Robust associations between white matter microstructure and general intelligence

Robust associations between white matter microstructure and general intelligence

Christina Stammen[a], Christoph Fraenz[a], Rachael G. Grazioplene[b], Caroline Schlüter[c], Viola Merhof[d], Wendy Johnson[e], Onur Güntürkün[c], Colin G. DeYoung[f], Erhan Genç*[a]

[a]: Department of Psychology and Neuroscience, Leibniz Research Centre for Working Environment and Human Factors (IfADo), Dortmund, Germany

[b]: Department of Psychology, Yale University, New Haven, Connecticut, USA

[c]: Department of Biopsychology, Institute of Cognitive Neuroscience, Ruhr University Bochum, Bochum, Germany

[d]: Chair of Research Methods and Psychological Assessment, University of Mannheim, Mannheim, Germany

[e]: Department of Psychology, University of Edinburgh, Edinburgh, UK

[f]: Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA

*: Corresponding author: Dr. Erhan Genç

Telephone: +49 231 1984520

Email: genc@ifado.de

Address: Department of Psychology and Neuroscience, Leibniz Research Centre for Working Environment and Human Factors (IfADo), Ardeystraße 67, 44139 Dortmund, Germany.

Running title: White matter microstructure and general intelligence

Abstract

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 multi-center approach to identify white matter regions that show replicable structure-function associations, employing data from four independent samples comprising over 2000 healthy participants. TBSS indicated 188 voxels exhibited significant positive associations between *g* factor scores and fractional anisotropy in all four data sets. Replicable voxels formed three 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 fractional anisotropy 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.

Keywords:

DWI, general intelligence, multi-center study, TBSS, white matter

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 (Aspara et al. 2018; Batty et al. 2007; Calvin et al. 2017; Calvin et al. 2011; Deary et al. 2010b; Hemmingsson et al. 2006; Whalley and Deary 2001). 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 (Cox et al. 2019; McDaniel 2005; Pietschnig et al. 2015), 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 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. 2014; Barbey et al. 2012; 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 (Cox et al. 2019; Jung and Haier 2007). 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 2014; Le Bihan and Breton 1985; Le Bihan et al. 1986) 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 (Assaf and Pasternak 2008; Basser and Pierpaoli 1996). 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; Friedrich et al. 2020; Jones et al. 2013; Le Bihan 2003). 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; Power et al. 2019; Tang et al. 2010), extracting them from whole fiber tracts (Bathelt et al. 2019; Booth et al. 2013; Clayden et al. 2012; Cox et al. 2019; Cremers et al. 2016; Dubner et al. 2019; Ferrer et al. 2013; Fuhrmann et al. 2020; Góngora et al. 2020; Holleran et al. 2020; Kennedy et al. 2021; Kievit et al. 2016; Kievit et al. 2014; Kievit et al. 2018; Kontis et al. 2009; Muetzel et al. 2015; Nestor et al. 2015; Ohtani et al. 2014; Penke et al. 2012; Penke et al. 2010; Simpson-Kent et al. 2020; Suprano et al. 2020; Urger et al. 2015; Yu et al. 2008), or by a whole-brain voxel-based approach (Allin et al. 2011; Chiang et al. 2009; Navas-Sanchez et al. 2014; Schmithorst 2009; Schmithorst et al. 2005). As summarized by Genç and Fraenz (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 four 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 while others did not. As proposed by Genç and Fraenz (2021), such inconsistent observations may be tackled by following a multi-center approach. To this end, multiple, independent data sets, typically collected by different research groups, are analyzed in the same way. Importantly, only those results which replicate across the majority (or all) of samples are considered robust. We followed this approach as

158 methodologically consistently as possible, searching for replicable observations among four
159 independent data sets comprising cross-sectional data from more than 2000 healthy
160 participants. Our group performed whole-brain TBSS analyses to examine the associations
161 between general intelligence, in the form of *g* factor scores, and FA separately on each
162 sample. Besides the aforementioned advantage of multi-center studies, another reason for
163 choosing this rather conservative approach was that a first-level combination (pooling all) of
164 our four data sets with not-identical behavioral measures was not possible since sample mean
165 *g* levels might differ and because imaging data were obtained on different scanners. However,
166 as *g* and FA values were available for all samples and relative values between subjects within
167 samples should be comparable, we were able to combine the data sets at a second level with
168 our multi-center approach. Data were collected at Ruhr-University Bochum (RUB) in Germany
169 with N = 557 (Genç et al. 2021), the Human Connectome Project (HCP) with N = 1061 (van
170 Essen et al. 2013), the University of Minnesota (UMN) with N = 251 (Grazioplene et al. 2016;
171 Grazioplene et al. 2015), and the Nathan Kline Institute (NKI) with N = 396 (Nooner et al.
172 2012). We compared observations to identify white matter areas that exhibited replicable
173 structure-function associations among data sets. As the overlap among multiple data sets'
174 results will be likely to include fewer areas than a single data set's results, our study might
175 yield relatively circumscribed but robust associations between white matter and *g*. This could
176 give the impression that only focal differences in FA are associated with differences in general
177 intelligence. However, if some white matter fiber tracts seem more commonly implicated in
178 successfully accomplishing cognitive tasks this will not mean other brain white matter areas
179 are irrelevant. Involvement of white matter throughout the brain for information transfer seems
180 relevant for intellectual performance as intelligence is more likely to emerge from a dynamic
181 system comprising interacting subcomponents (Barbey 2018).

Materials and Methods

Participants

Data set RUB.

The RUB sample encompassed 557 participants (see 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 five 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 Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657), funded by the 16 United States 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, 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 (see 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 (see 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 University of Minnesota 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, see 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 (see 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 Nathan Kline Institute 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).

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 ± standard deviation.

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 (Flanagan and Dixon 2013; Schneider and McGrew 2012). 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); 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 four 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). Non-significance 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 .05 and values of CFI and TLI greater than .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 to 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 two 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 (Bentler and Bonett 1980; Jöreskog 1969). 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 two 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 non-hierarchical single-factor solution. It correlated at $r = .976$ with the one shown in Figure 1. Since there was no substantial difference, we decided to use the hierarchical factor model (see Figure 1) and include all available intelligence measures.

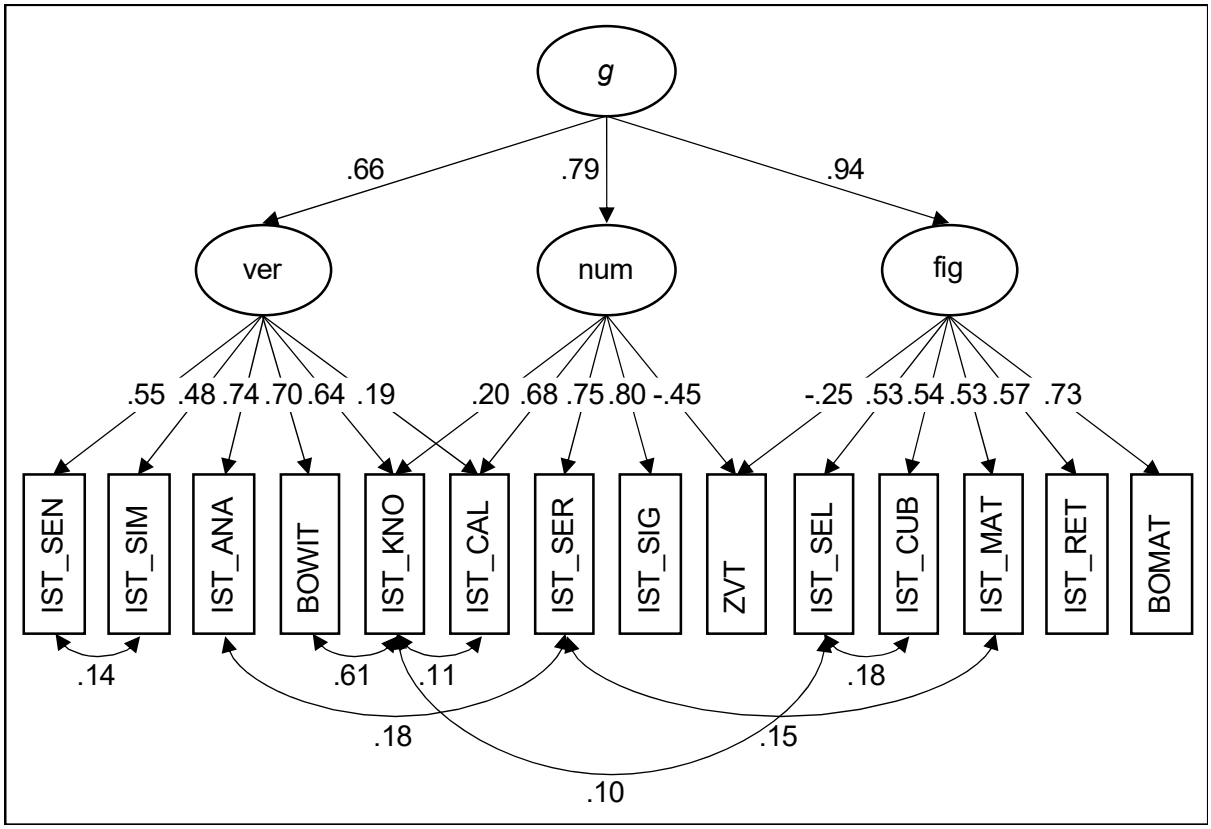


Figure 1. Confirmatory factor analytic model of the RUB data set. *g* = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, num = numerical intelligence as broad cognitive domain, fig = figural intelligence as broad cognitive domain, IST_SEN = subtest Sentence Completion 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 = Zahlen-Verbindungs-Test, 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.

