


The corpus callosum as anatomical marker of intelligence? A critical examination in a large-scale developmental study

René Westerhausen¹  · Charline-Marie Friesen¹ · Darius A. Rohani¹ ·
Stine K. Krogsrud¹ · Christian K. Tamnes¹ · Jon S. Skranes² · Asta K. Håberg^{3,4} ·
Anders M. Fjell^{1,5} · Kristine B. Walhovd^{1,5}

Received: 7 April 2017 / Accepted: 4 August 2017 / Published online: 11 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract Intellectual abilities are supported by a large-scale fronto-parietal brain network distributed across both cerebral hemispheres. This bihemispheric network suggests a functional relevance of inter-hemispheric coordination, a notion which is supported by a series of recent structural magnetic resonance imaging (MRI) studies demonstrating correlations between intelligence scores (IQ) and corpus-callosum anatomy. However, these studies also reveal an age-related dissociation: mostly positive associations are reported in adult samples, while negative associations are found in developing samples. In the present study, we re-examine the association between corpus callosum and intelligence measures in a large (734 datasets from 495 participants) developmental mixed cross-sectional and longitudinal sample (6.4–21.9 years) using raw test scores rather than deviation IQ measures to account for the

ongoing cognitive development in this age period. Analyzing mid-sagittal measures of regional callosal thickness, a positive association in the splenium of the corpus callosum was found for both verbal and performance raw test scores. This association was not present when the participants' age was considered in the analysis. Thus, we did not reveal any association that cannot be explained by a temporal co-occurrence of overall developmental trends in intellectual abilities and corpus callosum maturation in the present developing sample.

Keywords Corpus callosum · Intelligence · Development · Structure–function association · MRI · Longitudinal studies

Introduction

Intellectual and general cognitive abilities are supported by a large-scale brain network encompassing association cortices in frontal, parietal, and temporal lobes (Deary 2012; Deary et al. 2010; Jung and Haier 2007; Luders et al. 2009; Shaw 2007). The nodes constituting this network are distributed across both cerebral hemispheres, emphasizing the relevance of functional interaction between the hemispheres for performance in tasks demanding higher intellectual abilities (e.g., Belger and Banich 1998; Davis and Cabeza 2015; Welcome and Chiarello 2008, for review see Banich 2003). Consequently, it has long been suggested that the strength of the commissural connections—e.g., reflected in the mid-sagittal size of the corpus callosum—might serve as an anatomical marker of higher intellectual abilities (Hulshoff Pol et al. 2006; Men et al. 2014; Spitzka 1907; Strauss et al. 1994). A larger corpus callosum and/or thicker myelinated callosal axons would improve inter-hemispheric connectivity and, in turn, intellectual performance. In line with this

Electronic supplementary material The online version of this article (doi:10.1007/s00429-017-1493-0) contains supplementary material, which is available to authorized users.

✉ René Westerhausen
rene.westerhausen@psykologi.uio.no

¹ Department of Psychology, Center for Lifespan Changes in Brain and Cognition (LCBC), University of Oslo, Blindern, POB 1094, 0317 Oslo, Norway

² Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway

³ Department of Medical Imaging, St. Olav's Hospital, Trondheim, Norway

⁴ Department of Neuroscience, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

⁵ Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway

interpretation, a common genetic origin for corpus callosum size and intelligence has been suggested in studies on healthy twins and siblings (Hulshoff Pol et al. 2006). This general notion is supported by a series of recent magnetic resonance imaging (MRI) studies demonstrating correlations between intelligence coefficients (IQ; as measured with standard intelligence tests) and measures of mid-sagittal callosal area (e.g., Allin et al. 2007; Ganjavi et al. 2011; Hutchinson et al. 2009; Peterson et al. 2001), thickness (Luders et al. 2007, 2011), and microstructural integrity (i.e., fractional anisotropy; Chiang et al. 2009; Dunst et al. 2014; Hutchinson et al. 2009; Navas-Sanchez et al. 2014; Tang et al. 2010). These studies also reveal an apparent dissociation in the direction of the reported correlations as a function of age. Studies examining adult samples mostly report positive associations (Chiang et al. 2009; Dunst et al. 2014; Luders et al. 2007; Strauss et al. 1994; and selectively in female participants in Tang et al. 2010). On the other hand, studies examining children, adolescents, and young adults report negative associations (Allin et al. 2007; Ganjavi et al. 2011; Hutchinson et al. 2009; Luders et al. 2011; but see Nosarti et al. 2004). For example, Luders et al. (2007) studying a mostly adult sample (age range 16–44 years) observed significant positive correlations between IQ measures and thickness in the posterior half of the corpus callosum. Studying a developing sample (age range 6–17 years) with the same methodological approach, the same group found a significant negative correlations, again located in the posterior corpus callosum (Luders et al. 2011). It has been suggested that this developmental dissociation might be attributed to differences in the relative weight of intra- vs. inter-hemispheric processing (Ganjavi et al. 2011) and age-related changes in task demands (Hutchinson et al. 2009).

