, Davis SW, Griffiths J, Correia MM, Cam C, Henson RN. A watershed model of individual differences in fluid intelligence. *Neuropsychologia*. 2016:91:186–198.

Kievit RA, Fuhrmann D, Borgeest GS, Simpson-Kent IL, Henson RNA. The neural determinants of age-related changes in fluid intelligence: a pre-registered, longitudinal analysis in UK biobank [version 2; peer review: 3 approved]. *Wellcome Open Res*. 2018:3–38.

Kochunov P, Jahanshad N, Marcus D, Winkler A, Sprooten E, Nichols TE, Wright SN, Hong LE, Patel B, Behrens T, et al. Heritability of fractional anisotropy in human white matter: A comparison of

Human Connectome Project and ENIGMA-DTI data. *NeuroImage*. 2015:111:300–311.

Kolb B, Whishaw IQ. *Fundamentals of human neuropsychology*. New York: Worth Publishers; 2015.

Kontis D, Catani M, Cuddy M, Walshe M, Nosarti C, Jones D, Wyatt J, Rifkin L, Murray R, Allin M. Diffusion tensor MRI of the corpus callosum and cognitive function in adults born preterm. *Neuroreport*. 2009:20(4):424–428.

Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*. 2003:4(6):469–480.

Le Bihan D. Diffusion MRI: what water tells us about the brain. *EMBO Mol Med*. 2014:6(5):569–573.

Le Bihan D, Breton É. In vivo magnetic resonance imaging of diffusion. *C R Acad Sci Ser II*. 1985:301(15):1109–1112.

Le Bihan D, Breton É, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application of diffusion and perfusion in neurologic disorders. *Radiology*. 1986:161(2):401–407.

Lebel C, Beaulieu C. Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Hum Brain Mapp*. 2009:30(11):3563–3573.

Leng Y, Shi Y, Yu Q, Van Horn JD, Tang H, Li J, Xu W, Ge X, Tang Y, Han Y, et al. Phenotypic and genetic correlations between the lobar segments of the inferior fronto-occipital fasciculus and attention. *Sci Rep*. 2016:6(1):33015.

Liepmann D, Beauducel A, Brocke B, Amthauer R. *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). Manual*. Göttingen (Germany): Hogrefe; 2007.

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model*. 1999:6(1):1–55.

Malpas CB, Genc S, Saling MM, Velakoulis D, Desmond PM, O'Brien TJ. MRI correlates of general intelligence in neurotypical adults. *J Clin Neurosci*. 2016:24:128–134.

Marinsek N, Turner BO, Gazzaniga M, Miller MB. Divergent hemispheric reasoning strategies: reducing uncertainty versus resolving inconsistency. *Front Hum Neurosci*. 2014:8:839.

McCrimmon AW, Smith AD. Review of the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). *J Psychoeduc Assess*. 2012:31(3):337–341.

McDaniel M. Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence*. 2005:33(4):337–346.

McGue M, Bouchard TJ. Adjustment of twin data for the effects of age and sex. *Behav Genet*. 1984:14(4):325–343.

Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, Ugurbil K. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med*. 2010:63(5):1144–1153.

Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. Psychometric properties of the Penn computerized neurocognitive battery. *Neuropsychology*. 2015:29(2):235–246.

Mori S, Kaufmann WE, Davatzikos C, Stieltjes B, Amodei L, Fredericksen K, Pearlson GD, Melhem ER, Solaiyappan M, Raymond GV, et al. Imaging cortical association tracts in the human brain using diffusion-tensor-based axonal tracking. *Magn Res Med*. 2002:47(2):215–223.

Mori S, Wakana S, van Zijl PCM, Nagae-Poetscher LM. *MRI atlas of human white matter*. Amsterdam: Elsevier B. V; 2005.

Muetzel RL, Mous SE, van der Ende J, Blanken LM, van der Lugt A, Jaddoe VW, Verhulst FC, Tiemeier H, White T. White matter integrity and cognitive performance in school-age children: a population-based neuroimaging study. *NeuroImage*. 2015:119:119–128.

Navas-Sanchez FJ, Aleman-Gomez Y, Sanchez-Gonzalez J, Guzman-De-Villoria JA, Franco C, Robles O, Arango C, Desco M. White matter microstructure correlates of mathematical giftedness and intelligence quotient. *Hum Brain Mapp*. 2014:35(6):2619–2631.