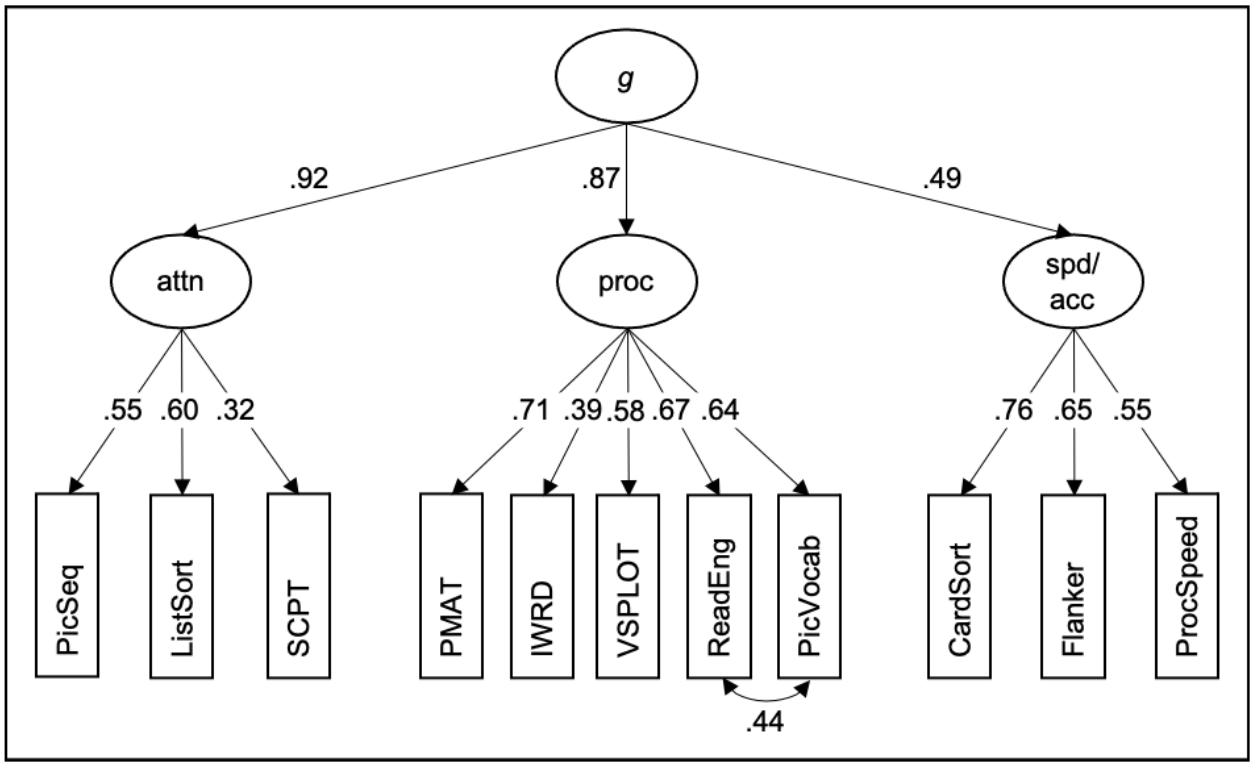


Figure 2. Confirmatory factor analytic model of the HCP data set. *g* = general factor of intelligence, attn = attention as broad cognitive domain, proc = processing as broad cognitive domain, spd/acc = speed/accuracy as broad cognitive domain, PicSeq = subtest Picture Sequence Memory Test of the NIH 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.

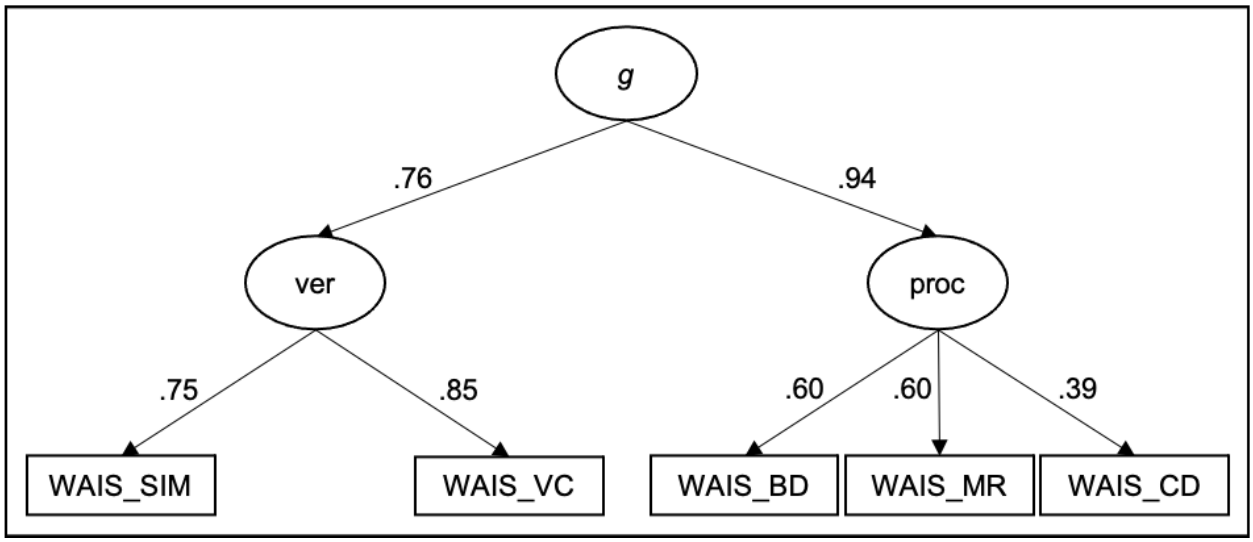


Figure 3. Confirmatory factor analytic model of the UMN data set. *g* = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, proc = processing as broad cognitive domain, WAIS_SIM = subtest Similarities of the WAIS-IV, WAIS_VC = subtest Vocabulary of the WAIS-IV, WAIS_BD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

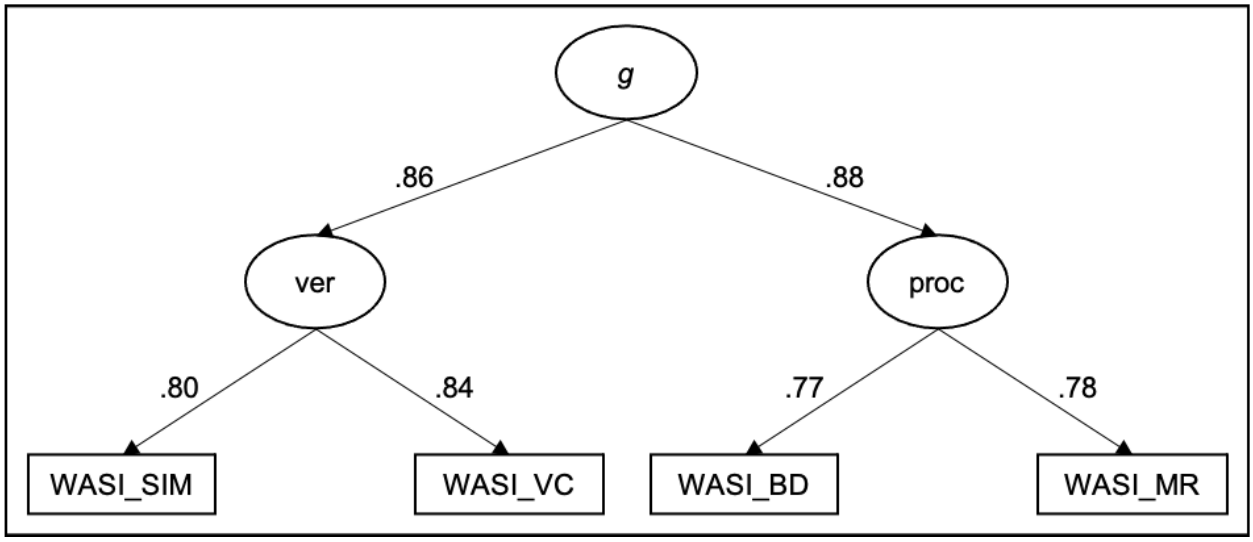


Figure 4. Confirmatory factor analytic model of the NKI data set. *g* = general factor of intelligence, ver = verbal intelligence as broad cognitive domain, proc = processing as broad cognitive domain, WASI_SIM = subtest Similarities of the WASI-II, WASI_VC = subtest Vocabulary of the WASI-II, WASI_BD = 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). *** $p < .001$.

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 (see Table 3). Most included cognitive tasks are presented in multiple-choice format. Reliability estimates (Cronbach's α) are between .88 and .96 for subtests and composite scores (Liepmann et al. 2007).

BOMAT-Advanced Short.

The Bochumer Matrizentest (BOMAT; Hossiep et al. 2001) is a non-verbal German intelligence test (see 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 (Fraenz et al. 2021; Genç et al. 2018; Genç et al. 2019; Hossiep et al. 2001; Jaeggi et al. 2008; Oelhafen et al. 2013). Split-half reliability of the BOMAT is .89 and Cronbach's α is .92 (Hossiep et al. 2001).

BOWIT.

The Bochumer Wissenstest (BOWIT; Hossiep and Schulte 2008) is a German 'general knowledge' questionnaire. It is available in two parallel test forms, in which each knowledge facet is represented by 14 multiple-choice questions (see Table 3). All participants completed both test forms, resulting in 308 items. In the BOWIT's manual split-half reliability is reported as .96, Cronbach's α .95, test-retest reliability .96, and parallel-form reliability .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 two short sample tasks and four assessed tasks (see Table 3). The reliability across the four tasks is reported as .95 in adults. The six-month retest-reliability is reported to be between .84 and .90 (Oswald and Roth 1987).