When interpreting these previous findings it also has to be considered that in all the above studies intelligence was quantified using age-standardized, norm-deviation IQ scores. These deviation IQ scores reflect the relative position within the norm group, and do not represent the absolute level of performance (Angoff 1984; Neisser 1997; Wechsler 1999). As a consequence differences in performance levels between norm groups (i.e., between participants converted with different conversion tables) are removed and set to the norm distribution's mean (usually 100; see e.g., Angoff 1984; Neisser 1997). The latter effect of the IQ conversion is especially pronounced when studying development samples since the provided conversion tables for children and adolescents cover very narrow age spans reflecting the rapid absolute intellectual development in this period of life. For example, the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler 1999), which was used by all of the above-mentioned developmental studies, provides conversion tables covering age spans of only 4 months for participants under the age of

16 year. As a result, the mean of the converted IQ scores will appear stable throughout childhood and adolescence (e.g., Burgaleta et al. 2014) although the absolute level of performance (e.g., the mean raw test scores) rises continuously in this period of life (Neisser 1997; Tamnes et al. 2010). While the conversion to IQ score allows evaluating age-appropriateness of an individual's intellectual abilities and rank stability within a cohort over time, only the use of raw test score will allow to study the development of intellectual functioning. Therefore, we here argue that to address the question of whether corpus-callosum morphology can serve as marker of intellectual and cognitive abilities during development, measures of absolute rather than norm-relative performance should be used.

The present study was designed to systematically re-examine the association of IQ test measures and corpus-callosum anatomy during development in a large mixed cross-sectional and longitudinal sample (734 datasets) by for the first time using raw test scores rather than deviation IQ. This was done in three analyses steps. First, to establish a general structure–function association, we related verbal raw test scores (*v*-RS) and performance raw test scores (*p*-RS) to regional callosal thickness measures. As brain size has previously been found to be positively correlated with both corpus callosum size (Ganjavi et al. 2011; Jäncke et al. 1999; Westerhausen et al. 2016) and intelligence measures (for a meta-analyses see McDaniel 2005; or Pietschnig et al. 2015) total intracranial volume (TIV) was included as an additional predictor in a second analysis step. Finally, in a third step, we examined the effect of chronological age on the structure–function association. This was necessary as raw intelligence test scores (Neisser 1997; Tamnes et al. 2010) and callosal thickness (e.g., Giedd et al. 1996; Luders et al. 2010b; Westerhausen et al. 2016, 2011) are characterized by a continuous age-related increase in the studied age period, and the temporal co-occurrence of these two developmental trends could confound the association (see e.g. Salthouse 2011).

Methods

Participants

Participants were drawn from two longitudinal imaging studies, the Norwegian Mother and Child Cohort Neurocognitive study (NeuroCogMoBa; e.g., Krogsrud et al. 2016) and Neurocognitive Development (NeuroCogDev; e.g., Tamnes et al. 2013), both coordinated by the Lifespan Changes in Brain and Cognition (LCBC) research center, University of Oslo, Norway. The analyzed sample included 734 datasets from 495 (250 female) participants, of which 239 (131 female) provided measures for a second time

point, and covering an age range between 6.4 and 21.9 years. This final sample represented a subsample of a total of 1085 data sets of which 83 data/participants were excluded due to a history of mental or neurological disorders, reported (by parents) to have premature birth (<37 weeks) and low birth weight (<2500 g), as well as due to low-quality of the T1-weighted MRI data. Another 263 datasets/participants had to be excluded as no WASI data was available or the available test data was incomplete. Thus, the final sample included all datasets of the above studies for which both high-quality MRI data and valid intelligence testing with the WASI were available.

The studies were approved by the Regional Committee for Medical and Health Research Ethics. Until the age of 16, the care legal guardians of the participants provided written informed consent. Additionally, informed assent was given in written form by participants above the age of 12, and in oral form by participants below the age of 12.

Intelligence assessment

Intelligence assessment for all participants and time points was conducted with a Norwegian version of the WASI (Wechsler 1999), including the subtests Vocabulary and Similarities for estimating verbal abilities as well as the subtests Block-Design and Matrix Reasoning for estimating performance (non-verbal) abilities. Raw subscale (RS) scores were calculated as weighted average raw test score, that is, the proportion of correct answers (i.e., the individual score divided by the maximum score that could be achieved in that subtest) was calculated per subtest and averaged across the relevant verbal (v-RS), performance (p-RS), or all four subtests (fs-RS), respectively. For comparison, age-standardized deviation IQ scores for verbal (v-IQ), performance (p-IQ), and estimated “full-scale” IQ (fs-IQ) were obtained using the age-appropriate conversion tables provided in the WASI test manual.