Nave KA. Myelination and support of axonal integrity by glia. *Nature*. 2010:468(7321):244–252.

Neisser U, Boodoo G, Bouchard TJ, Boykin AW, Brody N, Ceci SJ, Halpern DF, Loehlin JC, Perloff R, Sternberg RJ, et al. Intelligence: knowns and unknowns. *Am Psychol*. 1996:51(2):77–101.

Nestor PG, Ohtani T, Bouix S, Hosokawa T, Saito Y, Newell DT, Kubicki M. Dissociating prefrontal circuitry in intelligence and memory: neuropsychological correlates of magnetic resonance and diffusion tensor imaging. *Brain Imaging Behav*. 2015:9(4):839–847.

Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*. 2002:15(1):1–25.

Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, Panek LJ, Brown S, Zavitz ST, Li Q, et al. The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Front Neurosci*. 2012:6:152.

Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. Increased parietal activity after training of interference control. *Neuropsychologia*. 2013:51(13):2781–2790.

Ohtani T, Nestor PG, Bouix S, Saito Y, Hosokawa T, Kubicki M. Medial frontal white and gray matter contributions to general intelligence. *PLoS One*. 2014:9(12):e112691.

Oswald WD, Roth E. *Der Zahlen-Verbindungs-test (ZVT)*. Göttingen (Germany): Hogrefe Verlag für Psychologie; 1987.

Papez JW. A proposed mechanism of emotion. *Arch Neurol Psychiatr*. 1937:38(4):725–743.

Penke L, Munoz Maniega S, Murray C, Gow AJ, Hernandez MC, Clayden JD, Starr JM, Wardlaw JM, Bastin ME, Deary IJ. A general factor of brain white matter integrity predicts information processing speed in healthy older people. *J Neurosci*. 2010:30(22):7569–7574.

Penke L, Maniega SM, Bastin ME, Valdes Hernandez MC, Murray C, Royle NA, Starr JM, Wardlaw JM, Deary IJ. Brain white matter tract integrity as a neural foundation for general intelligence. *Mol Psychiatry*. 2012:17(10):1026–1030.

Pietschnig J, Penke L, Wicherts JM, Zeiler M, Voracek M. Meta-analysis of associations between human brain volume and intelligence differences: how strong are they and what do they mean? *Neurosci Biobehav Rev*. 2015:57:411–432.

Power MC, Su D, Wu A, Reid RI, Jack CR, Knopman DS, Coresh J, Huang J, Kantarci K, Sharrett AR, et al. Association of white matter microstructural integrity with cognition and dementia. *Neurobiol Aging*. 2019:83:63–72.

Raven JC, Court JH, Raven J. *Coloured progressive matrices. Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford (United Kingdom): Oxford Psychologists Press; 1990.

Rikhye RV, Wimmer RD, Halassa MM. Toward an integrative theory of thalamic function. *Annu Rev Neurosci*. 2018:41(1):163–183.

Roth B, Becker N, Romeyke S, Schäfer S, Domnick F, Spinath FM. Intelligence and school grades: a meta-analysis. *Intelligence*. 2015:53:118–137.

Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*. 1999:18(8):712–721.

Schmidt FL, Hunter J. General mental ability in the world of work: occupational attainment and job performance. *J Pers Soc Psychol*. 2004:86(1):162–173.

Schmithorst VJ. Developmental sex differences in the relation of neuroanatomical connectivity to intelligence. *Intelligence*. 2009:37(2):164–173.

Schmithorst VJ, Wilke M, Dardzinski BJ, Holland SK. Cognitive functions correlate with white matter architecture in a normal pediatric population: a diffusion tensor MRI study. *Hum Brain Mapp*. 2005:26(2):139–147.

Schneider WJ, McGrew KS. The Cattell-Horn-Carroll model of intelligence. In: Flanagan DP, Harrison PL, editors. *Contemporary intellectual assessment: theories, tests, and issues*. Third ed. New York: Guilford Press; 2012. pp. 99–144.

Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med*. 2012:67(5):1210–1224.

Simpson-Kent IL, Fuhrmann D, Bathelt J, Achterberg J, Borgeest GS, Kievit RA, Team C. Neurocognitive reorganization between crystallized intelligence, fluid intelligence and white matter microstructure in two age-heterogeneous developmental cohorts. *Dev Cogn Neurosci*. 2020:41:100743.

Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002:17(3):143–155.

Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. 2009:44(1):83–98.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004:23(Suppl 1): S208–S219.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006:31(4):1487–1505.

Spearman C. General intelligence, objectively determined and measured. *Am J Psychol*. 1904:15(2):201–292.

Strenze T. Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*. 2007:35(5):401–426.

Suprano I, Kocevar G, Stamile C, Hannoun S, Fourneret P, Revol O, Nusbaum F, Sappey-Marinier D. White matter microarchitecture and structural network integrity correlate with children intelligence quotient. *Sci Rep*. 2020:10(1):20722.

Swagerman SC, de Geus EJC, Kan KJ, van Bergen E, Nieuwboer HA, Koenis MMG, Hulshoff Pol HE, Gur RE, Gur RC, Boomsma DI. The computerized neurocognitive battery: validation, aging effects, and heritability across cognitive domains. *Neuropsychology*. 2016:30(1):53–64.

Takahashi M, Iwamoto K, Fukatsu H, Naganawa S, Iidaka T, Ozaki N. White matter microstructure of the cingulum and cerebellar peduncle is related to sustained attention and working memory: a diffusion tensor imaging study. *Neurosci Lett*. 2010:477(2): 72–76.

Tamnes CK, Ostby Y, Walhovd KB, Westlye LT, Due-Tonnessen P, Fjell AM. Intellectual abilities and white matter microstructure in development: a diffusion tensor imaging study. *Hum Brain Mapp*. 2010:31(10):1609–1625.

Tang CY, Eaves EL, Ng JC, Carpenter DM, Mai X, Schroeder DH, Condon CA, Colom R, Haier RJ. Brain networks for working memory and factors of intelligence assessed in males and females with fMRI and DTI. *Intelligence*. 2010:38(3):293–303.

Urger SE, De Bellis MD, Hooper SR, Woolley DP, Chen SD, Provenzale J. The superior longitudinal fasciculus in typically developing children and adolescents: diffusion tensor imaging and neuropsychological correlates. *J Child Neurol*. 2015:30(1):9–20.

van der Knaap LJ, van der Ham IJ. How does the corpus callosum mediate interhemispheric transfer? *A review Behav Brain Res*. 2011:223(1):211–221.

Varley R. Plasticity in high-order cognition: evidence of dissociation in aphasia. *Behav Brain Sci*. 2007:30(2):171–172.

Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang H, Dubey P, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*. 2007:36(3):630–644.

Wechsler D. *Wechsler Adult Intelligence Scale - Fourth edition (WAIS-IV)*. San Antonio, Texas: Pearson Assessment; 2008.

Wechsler D. *Wechsler Abbreviated Intelligence Scale - Second edition (WASI-II)*. San Antonio, Texas: NCS Pearson; 2011.

Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, Carlozzi NE, Slotkin J, Blitz D, Wallner-Allen K, et al. Cognition assessment using the NIH toolbox. *Neurology*. 2013: 80(Suppl. 3):S54–S64.

Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. *BMJ*. 2001:322(7290):819.

Wilcox RR. *Introduction to robust estimation and hypothesis testing*. San Diego: Academic Press; 1997.

Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *NeuroImage*. 2014:92(100):381–397.

Wood JN, Grafman J. Human prefrontal cortex: processing and representational perspectives. *Nat Rev Neurosci*. 2003:4(2):139–147.

Xu J, Moeller S, Strupp J, Auerbach EJ, Chen L, Feinberg DA, Ugurbil K, Yacoub E. Highly accelerated whole brain imaging using aligned-blipped-controlled-aliasing multiband EPI. *Proc Int Soc Magn Reson Med*. 2012:20:2306.

Yu C, Li J, Liu Y, Qin W, Li Y, Shu N, Jiang T, Li K. White matter tract integrity and intelligence in patients with mental retardation and healthy adults. *NeuroImage*. 2008:40(4):1533–1541.

Zagorsky JL. Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence*. 2007:35(5): 489–501.

Zemmoura I, Herbet G, Moritz-Gasser S, Duffau H. New insights into the neural network mediating reading processes provided by cortico-subcortical electrical mapping. *Hum Brain Mapp*. 2015:36(6):2215–2230.