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
2. IST_ANA	Find analogies	20	intelligence
3. IST_SIM	Recognize similarities	20	
4. IST_ CAL	Solve arithmetic calculations	20	Numerical
5. IST_SER	Complete number series	20	intelligence

6. IST_SIG	Add arithmetic signs to mathematical equations	20	
7. IST_SEL	Select and reassemble parts of a cut-up figure	20	Figural
8. IST_CUB	Mentally rotate and match three-dimensional objects	20	intelligence
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-by-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.

370 **Data set HCP.**

371 ***Penn CNB.***

372 Four subtests from the University of Pennsylvania Computerized Neurocognitive Battery
373 (PennCNB; Gur et al. 2001; Gur et al. 2010; Moore et al. 2015) were used to assess
374 intelligence (see Table 4). These included the Penn Matrix Reasoning Task (PMAT), the Short
375 Penn Continuous Performance Test (SCPT), the Variable Short Penn Line Orientation Test
376 (VSPLOT), and the Penn Word Memory Test (IWRD). The reliability estimates (Cronbach's α)
377 for all subtests of the Penn CNB are reported to be between .55 and .98 (Gur et al. 2010).
378 Internal consistency was reported in a Dutch study to have a median Cronbach's α of .86
379 across all Penn CNB subtests (Swagerman et al. 2016).

380 ***NIH Toolbox.***

381 Seven subtests from the NIH Toolbox for the Assessment of Neurological and Behavioral
382 Function (http://www.nihtoolbox.org; Gershon et al. 2013; Heaton et al. 2014; Weintraub et al.
383 2013) were selected to assess intelligence (see Table 4). These were the Flanker Inhibitory
384 Control and Attention Test (Flanker), the Dimensional Change Card Sort Test (CardSort), the
385 List Sorting Working Memory Test (ListSort), the Picture Sequence Memory Test (PicSeq),
386 the Oral Reading Recognition Test (ReadEng), the Picture Vocabulary Test (PicVocab), and
387 the Pattern Comparison Processing Speed Test (ProcSpeed). The NIH Toolbox has been
388 validated with several American samples (Heaton et al. 2014; Weintraub et al. 2013). For the
389 subtests, Weintraub et al. (2013) reported test-retest reliabilities (intraclass correlation
390 coefficients) between r = .78 and .99. Heaton et al. (2014) built and analyzed composite scores
391 and found acceptable internal consistency (Cronbach's α between .77 and .84) as well as
392 excellent test-retest reliabilities between r = .86 and .92.

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

Intelligence test	Task description	No. of items	Construct measured

PennCNB			
1. PMAT	Solve matrix-reasoning problems (2-by-2, 3-by-3, or 1-by-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 two dimensions (e.g., shape and color) to one of two 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	Stop criterion: failure in two	Working memory capacity

	same category; second condition: stimuli belong to two categories and must be repeated in order of size as well as category-specific)	trials of the same length	
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 four 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

PennCNB = University of Pennsylvania Computerized Neurocognitive Battery. PMAT = subtest Penn Matrix Reasoning Task of the PennCNB. SCPT = subtest Short Penn Continuous Performance Test of the PennCNB. VSPLOT = subtest Variable Short Penn Line Orientation Test of the PennCNB. IWRD = subtest Penn Word Memory Test of the PennCNB. 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.

Data set UMN.

WAIS-IV.

Intelligence was assessed using five subtests (see 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 .84 and .94 and test-retest reliabilities to range between *r* = .69 and .91 (Wechsler 2008).

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 two-dimensional pattern with several three-dimensional building blocks	14	Perceptual reasoning
2. WAIS_MR	Solve matrix-reasoning problems	26	
3. WAIS_SIM	Describe the qualitative similarity between two words	18	Verbal comprehension
4. WAIS_VC	Define or describe words or concepts	30	
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.

Data set NKI.

The Wechsler Abbreviated Scale of Intelligence, second edition (WASI-II; Wechsler 2011), measured intelligence. The inventory has four 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 (see 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 = .87 and .91 in the child norming sample (6-16 years) and between r = .90 and .92 in the adult norming sample (17-90 years). Test-retest reliability was r = .79 in the child sample and .94 in the adult sample. The interrater reliabilities of the four subtests were between r = .94 and .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 datasets 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, three of our four samples leaned heavily towards 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 diffusion-weighted imaging data

Data set RUB.

All images were collected on a Philips 3T Achieva scanner at Bergmannsheil Hospital in Bochum, Germany, using a 32-channel head coil. Diffusion-weighted images were acquired using echo planar imaging (see Table 6). Diffusion weighting was uniformly distributed along 60 directions using a *b*-value of 1000 s/mm^2. Additionally, six 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 three 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 3T 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 (see Table 6; Feinberg et al. 2010; Moeller et al. 2010; Setsompop et al. 2012; Xu et al. 2012). The complete diffusion-weighted imaging session was divided into six runs, each lasting approximately nine minutes and 50 seconds (total acquisition time of about one hour). The six runs represented three 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 six 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 3T Siemens Trio scanner at the Center for Magnetic Resonance Research (CMRR) at the University of Minnesota in Minneapolis, using a 12-channel head coil. Diffusion-weighted images were acquired using echo planar imaging (see

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 Nathan Kline Institute in Orangeburg, New York. Diffusion-weighted images were acquired using echo planar imaging (see Table 6). Diffusion weighting was uniformly distributed along 128 directions using a *b*-value of 1500 s/mm^2. In addition, nine volumes without diffusion weighting (b = 0 s/mm^2) were obtained. The total acquisition time was five minutes, 58 seconds.

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 x 112	2 x 2 x 2
HCP	5520	89.5	78°	111	145 x 174	1.25 x 1.25 x 1.25
UMN	7900	86	90°	64	128 x 128	2 x 2 x 2
NKI	2400	85	90°	64	106 x 90	2 x 2 x 2

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 (BET; Smith 2002). Then, FA images were created by fitting tensor models to the raw diffusion data using FMRIB's Diffusion Toolbox (FDT). We transformed the resulting FA images into a common space via FMRIB's Nonlinear Image Registration Tool (FNIRT; 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 Human Connectome Project and is available in Montreal Neurologic Institute (MNI) 152 standard space (1 x 1 x 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 inter-subject 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 5,000 permutations for each analysis. Within the white matter skeleton of each data set, we used a general linear model (GLM) 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 $p < .05$. 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 four binarized maps depicting positive contrasts and the four binarized maps depicting negative contrasts (see Figure 5). This resulted in two 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 four 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 (Hua et al. 2008; Mori et al. 2005; Wakana et al. 2007). 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.

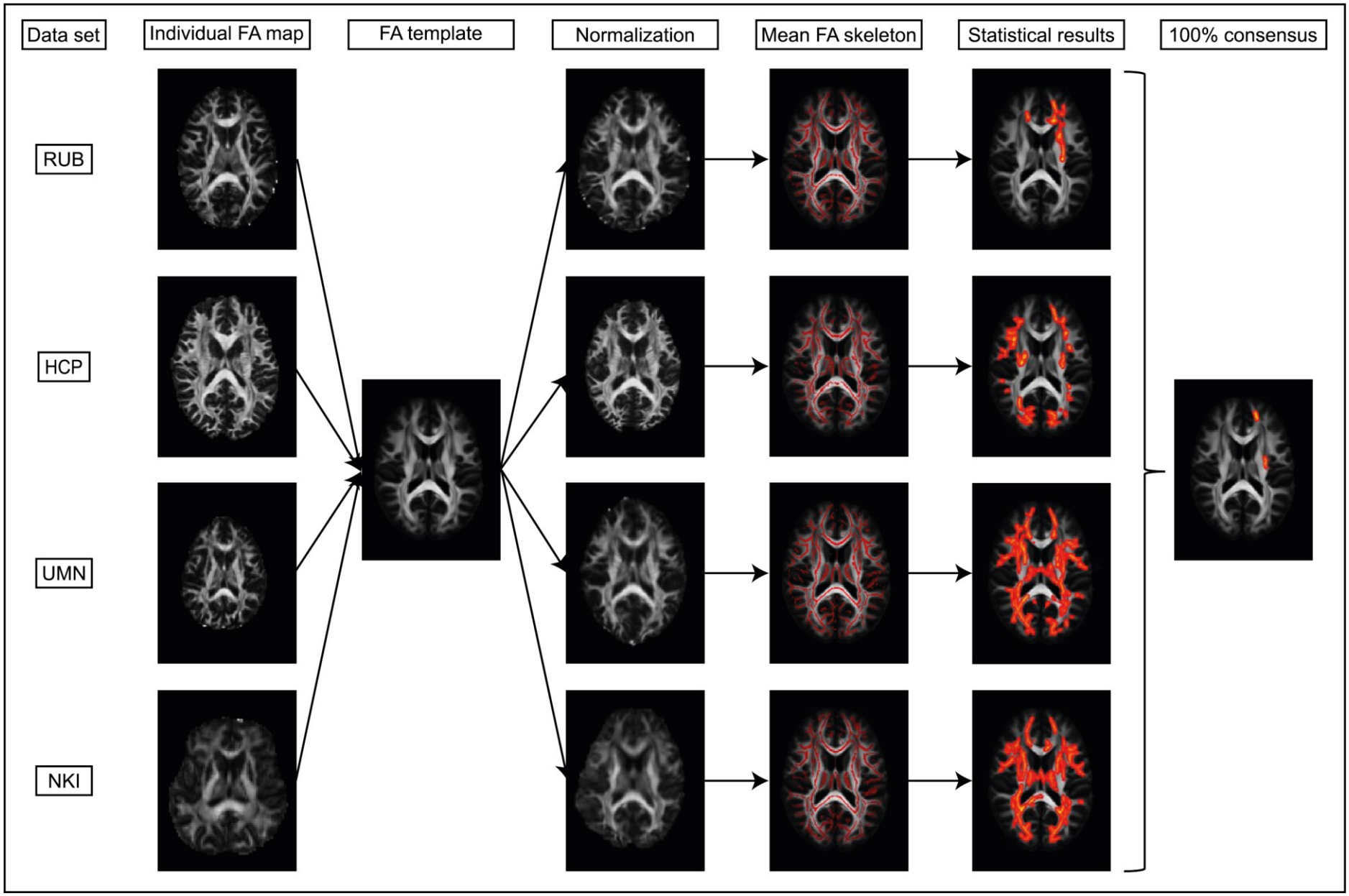

Data set
Individual FA map
FA template
Normalization
Mean FA skeleton
Statistical results
100% consensus
RUB
HCP
UMN
NKI

Figure 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 four data sets.