For illustration, we performed an explorative regression analysis determining the empirical association between RS and standardized IQ measures. Both linear and quadratic effects of RS on IQ measures were included in the model. This analysis was done separately for participants tested the first and second time, respectively. Although all models were statistically significant (all $p < 0.001$), the association between RS and IQ scores was, as expected, far from perfect, explaining at maximum 42% of the variance. In detail, at baseline testing ($n = 495$) coefficient of determination was $R^2 = 0.37$ for the prediction of v-IQ, $R^2 = 0.35$ for the prediction of p-IQ, and $R^2 = 0.31$ for the prediction of fs-IQ. At follow-up testing ($n = 239$), it was $R^2 = 0.26$ for v-IQ, $R^2 = 0.42$ for p-IQ, and $R^2 = 0.31$ for the prediction of fs-IQ. Across all analyses both the linear (all regression weights positive) and quadratic predictors

(all negative regression weights) were significant (all $p < 0.05$), except for the v-RS analysis at follow-up testing in which the quadratic prediction was not significant ($p = 0.41$). This overall pattern of association underlines the expected non-equivalence of RS and deviation IQ in a developmental sample (Neisser 1997). Thus, in accordance with the objectives of the present study we focused on the raw test scores for further analyses.

To further describe the developmental trajectories of v-RS and p-RS, we also conducted two separate linear-mixed model analyses (restricted maximum likelihood estimations, REML) using chronological age as linear (Age) and quadratic (Age squared) predictors (fixed effects), as well as allowing for different intercepts between participants (random effect term). In both cases, Age (v-RS: $\beta_{\text{Age}} = 0.049$; $t_{731} = 47.3$, $p < 0.0001$; p-RS: $\beta_{\text{Age}} = 0.061$; $t_{731} = 37.8$, $p < 0.0001$) and Age-squared (v-RS: $\beta_{\text{Age squared}} = -0.0025$; $t_{731} = -14.4$, $p < 0.0001$; p-RS: $\beta_{\text{Age squared}} = -0.0036$; $t_{731} = -13.8$, $p < 0.0001$) contributed significantly to the prediction, together describing the monotonous increase of level of performance which flattens towards young adulthood (see Fig. 1).

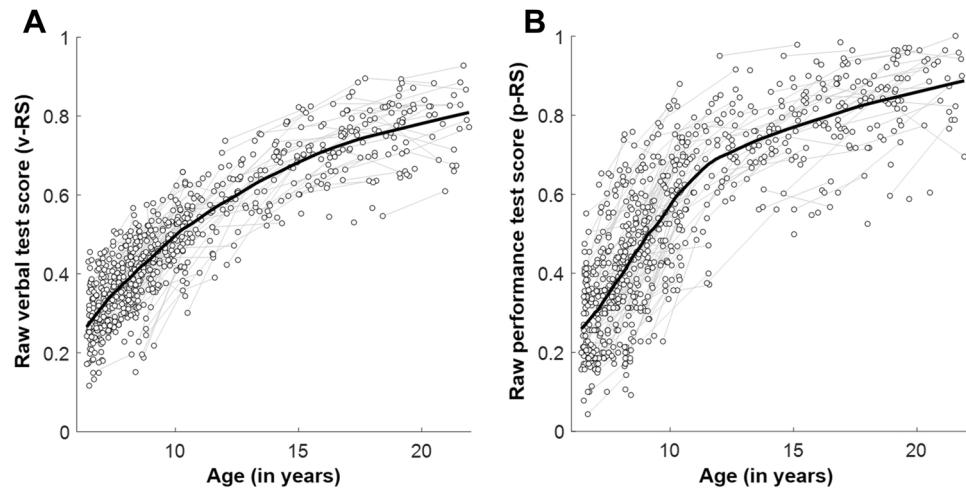
MRI acquisition

The MRI data used in the present analysis stems from two different research projects (NeuroCogMoBa, NeuroCogDev) and two different sites (Rikshospitalet, Oslo, and St. Olav’s Hospital, Trondheim), but all MRI scans were acquired with the same scanner model and equipment (1.5T Siemens Avanto, a 12-channel head coil) as well as using the very same pulse sequences. T1 weighted images were acquired with the following parameters: echo time, TE = 3.61 ms, repetition time, TR = 2400 ms, inversion time, TI = 1000 ms, and a flip angle of 8 degrees. For each participant 160 sagittal slices with a thickness of 1.2 mm were taken, with an image resolution of $1.25 \times 1.25 \times 1.2 \text{ mm}^3$, a field view of $240 \times 240 \text{ mm}^2$ and a 192×192 scan matrix. The data acquisition for the MOBA participants was done using a parallel imaging technique (iPAT, GRAPPA factor 2) acquiring multiple (between 2 and 4) T1 volumes in a short scan time (4 min and 18 s per volume). Of the final 734 datasets 542 datasets were acquired in Oslo and 192 in Trondheim. Raw image quality was assessed based on visual inspection and performed by two experienced examiners (D.A.R., S.K.K.) and the best dataset was used for the analysis.