Additional exploratory analyses

We also took an exploratory and more liberal approach by creating brain maps including all voxels that exhibited significant associations in three out of four 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 (see Figures 1 to 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 four 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

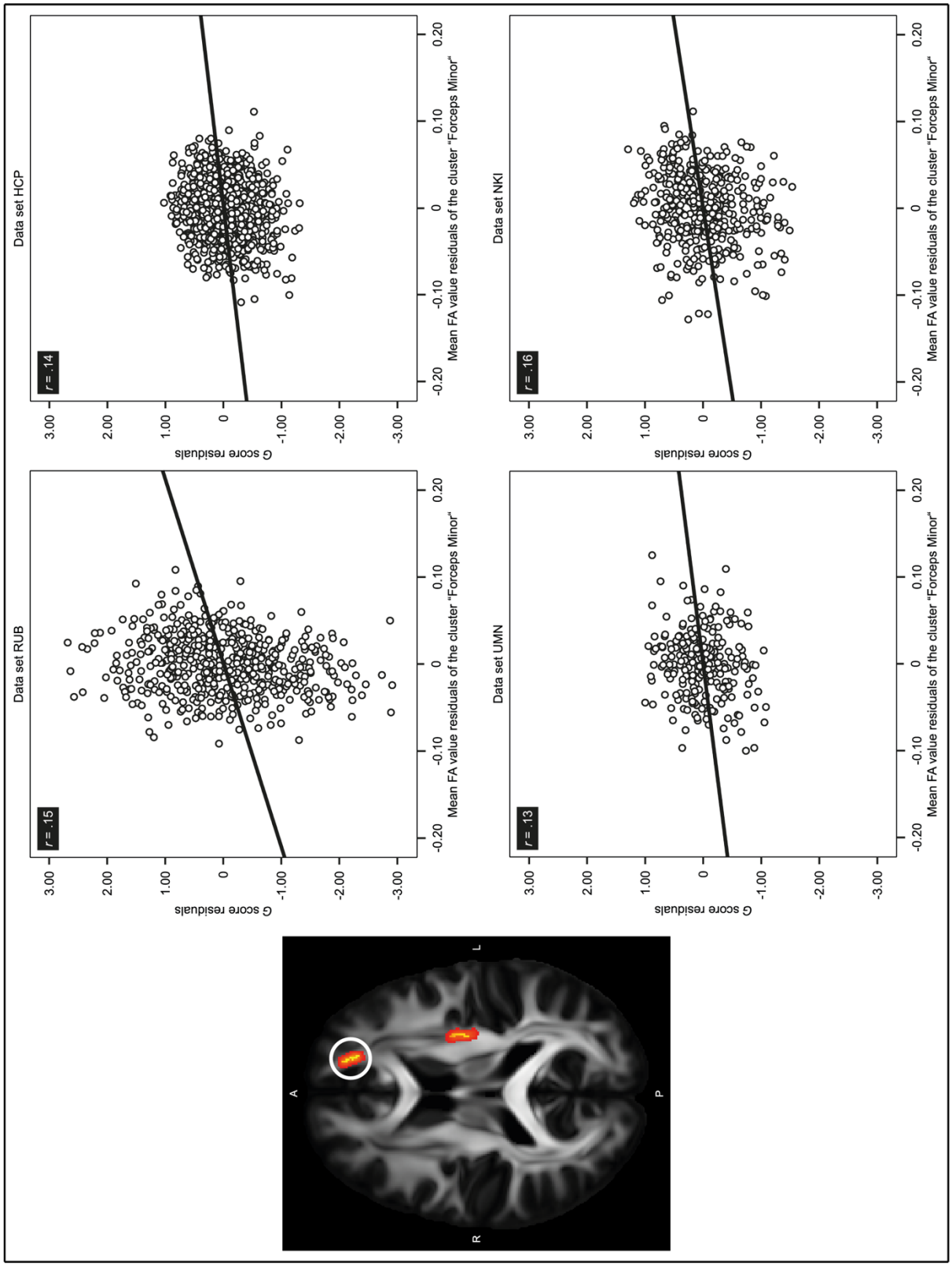
562 factor related to only verbal abilities (see Figure 2). Nonetheless, we still tried to include this
563 sample in our comparison of first-order intelligence factors. Hereby, we tested whether there
564 was a robust relation between FA and verbal abilities by combining the results of the first-
565 order intelligence factors ver (RUB, UMN, and NKI) and proc (HCP) (see Figures 1 to 4). For
566 processing abilities, we combined the first-order intelligence factors fig (RUB) and proc (HCP,
567 UMN, and NKI).
568

Results

Relations between *g* and FA

Main analysis with 100% consensus.

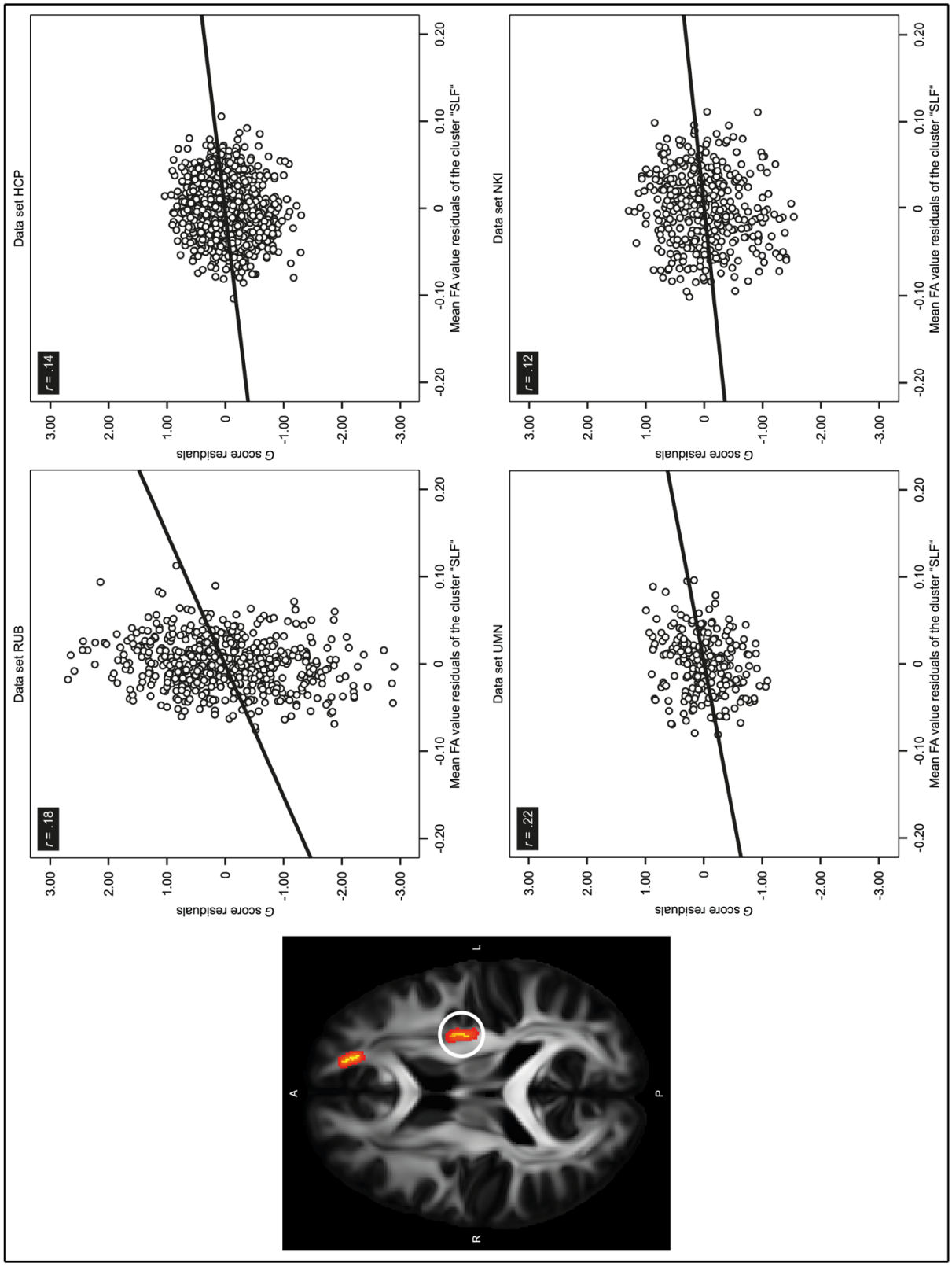
No voxels exhibited significant negative associations between *g* and FA in any of the four data sets. In total 188 individual voxels, 0.12% of the white matter skeleton, exhibited significant positive associations between *g* and FA in all four data sets, controlling age, sex, age*sex, age^2, and age^2*sex (for the results of the single data sets, see Supplemental Figure 1). These voxels could be pooled into three 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 Figure 6 (RUB: r = .15; HCP: r = .14; UMN: r = .13; NKI: r = .16). The second cluster "SLF" comprised 79 voxels and was located around the superior longitudinal fasciculus in the left hemisphere. Figure 7 shows the four scatter plots illustrating the associations between this cluster's mean FA and *g* (RUB: r = .18; HCP: r = .14; UMN: r = .22; NKI: r = .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 four scatter plots illustrating the associations between this cluster’s mean FA and *g* are shown in Figure 8 (RUB: r = .14; HCP: r = .12; UMN: r = .13; NKI: r = .13).