Measurement of callosal thickness

Corpus callosum morphology was assessed using thickness measurements determined on callosal outline on the mid-sagittal cross-sectional surface area on the T1-weighted

Fig. 1 Intellectual development. Spaghetti plot illustrating the development of **a** raw verbal (v-RS) and **b** raw performance (p-RS) test scores for 495 participants of which 239 provided measures for two time points (connected with gray lines). Solid black line represents the locally weighted scatterplot smoothing (LOWESS, smoothing factor, $f = 0.25$)



images. This approach was preferred over the traditional geometrical subdivision since it is less dependent on inter-individual differences in the curvature of the corpus callosum and offers a better regional specificity (Luders et al. 2006). The present implementation of the approach has been described elsewhere (Westerhausen et al. 2016), and is briefly summarized here. First, all individual images were coregistered to a template using a rigid-body transformation and segmented in native space (using SPM12; <http://www.fil.ion.ucl.ac.uk/spm/>). Then, in a semi-automated procedure, the corpus callosum was identified on the midsagittal slice of the segmented white-matter volume and manually adjusted where necessary (to remove “non-callosal” voxels, e.g., fornix). The resulting callosal mask was extracted as binary image and an outline was created by removing all non-border voxels from the segmented corpus callosum. The tip of the rostrum (posterior-most voxel of the in-bend rostrum in the anterior half) and base of the splenium were identified (ventral-most voxel in the posterior half). Subsequently, the outline was rotated so that the imagined line connecting rostrum tip and base of the splenium were horizontally oriented. Rostrum tip and the base of the splenium determined the division into a ventral (“lower”) and dorsal (“upper”) outline. The midline between ventral and dorsal outline was determined based on support points spaced equidistantly on the two outlines. Finally, the midline was resampled into 60 equidistant points along the midline, orthogonal to which callosal thickness was determined as distance between ventral and dorsal outline. The resulting 60 regional thickness measures for each participant and time point were used for the statistical analysis.

Measurement of brain compartment volume

Total intracranial volume, TIV, estimates were obtained by means of automated segmentation routines (“tissue volumes” utility) using SPM12, and was defined as the sum of

grey matter, white matter, and CSF compartments, including forebrain, midbrain, hindbrain, and cerebellum. The inferior limit for TIV estimation was set at the boarder of the cerebellum. This approach was selected as it has demonstrated high external validity (Hansen et al. 2015). The mean TIV was 1.43 L (SD = 0.14) across all subjects and time points, ranging from 1.10 to 2.01 L.

Statistical analysis

To integrate cross-sectional and longitudinal data in the same analysis, a linear mixed model design (Verbeke and Molenberghs 2009) was conducted for the regional thickness analysis. To establish a general association between RS intelligence measures and callosal thickness, the first set of analyses (step 1) included a fixed-effect part with the predictors for the respective Test Score (either v-RS, or p-RS; demeaned continuous variable), Sex, as well as the interaction of these two variables. The random effect part consisted of Participant (allowing for different intercepts) to account for the repeated measures available for 239 of the participants. Additionally, handedness (coded as right-handed vs. non-right-handed) and Site were added as nuisance variables. Handedness was included as it is related to differences in corpus callosum size and thickness (Luders et al. 2010a; Westerhausen et al. 2004). The predictor Site was included to account for possible differences resulting from using different MR scanners, an approach which can be considered appropriate as the here aggregated studies were conducted on the same scanning platforms and used the same scanning sequence (Chen et al. 2014). The regional thickness in each of the 60 segments, served as dependent variable. In analysis step 2, the above model was extended by introducing TIV as additional covariate. In step 3, chronological age was considered by adding predictors representing linear (Age) and quadratic relations (Age squared), as well as the interaction of Test Score and

Age. The same three-step analysis was conducted using the fs-RS. However, as fs-RS was naturally highly correlated with the two constituting sub-test scores (see section Intelligence Assessment), we here only present analyses for v-RS and p-RS in detail (see Supplement Fig. 1 for the analyses of the fs-RS data). Although the use of deviation-IQ scores was not considered appropriate for the present research question the three analysis steps were also conducted for the respective IQ scores (v-IQ, p-IQ, fs-IQ) to allow the interested reader to compare the present to previous findings. The results of these analyses are provided in Supplement Fig. 2.

All analyses were conducted utilizing restricted maximum likelihood estimations (using a full covariance matrix; Cholesky parameterization), and the models were fitted with the “fitlme” function provided in MATLAB (R2015b, The MathWorks, Natick, MA, USA). The significance level was adjusted to a false-discovery-rate (FDR) of 0.05 using Benjamini–Hofberg–Yekutieli (Benjamini and Yekutieli 2001) procedure (i.e., not assuming positive dependency of the tests within one analysis). The results are presented with unstandardized regression weights (β) to express the effect on callosal thickness in absolute measures (i.e., mm) as well as using percentage explained variance (ω^2) as relative measure.