591

592 **Figure 6**. Associations between *g* and mean FA values from the cluster “Forceps minor”. The image on
593 the left side shows the voxel cluster named "Forceps minor" (encircled). The FA values of these voxels
594 were significantly positively associated with *g* in all four data sets (independent of effects of age, sex,
595 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 four 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.



602

603 **Figure 7**. Associations between *g* and mean FA values from the cluster "SLF". The image on the left
604 side shows the voxel cluster named "SLF" (encircled). The FA values of these voxels were significantly
605 positively associated with *g* in all four data sets (independent of the effects of age, sex, age*sex, age^2,
606 and age^2*sex). The voxels were located around the superior longitudinal fasciculus in the left

hemisphere. The right side of the figure shows four 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.

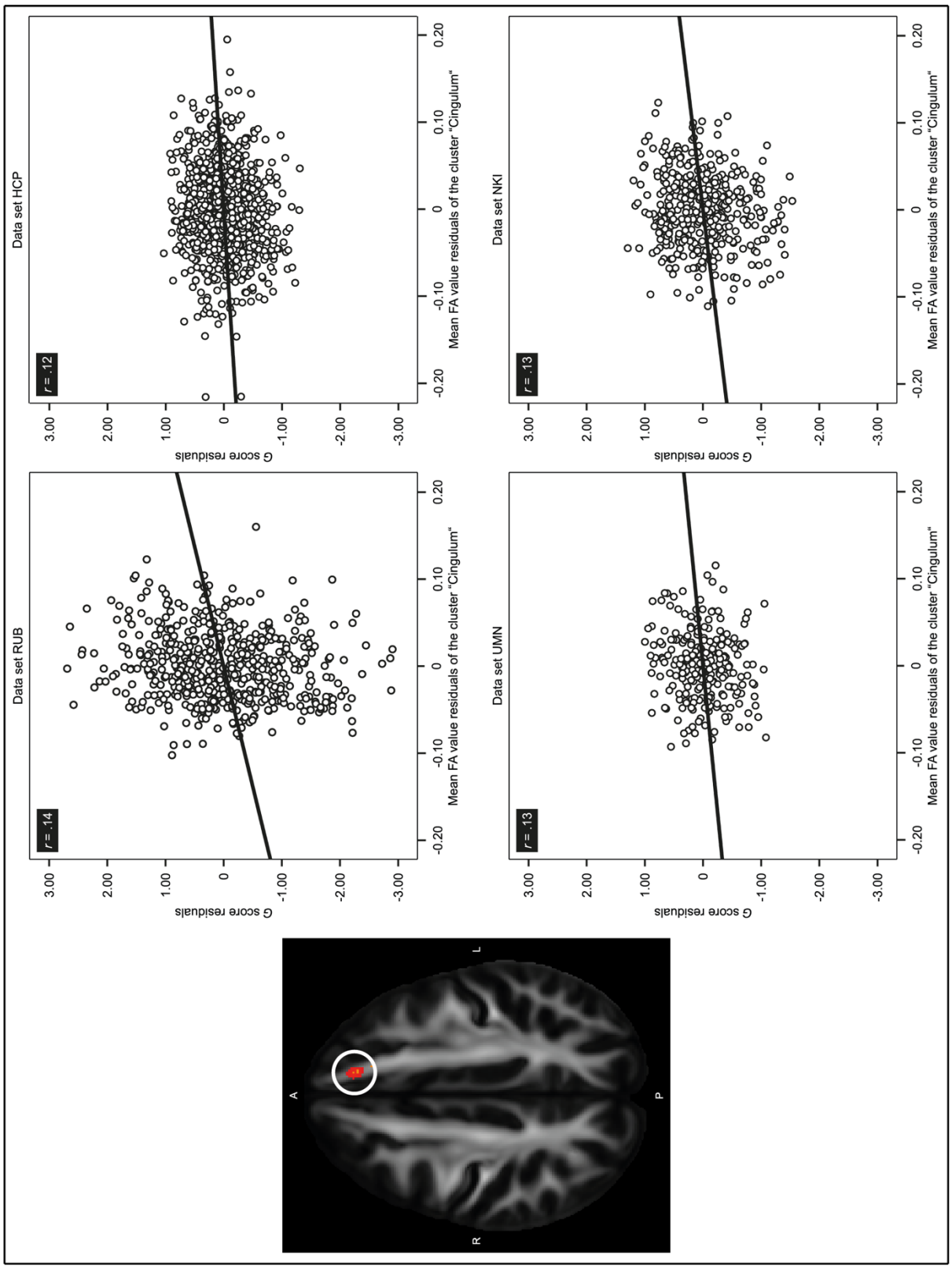


Figure 8. Associations between *g* and mean FA values from the cluster "Cingulum". The image on the left side shows the voxel cluster named "Cingulum" (encircled). The FA values of these voxels were significantly positively associated with *g* in all four data sets (independent of the effects of age, 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 four 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.

Exploratory approach with 75% consensus.

The more liberal approach, requiring results to replicate in three of the four 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 Supplemental Figure 2, these voxels were widely scattered across the skeleton. Table S1 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 four 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

639 exhibited significant positive relations between *g* and FA across all four data sets, controlling
640 age, sex, age*sex, age^2, and age^2*sex. These voxels formed three contiguous clusters. The
641 first was located around the forceps minor, crossing with extensions of the anterior thalamic
642 radiation, the cingulum-cingulate gyrus, and the inferior fronto-occipital fasciculus in the left
643 hemisphere. The second was located around the left-hemispheric superior longitudinal
644 fasciculus. The third was located around the left-hemispheric cingulum-cingulate gyrus,
645 crossing with extensions of the anterior thalamic radiation and the inferior fronto-occipital
646 fasciculus.
647 There were no voxels exhibiting significant negative associations between *g* and FA in any of
648 the four data sets. This was consistent with previous research. Multiple studies have examined
649 the associations between various measures of intelligence and FA using various approaches
650 including ROI-based, tract-based, whole-brain-based, and TBSS-based analyses. Despite
651 these differences in design, these studies almost exclusively reported positive associations
652 (Genç and Fraenz 2021). This suggests that individuals with higher intelligence scores tend
653 to have white matter with stronger anisotropic diffusion patterns. However, as FA is a metric
654 aggregating many tissue properties (Beaulieu 2002; Friedrich et al. 2020; Jones et al. 2013;
655 Le Bihan 2003), the exact neurobiological underpinnings driving FA signal differences remain
656 unclear. We can thus only speculate about how higher FA values link to higher *g*. Causal
657 implications could not be drawn from our analyses. Previous studies examining healthy older
658 people suggested that information processing efficiency might mediate the association
659 between FA values and *g* (Deary et al. 2006; Penke et al. 2010). Whether this finding extends
660 to other age groups remains to be seen, but it provides first indications that higher *g* values
661 might emerge from faster, more direct, or more parallel information processing. As
662 summarized by Friedrich et al. (2020), myelination and fiber density have been considered
663 two likely neurobiological contributors to FA. Higher FA values might create links with higher
664 mental speed via greater underlying myelination enabling faster conduction velocity (Nave
665 2010). More direct information transfer throughout the brain might rely on higher FA values
666 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 fiber tracts have been reported in previous TBSS-studies (Dunst et al. 2014; Malpas et al. 2016; Tamnes et al. 2010).

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 (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) (Karolis et al. 2019; Kolb and Whishaw 2015) and the two 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 two 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 (Basten et al. 2015; Deary et al. 2010a; Jung and Haier 2007). 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 (Aggleton et al. 2010; Behrens et al. 2003). 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 (Dehghani and Wimmer 2019; Rikhye et al. 2018).

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 towards the temporal lobe (Bubb et al. 2018; Catani and Thiebaut de Schotten 2008). Its fibers form intracortical connections between the 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 (Buyanova and Arsalidou 2021; Catani and Thiebaut de Schotten 2008). The cingulum appears to be involved in various cognitive domains such as cognitive control, attention, executive functions, memory, language, and visual-spatial functions (Bettcher et al. 2016; Bubb et al. 2018; Buyanova and Arsalidou 2021; Kantarci et al. 2011; Takahashi et al. 2010).

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 (Buyanova and Arsalidou 2021; Catani and Thiebaut de Schotten 2008; Leng et al. 2016).

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 (Cox et al. 2019; Dick and Tremblay 2012; Kamali et al. 2014). 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 four 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

751 constituting their ‘P-FIT’ network. More precisely, they suggested that brain regions in the
752 temporal and occipital lobes are crucial in successfully recognizing and initially processing
753 sensory information. Subsequently, they presumed that the parietal cortex is essential for the
754 interpretation, abstraction, and elaboration of the information’s symbolic content. The parietal
755 cortex is believed to interact with frontal regions, which are thought to orchestrate generation
756 and testing of potential solutions to given problems. Once a solution has been selected, it is
757 thought that the anterior cingulate cortex chooses an appropriate reaction and inhibits
758 alternative responses.
759 Based on this, Jung & Haier (2007) proposed that the rapid and error-free transfer of
760 information from posterior to frontal brain areas depends on underlying white matter integrity.
761 They also emphasized the importance of information exchange between parietal and frontal
762 association areas, which would highlight a role for the superior longitudinal fasciculus (Jung
763 and Haier 2007). Therefore, our observations relating the superior longitudinal fasciculus to
764 general intelligence supported the P-FIT model. Our cingulum observations fit within the P-
765 FIT network. As noted by Fraenz et al. (2021), the P-FIT network is not organized exclusively
766 intra-hemispherically. Hence, interhemispheric information transfer between prefrontal areas,
767 e.g. via the forceps minor, seems to be consistent as well. The P-FIT model does not propose
768 direct connections between occipital and (orbito-)frontal areas. However, our observations,
769 highlighting the importance of the inferior fronto-occipital fasciculus, did not necessarily
770 contradict the model, given that this fiber tract also connects distal cortical regions of the P-
771 FIT network. Instead, additional connections offer the possibility of more parallel flows of
772 information. Since individuals who score identically in an intelligent test may use different
773 cognitive strategies as well as different brain structures to reach their performance level
774 (Deary et al. 2010a), there may be more than one adequate solution path and overall good
775 brain function may be more important for general intelligence than using any specific parts
776 well.
777 Jung and Haier (2007) assumed that brain regions beyond the cerebral cortex, such as
778 thalamus, hippocampus, and cerebellum, are involved only in rather basic functions. Hence,

779 they believed that they would not contribute to interindividual intelligence differences
780 significantly. However, more recent studies indicate that the thalamus and the hippocampus
781 as well as their connections could play more important roles in reasoning than originally
782 thought (Bohlken et al. 2014; Cox et al. 2019; Deary et al. 2022; Dehghani and Wimmer 2019;
783 Rikhye et al. 2018). Our observations, involving the anterior thalamic radiation, supported
784 these studies in suggesting that the P-FIT model (Jung and Haier 2007) needs some updating,
785 which is only to be expected after 15 years more research.
786 We initially took a rather conservative analytical approach. To be considered for discussion,
787 voxels had to exhibit significant associations between *g* and FA across all four data sets (100%
788 consensus). A more liberal threshold (75% consensus) yielded about 44 times more voxels.
789 This was simply because more datasets inevitably vary in more ways. Moreover, as illustrated
790 in Supplemental Figure 2, significant voxel clusters were no longer exclusively located in the
791 left hemisphere. However, Table S1 indicates that more significant voxels could be assigned
792 to fiber tracts in the left hemisphere (59.3%, out of nine fiber tracts with left-right symmetry
793 seven had more voxels in the left hemisphere). As the left and the right hemisphere differ in
794 their specialized functions (Karolis et al. 2019; Kolb and Whishaw 2015; Marinsek et al. 2014),
795 both hemispheres and their functional interaction are relevant for intelligent behavior.
796 The additional exploratory analyses of different first-order intelligence factors did not lead to
797 any overlapping results in even two data sets. Our observations were not consistent with
798 Tamnes et al. (2010), who reported significant positive associations between FA and
799 verbal/nonverbal reasoning abilities. This could be because the first-order intelligence factors
800 differed among samples (see Figures 1-4). As they include much less information than *g*,
801 differences in the specific tasks might have impacted these factors' contents more than they
802 did *g*. But our results could also differ from Tamnes et al.'s because our analyses of these
803 narrower intelligence factors controlled *g* itself, which theirs did not, so we examined only
804 factor-specific variance. *g* explains about 40% of total variance in typical test batteries (Deary
805 et al. 2010a), in our cases 32-65% (see Table 2). To resolve such inconsistencies, future
806 studies should also focus on specific intelligence factors, though 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 (see Supplemental Figure S1). 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 four 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 Supplemental Figure 1, the RUB data set was the most common exception to 100% overlap. This data set differed from the other, more similar three in several aspects: the sample had been collected in Germany and therefore influenced by German pedagogies (vs. USA), 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 (see Figures 6-8). As two other (HCP and UMN) of our four samples leaned heavily towards 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, two 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 non-verbal 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 diffusion-weighted imaging 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 four 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 four 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.

Funding

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

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

References

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

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

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

Andersson JLR, Jenkinson M, Smith SM. 2007a. Non-linear optimisation. FMRIB technical report TR07JA1.

Andersson JLR, Jenkinson M, Smith SM. 2007b. Non-linear registration aka Spatial normalisation. FMRIB technical report TR07JA2.

Ashburner J, Friston KJ. 2000. Voxel-based morphometry — The methods. NeuroImage. 11:805-821.

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

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

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

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

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

Basser PJ, Pierpaoli C. 1996. Microstructural and physiological features of tissues elucidated by Quantitative-Diffusion-Tensor MRI. J Magn Reson. Series B 111(3):209-219.

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

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

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

Beaulieu C. 2002. The basis of anisotropic water diffusion in the nervous system - a technical review. NMR Biomed. 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. . 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat Neurosci. 6(7):750-757.

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

Bettcher BM, Mungas D, Patel N, Elofson J, Dutt S, Wynn M, Watson CL, Stephens M, Walsh CM, Kramer JH. 2016. Neuroanatomical substrates of executive functions: Beyond prefrontal structures. Neuropsychologia. 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. 2014. Genes contributing to subcortical volumes and intellectual ability implicate the thalamus. Hum Brain Mapp. 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. . 2013. Brain white matter tract integrity and cognitive abilities in community-dwelling older people: the Lothian Birth Cohort, 1936. Neuropsychology. 27(5):595-607.

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

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

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

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

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

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

Catani M, Thiebaut de Schotten M. 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex. 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. . 2009. Genetics of brain fiber architecture and intellectual performance. J Neurosci. 29(7):2212-2224.

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

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

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

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

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

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

Deary IJ, Weiss A, Batty GD. 2010b. 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. 11(2):53-79.

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

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

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

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

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

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

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

978 Filley C. 2012. The behavioral neurology of white matter. New York: Oxford University Press.
979 Flanagan DP, Dixon SG. 2013. The Cattell-Horn-Carroll theory of cognitive abilities. In: Reynolds
980 CR, Vannest KJ, Fletcher-Janzen E, editors. Encyclopedia of Special Education.
981 Hoboken, New Jersey: John Wiley & Sons, Inc.
982 Fraenz C, Schlüter C, Friedrich P, Jung RE, Güntürkün O, Genç E. 2021. Interindividual
983 differences in matrix reasoning are linked to functional connectivity between brain regions
984 nominated by Parieto-Frontal Integration Theory. Intelligence. 87:101545.
985 Friederici AD. 2009. Pathways to language: fiber tracts in the human brain. Trends Cogn Sci.
986 13(4):175-181.
987 Friedrich P, Fraenz C, Schlüter C, Ocklenburg S, Madler B, Güntürkün O, Genç E. 2020. The
988 relationship between axon density, myelination, and fractional anisotropy in the human
989 corpus callosum. Cereb Cortex. 30(4):2042-2056.
990 Frye RE, Hasan K, Malmberg B, Desouza L, Swank P, Smith K, Landry S. 2010. Superior
991 longitudinal fasciculus and cognitive dysfunction in adolescents born preterm and at term.
992 Dev Med Child Neurol. 52(8):760-766.
993 Fuhrmann D, Simpson-Kent IL, Bathelt J, Team C, Kievit RA. 2020. A hierarchical watershed
994 model of fluid intelligence in childhood and adolescence. Cereb Cortex. 30(1):339-352.
995 Genç E, Bergmann J, Singer W, Kohler A. 2011a. Interhemispheric connections shape subjective
996 experience of bistable motion. Curr Biol. 21(17):1494-1499.
997 Genç E, Bergmann J, Tong F, Blake R, Singer W, Kohler A. 2011b. Callosal connections of
998 primary visual cortex predict the spatial spreading of binocular rivalry across the visual
999 hemifields. Front Hum Neurosci. 5:161.
1000 Genç E, Fraenz C. 2021. Diffusion-weighted imaging of intelligence. In: Barbey AK, Karama S,
1001 Haier RJ, editors. The Cambridge Handbook of Intelligence and Cognitive Neuroscience.
1002 1. ed. New York: Cambridge University Press. p. 191-209.
1003 Genç E, Fraenz C, Schlüter C, Friedrich P, Hossiep R, Voelkle MC, Ling JM, Güntürkün O, Jung
1004 RE. 2018. Diffusion markers of dendritic density and arborization in gray matter predict
1005 differences in intelligence. Nat Commun. 9(1):1905.

Genç E, Fraenz C, Schlüter C, Friedrich P, Voelkle MC, Hossiep R, Güntürkün O. 2019. The neural architecture of general knowledge. EJP. 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. . 2021. Polygenic scores for cognitive abilities and their association with different aspects of general intelligence—A deep phenotyping approach. Mol Neurobiol. 58:4145-4156.

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

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

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

Gottfredson LS. 1997. Why g matters: The complexity of everyday life. Intelligence. 24(1):79-132.

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

Grazioplene RG, S GR, Gray JR, Rustichini A, Jung RE, DeYoung CG. 2015. Subcortical intelligence: caudate volume predicts IQ in healthy adults. Hum Brain Mapp. 36(4):1407-1416.