Across analyses, the effects of interest were: (a) the main effect of Test Score (v-RS or p-RS), indicating an overall structure–function association (analysis steps 1–3); (b) the interaction of Test Score with Sex, indicative of slope difference between male and female subjects (step 1–3); (c) the interaction of Test Score with Age, indicative of changes in the slope during development (step 3). For the effects of interest, the main analyses were supplemented with a power analysis using G*Power 3 (Faul et al. 2009), whereby a sensitivity analysis was conducted to determine the smallest effect size which can be detected with a power of 0.80 and 0.95. The power analysis was calculated assuming an α -level of 0.00083 (i.e., representing a scenario with FDR threshold reaching Bonferroni threshold for 60 tests), two-tailed testing, and considering the fixed-effect part only. The smallest detectable effect size was Cohen’s $f^2 = 0.024$ and $f^2 = 0.034$, for 0.80 and 0.95 test power, respectively (only marginally differing from step 1 to step 3, with 5–9 predictors, respectively) which is equivalent to an explained variance of $\omega^2 = 0.023$ – 0.033 . That is, population effects larger than 2.3 or 3.3% can be excluded for non-significant associations.

Results

The analysis 1 for both p-RS and v-RS revealed a significant positive association of test performance and callosal thickness in the posterior corpus callosum including most

parts of the splenium section (see upper row, left panel, Figs. 2, 3). For p-RS the strongest association was found in segment 58 with an explained variance of $\omega^2 = 0.07$ ($\beta_{p-RS} = 1.00$; $t_{728} = 5.01$, $p < 0.0001$, see Fig. 4). Comparably, for v-RS the strongest association was also found in segment 57 ($\beta_{v-RS} = 1.01$; $t_{728} = 5.03$, $p < 0.0001$; $\omega^2 = 0.07$, see Fig. 4). The v-RS analysis additionally revealed a negative association in the genu of the corpus callosum, that is, in segment 13 ($\beta_{v-RS} = -0.55$; $t_{728} = -3.14$, $p = 0.0018$, $\omega^2 = 0.03$, see Fig. 2). In addition to the main effect, both analyses also revealed a significant interaction of test score and Sex indicating differences in the slope of the association between the sexes. In the p-RS analysis the interaction was found in extended areas in the genu as well as in the truncus region with the maximum effect being located in segment 9 ($\beta_{p-RS*Sex} = 1.44$; $t_{728} = 4.12$, $p < 0.0001$, $\omega^2 = 0.05$). Across all significant segments, the interaction was driven by a more positive slope in female than in male participants (see Fig. 4d). In the v-RS analysis the area of significant interactions was restricted to the genu region with the maximum being located in segment 12 ($\beta_{v-RS*Sex} = 1.09$; $t_{728} = 4.17$, $p < 0.0001$, $\omega^2 = 0.05$). Also here the interaction was driven by a more positive slope in female than male participants (see Fig. 4b).

In analysis step 2, including TIV as covariate, for both intelligence measures a positive association with thickness in posterior callosal segments was found (see Figs. 2, 3, second row). Comparable to analysis step 1, the strongest association was located in segment 57 for the v-RS analysis ($\beta_{v-RS} = 1.13$; $t_{727} = 5.54$, $p < 0.0001$; $\omega^2 = 0.09$, see Fig. 5) and in segment 58 for the p-RS analysis (for p-RS: $\beta_{p-RS} = 1.07$; $t_{727} = 5.36$, $p < 0.0001$; $\omega^2 = 0.08$). However, including TIV in the model, the interaction of Sex and Test Score was no longer significant (for v-RS all $|\beta_{v-RS*Sex}| < 0.51$; all $|t_{727}| < 1.66$; and for p-RS all $|\beta_{p-RS*Sex}| < 0.69$; all $|t_{727}| < 2.08$; all non-significant using FDR correction).

Finally, including Age covariates (i.e., linear term, quadratic term, interaction of test score and Age) in analysis step 3, the main effect of test score vanished for both test subscores (for v-RS all $|\beta_{v-RS}| < 1.04$, all $|t_{724}| < 2.41$; for p-RS all $|\beta_{p-RS}| < 0.76$, all $|t_{724}| < 2.00$; all non-significant using FDR correction) for both p-RS and v-RS analyses. In addition, the interaction of test score and Sex did not reach significance (for v-RS all $|\beta_{v-RS*Sex}| < 0.47$, all $|t_{724}| < 1.53$; and for p-RS all $|\beta_{p-RS*Sex}| < 0.70$, all $|t_{724}| < 2.11$). In neither of the two analyses a significant interaction of Test Score and Age was found (for v-RS all $|\beta_{v-RS*Age}| < 0.20$, all $|t_{724}| < 2.72$; and for p-RS all $|\beta_{p-RS*Age}| < 0.22$, all $|t_{724}| < 2.52$; all non-significant using FDR correction, see Fig. 6), that is, the association of intelligence and callosal thickness was not significantly modulated by Age.