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

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

1034 Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, Beaumont J, Casaletto
1035 KB, Conway K, Slotkin J et al. . 2014. Reliability and validity of composite scores from the
1036 NIH Toolbox Cognition Battery in adults. J Int Neuropsychol Soc. 20(6):588-598.
1037 Hemmingsson T, Melin B, Allebeck P, Lundberg I. 2006. The association between cognitive ability
1038 measured at ages 18-20 and mortality during 30 years of follow-up--a prospective
1039 observational study among Swedish males born 1949-51. Int J Epidemiol. 35(3):665-670.
1040 Hidese S, Ota M, Matsuo J, Ishida I, Hiraishi M, Yokota Y, Hattori K, Yomogida Y, Kunugi H.
1041 2020. Correlation between the Wechsler adult intelligence scale- 3rd edition metrics and
1042 brain structure in healthy individuals: A whole-brain magnetic resonance imaging study.
1043 Front Hum Neurosci. 14:211.
1044 Hoeft F, Barnea-Goraly N, Haas BW, Golarai G, Ng D, Mills D, Korenberg J, Bellugi U, Galaburda
1045 A, Reiss AL. 2007. More is not always better: increased fractional anisotropy of superior
1046 longitudinal fasciculus associated with poor visuospatial abilities in Williams syndrome. J
1047 Neurosci. 27(44):11960-11965.
1048 Holleran L, Kelly S, Alloza C, Agartz I, Andreassen OA, Arango C, Banaj N, Calhoun V, Cannon
1049 D, Carr V et al. . 2020. The relationship between white matter microstructure and general
1050 cognitive ability in patients with schizophrenia and healthy participants in the ENIGMA
1051 consortium. Am J Psychiatry. 177(6):537-547.
1052 Hossiep R, Hasella M, Turck D. 2001. BOMAT-advanced-short version: Bochumer Matrizentest.
1053 Göttingen (Germany): Hogrefe.
1054 Hossiep R, Schulte M. 2008. BOWIT: Bochumer Wissenstest. Göttingen (Germany): Hogrefe.
1055 Hu Lt, Bentler PM. 1999. Cutoff criteria for fit indexes in covariance structure analysis:
1056 Conventional criteria versus new alternatives. Struct Equ Model. 6(1):1-55.
1057 Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, Calabresi PA, Pekar JJ, van Zijl PC, Mori
1058 S. 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy
1059 and tract-specific quantification. NeuroImage. 39(1):336-347.
1060 Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training
1061 on working memory. Proc Natl Acad Sci U S A. 105(19):6829-33.

1062 Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II. 2004. Just one g: consistent
1063 results from three test batteries. Intelligence. 32(1):95-107.
1064 Johnson W, Nijenhuis Jt, Bouchard TJ. 2008. Still just 1 g: Consistent results from five test
1065 batteries. Intelligence. 36(1):81-95.
1066 Jones DK, Knösche TR, Turner R. 2013. White matter integrity, fiber count, and other fallacies:
1067 the do's and don'ts of diffusion MRI. NeuroImage. 73:239-254.
1068 Jöreskog KG. 1969. A general approach to confirmatory maximum likelihood factor analysis.
1069 Psychometrika. 34(2):183-202.
1070 Jung RE, Haier RJ. 2007. The Parieto-Frontal Integration Theory (P-FIT) of intelligence:
1071 Converging neuroimaging evidence. Behav Brain Sci. 30(2):135-154.
1072 Kamali A, Flanders AE, Brody J, Hunter JV, Hasan KM. 2014. Tracing superior longitudinal
1073 fasciculus connectivity in the human brain using high resolution diffusion tensor
1074 tractography. Brain Struct Funct. 219(1):269-281.
1075 Kantarci K, Senjem ML, Avula R, Zhang B, Samikoglu AR, Weigand SD, Przybelski SA,
1076 Edmonson HA, Vemuri P, Knopman DS et al. . 2011. Diffusion tensor imaging and
1077 cognitive function in older adults with no dementia. Neurology. 77:26-34.
1078 Karolis VR, Corbetta M, Thiebaut de Schotten M. 2019. The architecture of functional
1079 lateralisation and its relationship to callosal connectivity in the human brain. Nat Commun.
1080 10(1):1417.
1081 Kennedy E, Poppe T, Tottman A, Harding J. 2021. Neurodevelopmental impairment is associated
1082 with altered white matter development in a cohort of school-aged children born very
1083 preterm. NeuroImage Clin. 31:102730.
1084 Kievit RA, Davis SW, Griffiths J, Correia MM, Cam C, Henson RN. 2016. A watershed model of
1085 individual differences in fluid intelligence. Neuropsychologia. 91:186-198.
1086 Kievit RA, Davis SW, Mitchell DJ, Taylor JR, Duncan J, Cam CANRT, Henson RN, Cam CANRT.
1087 2014. Distinct aspects of frontal lobe structure mediate age-related differences in fluid
1088 intelligence and multitasking. Nat Commun. 5:5658.

1089 Kievit RA, Fuhrmann D, Borgeest GS, Simpson-Kent IL, Henson RNA. 2018. The neural
1090 determinants of age-related changes in fluid intelligence: a pre-registered, longitudinal
1091 analysis in UK Biobank. Wellcome Open Res.3-38.
1092 Kochunov P, Jahanshad N, Marcus D, Winkler A, Sprooten E, Nichols TE, Wright SN, Hong LE,
1093 Patel B, Behrens T et al. . 2015. Heritability of fractional anisotropy in human white matter:
1094 A comparison of Human Connectome Project and ENIGMA-DTI data. NeuroImage.
1095 111:300-311.
1096 Kolb B, Whishaw IQ. 2015. Fundamentals of human neuropsychology. New York: Worth
1097 Publishers.
1098 Kontis D, Catani M, Cuddy M, Walshe M, Nosarti C, Jones D, Wyatt J, Rifkin L, Murray R, Allin
1099 M. 2009. Diffusion tensor MRI of the corpus callosum and cognitive function in adults born
1100 preterm. NeuroReport. 20(4):424-428.
1101 Le Bihan D. 2003. Looking into the functional architecture of the brain with diffusion MRI. Nat Rev
1102 Neurosci. 4(6):469-480.
1103 Le Bihan D. 2014. Diffusion MRI: what water tells us about the brain. EMBO Mol Med. 6(5):569-
1104 573.
1105 Le Bihan D, Breton É. 1985. In vivo magnetic resonance imaging of diffusion. C R Acad Sci Ser
1106 II. 301(15):1109-1112.
1107 Le Bihan D, Breton É, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. 1986. MR imaging
1108 of intravoxel incoherent motions: Application of diffusion and perfusion in neurologic
1109 disorders. Radiology. 161(2):401-407.
1110 Lebel C, Beaulieu C. 2009. Lateralization of the arcuate fasciculus from childhood to adulthood
1111 and its relation to cognitive abilities in children. Hum Brain Mapp. 30(11):3563-3573.
1112 Leng Y, Shi Y, Yu Q, Van Horn JD, Tang H, Li J, Xu W, Ge X, Tang Y, Han Y et al. . 2016.
1113 Phenotypic and genetic correlations between the lobar segments of the inferior fronto-
1114 occipital fasciculus and attention. Sci Rep. 6:33015.
1115 Liepmann D, Beauducel A, Brocke B, Amthauer R. 2007. Intelligenz-Struktur-Test 2000 R (I-S-T
1116 2000 R). Manual. Göttingen (Germany): Hogrefe.

1117 Malpas CB, Genc S, Saling MM, Velakoulis D, Desmond PM, O'Brien TJ. 2016. MRI correlates
1118 of general intelligence in neurotypical adults. J Clin Neurosci. 24:128-134.
1119 Marinsek N, Turner BO, Gazzaniga M, Miller MB. 2014. Divergent hemispheric reasoning
1120 strategies: reducing uncertainty versus resolving inconsistency. Front Hum Neurosci.
1121 8:839.
1122 McCrimmon AW, Smith AD. 2012. Review of the Wechsler Abbreviated Scale of Intelligence,
1123 Second Edition (WASI-II). J Psychoeduc Assess. 31(3):337-341.
1124 McDaniel M. 2005. Big-brained people are smarter: A meta-analysis of the relationship between
1125 in vivo brain volume and intelligence. Intelligence. 33(4):337-346.
1126 McGue M, Bouchard TJ. 1984. Adjustment of twin data for the effects of age and sex. Behav.
1127 Genet. 14(4):325-343.
1128 Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, Ugurbil K. 2010. Multiband
1129 multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with
1130 application to high spatial and temporal whole-brain fMRI. Magn Reson Med. 63(5):1144-
1131 1153.
1132 Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. 2015. Psychometric properties of the
1133 Penn Computerized Neurocognitive Battery. Neuropsychology. 29(2):235-246.
1134 Mori S, Kaufmann WE, Davatzikos C, Stieltjes B, Amodei L, Fredericksen K, Pearlson GD,
1135 Melhem ER, Solaiyappan M, Raymond GV et al. . 2002. Imaging cortical association tracts
1136 in the human brain using diffusion-tensor-based axonal tracking. Magn Res Med.
1137 47(2):215-223.
1138 Mori S, Wakana S, van Zijl PCM, Nagae-Poetscher LM. 2005. MRI atlas of human white matter.
1139 Elsevier B. V.
1140 Muetzel RL, Mous SE, van der Ende J, Blanken LM, van der Lugt A, Jaddoe VW, Verhulst FC,
1141 Tiemeier H, White T. 2015. White matter integrity and cognitive performance in school-
1142 age children: A population-based neuroimaging study. NeuroImage. 119:119-128.