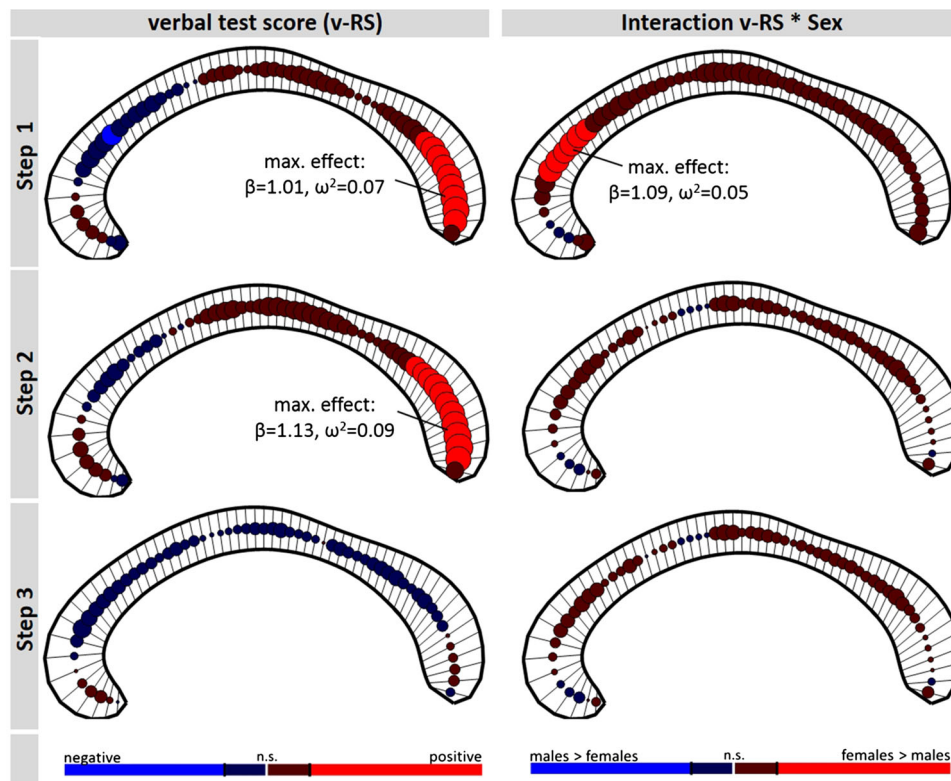


Fig. 2 Association raw verbal test score (v-RS) and regional callosal thickness (*left column*) and the interaction of v-RS with Sex (*right column*). The rows represent the three analysis steps (full models described in “Method” section), with the statistical design of step 2 compared to step 1 additionally including TIV as covariate, and step 3 compared to step 2 additionally including age-related variables (i.e., Age linear, Age squared, and the interaction of v-RS and Age). At

each of the 60 segments of the corpus callosum, the direction and magnitude of the association is visualized by a *circle*, whereby the size of the *circle* is proportional to empirical *t* value and the *color*, *red* vs *blue*, codes positive and negative associations, respectively. *Lighter red* and *lighter blue* indicate significant associations, with the significance level being adjusted to a false-discovery-rate (FDR) of 0.05. The anterior corpus callosum is on the *left side* of each panel

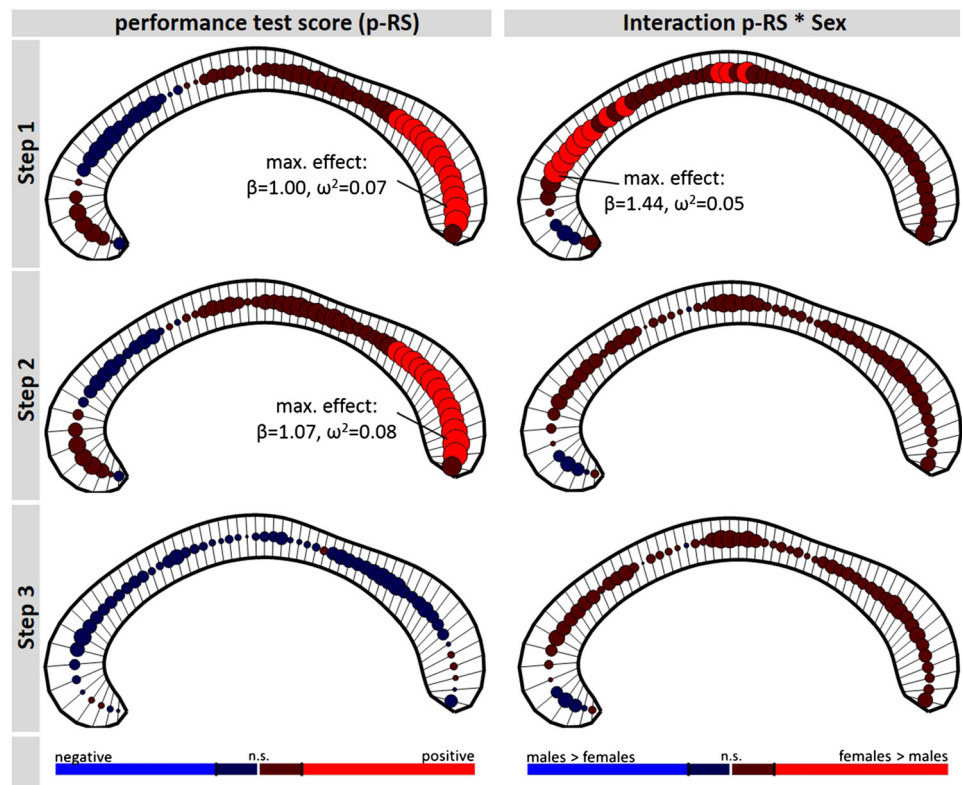
Discussion

The present study re-examined the relationship between intellectual development and callosal maturation during childhood, adolescence, and young adulthood. When not accounting for chronological age, a positive association in the splenium of the corpus callosum was detected for both p-RS and v-RS. At the thickness segments with the largest effect size, an increase in p-RS or v-RS by one standard deviation was reflected by an increase in regional callosal thickness of about 1 mm. This positive association was found irrespective of whether individual differences in brain size were accounted for in the analysis or not. Considering the known topographical organization of the corpus callosum, the splenium region interconnects temporal, parietal, and occipital cortices, including sensory and higher association areas (Putnam et al. 2010; Schmahmann and Pandya 2006; Westerhausen et al. 2009). Temporal, parietal, and occipital regions have shown associations to IQ scores in developing (Burgaleta et al. 2014; Karama et al. 2011; Menary et al. 2013; Shaw et al. 2006) as well as

adult samples (Choi et al. 2008; Haier et al. 2004; Narr et al. 2007). Furthermore, the present posterior callosal association might reflect especially the maturation of the “parietal functions” within parieto-frontal integration theory of intelligence (PFIT; Jung and Haier 2007), and be associated in particular with early processing and integration of sensory information. A “non-functional” interpretation is also conceivable; any callosal variability is likely to reflect differences in the developing cortical architecture but might not itself contribute to the level of performance (Strauss et al. 1994). That is, the corpus callosum would serve as a “marker” for an “intelligent brain”.

However, the above interpretation deserves further clarification. When introducing chronological age into the statistical model, the suggested associations between callosal thickness and raw intelligence scores disappeared. In the studied period of life, between 6 and 22 years, it is known that intelligence test performance when expressed as raw scores (Tamnes et al. 2010, see also Fig. 1) and callosal thickness (e.g., Giedd et al. 1996; Luders et al. 2010b; Westerhausen et al. 2016) are characterized by a

Fig. 3 Association of raw performance test score (p-RS) and of regional callosal thickness (*left column*) and the interaction of p-RS with Sex (*right column*). As in Fig. 2, rows represent the three analysis steps. At each callosal segment, the direction and magnitude of the association is visualized by a circle. The size of the circle is proportional to empirical *t* value and positive and negative associations are coded red and blue respectively. Lighter red and lighter blue indicate significant associations [significance level adjusted to a false-discovery-rate (FDR) of 0.05]. The anterior corpus callosum is on the left side of each panel



continuous, monotonous age-related increase. Consequently, the temporal co-occurrence of these two developmental trends alone could drive the present structure–function association without the two variables necessarily being interrelated (e.g., Salthouse 2011, for a more general discussion). Thus, to examine the interdependence beyond the co-occurring age-related covariance, we introduced chronological age to the design, which yielded the effect of Test Score for both p-RS and v-RS non-significant. This “null” finding by itself does not unequivocally exclude the existence of a true functional association between intelligence measures and callosal thickness. Assuming, for example, that callosal maturation would be the only determinant of intelligence test performance, both variables would exhibit a strong covariation of rates of change during development, a correction for age would likely leave the two variables uncorrelated. However, the also non-significant interaction of Test Score and Age further indicates that the lack of covariance of Test Score and callosal thickness is invariant across all ages studied (Hofer et al. 2006; Salthouse 2011). Thus, the lack of a significant effect of Test Score together with the non-significant interaction of Test Score and Age, indicate that for no chronological age a significant structure–function association was found. Given sufficient test power to exclude effects larger than 2–3% explained variance, the present findings render any substantial functional relevance of

macrostructural callosal variability for intelligence test performance unlikely, within the studied age range.