1143 Navas-Sanchez FJ, Aleman-Gomez Y, Sanchez-Gonzalez J, Guzman-De-Villoria JA, Franco C,
1144 Robles O, Arango C, Desco M. 2014. White matter microstructure correlates of
1145 mathematical giftedness and intelligence quotient. Hum Brain Mapp. 35(6):2619-2631.
1146 Nave KA. 2010. Myelination and support of axonal integrity by glia. Nature. 468(7321):244-52.
1147 Neisser U, Boodoo G, Bouchard TJ, Boykin AW, Brody N, Ceci SJ, Halpern DF, Loehlin JC,
1148 Perloff R, Sternberg RJ et al. . 1996. Intelligence: Knowns and unknowns. Am Psychol.
1149 51(2):77-101.
1150 Nestor PG, Ohtani T, Bouix S, Hosokawa T, Saito Y, Newell DT, Kubicki M. 2015. Dissociating
1151 prefrontal circuitry in intelligence and memory: neuropsychological correlates of magnetic
1152 resonance and diffusion tensor imaging. Brain Imaging Behav. 9(4):839-847.
1153 Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: A
1154 primer with examples. Hum Brain Mapp. 15:1-25.
1155 Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, Panek LJ, Brown S,
1156 Zavitz ST, Li Q et al. . 2012. The NKI-Rockland sample: A model for accelerating the pace
1157 of discovery science in psychiatry. Front Neurosci. 6:152.
1158 Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. 2013. Increased parietal
1159 activity after training of interference control. Neuropsychologia. 51(13):2781-90.
1160 Ohtani T, Nestor PG, Bouix S, Saito Y, Hosokawa T, Kubicki M. 2014. Medial frontal white and
1161 gray matter contributions to general intelligence. PLoS One. 9(12):e112691.
1162 Oswald WD, Roth E. 1987. Der Zahlen-Verbindungs-Test (ZVT). Göttingen (Germany): Hogrefe
1163 Verlag für Psychologie.
1164 Papez JW. 1937. A proposed mechanism of emotion. Arch Neurol Psychiatry. 38:725-743.
1165 Penke L, Maniega SM, Bastin ME, Valdes Hernandez MC, Murray C, Royle NA, Starr JM,
1166 Wardlaw JM, Deary IJ. 2012. Brain white matter tract integrity as a neural foundation for
1167 general intelligence. Mol Psychiatry. 17(10):1026-1030.
1168 Penke L, Munoz Maniega S, Murray C, Gow AJ, Hernandez MC, Clayden JD, Starr JM, Wardlaw
1169 JM, Bastin ME, Deary IJ. 2010. A general factor of brain white matter integrity predicts
1170 information processing speed in healthy older people. J Neurosci. 30(22):7569-7574.

1171 Pietschnig J, Penke L, Wicherts JM, Zeiler M, Voracek M. 2015. Meta-analysis of associations
1172 between human brain volume and intelligence differences: How strong are they and what
1173 do they mean? Neurosci Biobehav Rev. 57:411-432.
1174 Power MC, Su D, Wu A, Reid RI, Jack CR, Knopman DS, Coresh J, Huang J, Kantarci K, Sharrett
1175 AR et al. . 2019. Association of white matter microstructural integrity with cognition and
1176 dementia. Neurobiol Aging. 83:63-72.
1177 Raven JC, Court JH, Raven J. 1990. Coloured progressive matrices. Manual for Raven's
1178 Progressive Matrices and Vocabulary Scales. Oxford (United Kingdom): Oxford
1179 Psychologists Press.
1180 Rikhye RV, Wimmer RD, Halassa MM. 2018. Toward an integrative theory of thalamic function.
1181 Annu Rev Neurosci. 41:163-183.
1182 Roth B, Becker N, Romeyke S, Schäfer S, Domnick F, Spinath FM. 2015. Intelligence and school
1183 grades: A meta-analysis. Intelligence. 53:118-137.
1184 Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ. 1999. Nonrigid registration
1185 using free-form deformations: application to breast MR images. IEEE Trans Med Imaging.
1186 18:712-721.
1187 Schmidt FL, Hunter J. 2004. General mental ability in the world of work: occupational attainment
1188 and job performance. J Pers Soc Psychol. 86(1):162-173.
1189 Schmithorst VJ. 2009. Developmental sex differences in the relation of neuroanatomical
1190 connectivity to intelligence. Intelligence. 37(2):164-173.
1191 Schmithorst VJ, Wilke M, Dardzinski BJ, Holland SK. 2005. Cognitive functions correlate with
1192 white matter architecture in a normal pediatric population: a diffusion tensor MRI study.
1193 Hum Brain Mapp. 26(2):139-147.
1194 Schneider WJ, McGrew KS. 2012. The Cattell-Horn-Carroll model of intelligence. In: Flanagan
1195 DP, Harrison PL, editors. Contemporary intellectual assessment: Theories, tests, and
1196 issues. Third Edition ed. New York: Guilford Press. p. 99-144.

1197 Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. 2012. Blipped-controlled
1198 aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced
1199 g-factor penalty. Magn Reson Med. 67(5):1210-1224.
1200 Simpson-Kent IL, Fuhrmann D, Bathelt J, Achterberg J, Borgeest GS, Kievit RA, Team C. 2020.
1201 Neurocognitive reorganization between crystallized intelligence, fluid intelligence and
1202 white matter microstructure in two age-heterogeneous developmental cohorts. Dev Cogn
1203 Neurosci. 41:100743.
1204 Smith SM. 2002. Fast robust automated brain extraction. Hum Brain Mapp. 17(3):143-155.
1205 Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE,
1206 Ciccarelli O, Cader MZ, Matthews PM et al. . 2006. Tract-based spatial statistics:
1207 voxelwise analysis of multi-subject diffusion data. NeuroImage. 31(4):1487-1505.
1208 Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister
1209 PR, De Luca M, Drobnjak I, Flitney DE et al. . 2004. Advances in functional and structural
1210 MR image analysis and implementation as FSL. NeuroImage. 23 Suppl 1:S208-S219.
1211 Smith SM, Nichols TE. 2009. Threshold-free cluster enhancement: addressing problems of
1212 smoothing, threshold dependence and localisation in cluster inference. NeuroImage.
1213 44(1):83-98.
1214 Spearman C. 1904. „General intelligence," objectively determined and measured. Am J Psychol.
1215 15(2):201-292.
1216 Strenze T. 2007. Intelligence and socioeconomic success: A meta-analytic review of longitudinal
1217 research. Intelligence. 35(5):401-426.
1218 Suprano I, Kocevar G, Stamile C, Hannoun S, Fourneret P, Revol O, Nusbaum F, Sappey-
1219 Marinier D. 2020. White matter microarchitecture and structural network integrity correlate
1220 with children intelligence quotient. Sci Rep. 10(1):20722.
1221 Swagerman SC, de Geus EJC, Kan KJ, van Bergen E, Nieuwboer HA, Koenis MMG, Hulshoff
1222 Pol HE, Gur RE, Gur RC, Boomsma DI. 2016. The Computerized Neurocognitive Battery:
1223 Validation, aging effects, and heritability across cognitive domains. Neuropsychology.
1224 30(1):53-64.

1225 Takahashi M, Iwamoto K, Fukatsu H, Naganawa S, Iidaka T, Ozaki N. 2010. White matter
1226 microstructure of the cingulum and cerebellar peduncle is related to sustained attention
1227 and working memory: A diffusion tensor imaging study. Neurosci Lett. 477(2):72-76.
1228 Tamnes CK, Ostby Y, Walhovd KB, Westlye LT, Due-Tonnessen P, Fjell AM. 2010. Intellectual
1229 abilities and white matter microstructure in development: a diffusion tensor imaging study.
1230 Hum Brain Mapp. 31(10):1609-1625.
1231 Tang CY, Eaves EL, Ng JC, Carpenter DM, Mai X, Schroeder DH, Condon CA, Colom R, Haier
1232 RJ. 2010. Brain networks for working memory and factors of intelligence assessed in
1233 males and females with fMRI and DTI. Intelligence. 38(3):293-303.
1234 Urger SE, De Bellis MD, Hooper SR, Woolley DP, Chen SD, Provenzale J. 2015. The superior
1235 longitudinal fasciculus in typically developing children and adolescents: diffusion tensor
1236 imaging and neuropsychological correlates. J Child Neurol. 30(1):9-20.
1237 van der Knaap LJ, van der Ham IJ. 2011. How does the corpus callosum mediate
1238 interhemispheric transfer? A review. Behav Brain Res. 223(1):211-221.
1239 van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Consortium WU-MH.
1240 2013. The WU-Minn Human Connectome Project: an overview. NeuroImage. 80:62-79.
1241 van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R, Chang A, Chen L,
1242 Corbetta M, Curtiss SW et al. . 2012. The Human Connectome Project: a data acquisition
1243 perspective. NeuroImage. 62(4):2222-2231.
1244 Varley R. 2007. Plasticity in high-order cognition: Evidence of dissociation in aphasia. Behav
1245 Brain Sci. 30(2):171-172.
1246 Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang
1247 H, Dubey P et al. . 2007. Reproducibility of quantitative tractography methods applied to
1248 cerebral white matter. NeuroImage. 36(3):630-644.
1249 Wechsler D. 2008. Wechsler Adult Intelligence Scale - Forth edition (WAIS-IV). San Antonio,
1250 Texas: Pearson Assessment.
1251 Wechsler D. 2011. Wechsler Abbreviated Intelligence Scale - Second edition (WASI-II). San
1252 Antonio, Texas: NCS Pearson.

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

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

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

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

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

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

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

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

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