The non-relevance of individual differences in callosal thickness for intelligence test performance is, on one hand, not surprising as it is in line with studies on patients with partial or complete callosotomy (e.g., Mamelak et al. 1993; Oguni et al. 1991; Tanriverdi et al. 2009). Pre-post-surgery comparisons usually fail to find a substantial decrease in IQ scores as consequence of the callosal transection. For example, Mamelak et al. (1993) examined 15 epilepsy patients aged between 9 and 31 years (i.e., also covering the age range examined here) and did neither find substantial change in verbal (mean change in v-IQ: 0.38, SD = 3.52; from Table 3, p. 692, of Mamelak et al. 1993) nor in performance IQ (mean change in p-IQ: 1.23, SD = 3.56). Nevertheless, it has to be kept in mind that findings in callosotomy patients are not fully representative, as the patients’ pre-surgical brain anatomy certainly deviates from a healthy brain including plastic adaption processes to the epilepsy-related changes (Campbell et al. 1981). On the other hand, the present findings are partly at odds with a series of earlier studies demonstrating that inter-individual difference in callosal architecture are predictive of functional inter-hemispheric integration in tasks assessing cognitive-control functions (Davis and Cabeza 2015; Huster et al. 2011; Kompus et al. 2011; Schulte et al. 2006). However, a critical component of these tasks was

Fig. 4 Spaghetti plots at location of maximum effect of analysis step 1. Plot **a** shows the main effect of verbal raw (v-RS) test scores on callosal thickness (residualized for all other effects) in segment 57 (red line). Plot **b** illustrates the interaction of v-RS with Sex with filled circles representing female (red line) and open circles representing male (blue line) participants. Plot **c** illustrates the main effect of performance raw (p-RS) test scores on residualized callosal thickness in segment 58 (red line). Plot **d** shows the interaction of p-RS with Sex with filled circles representing female (red line) and open circles representing male (blue line) participants

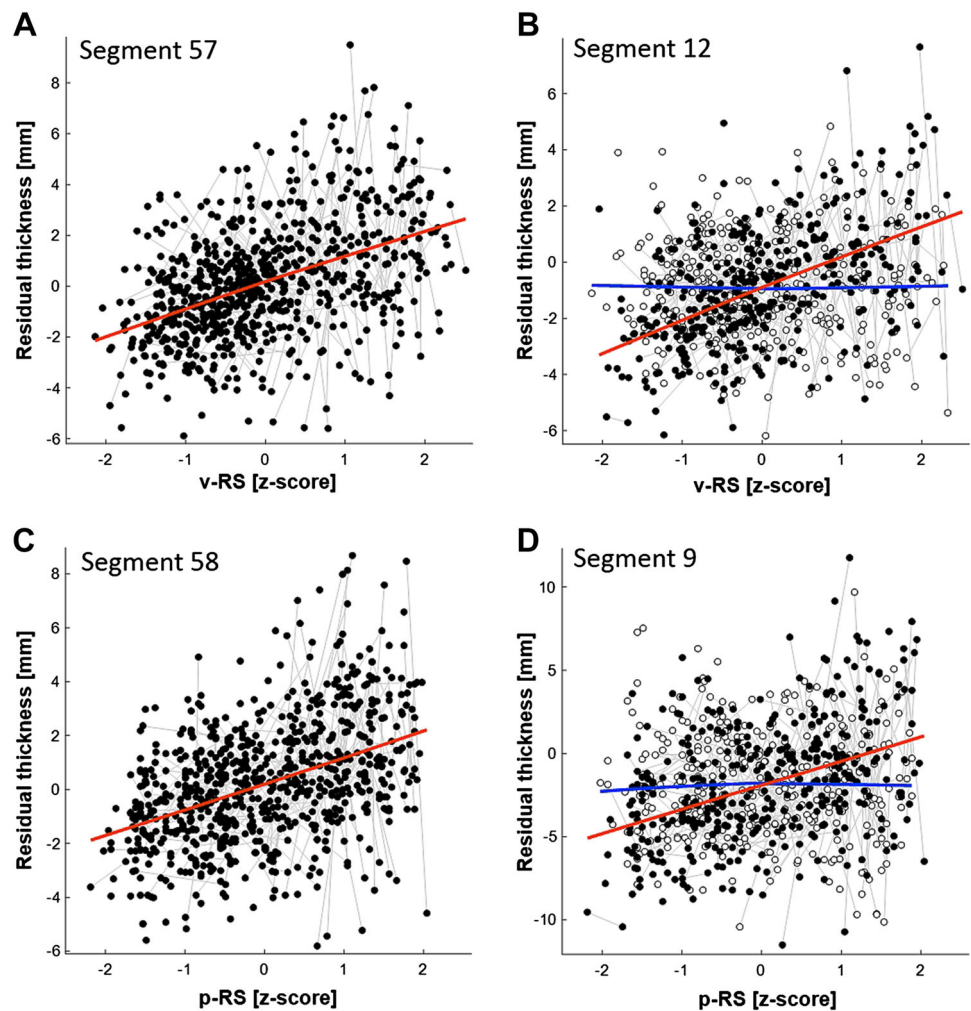
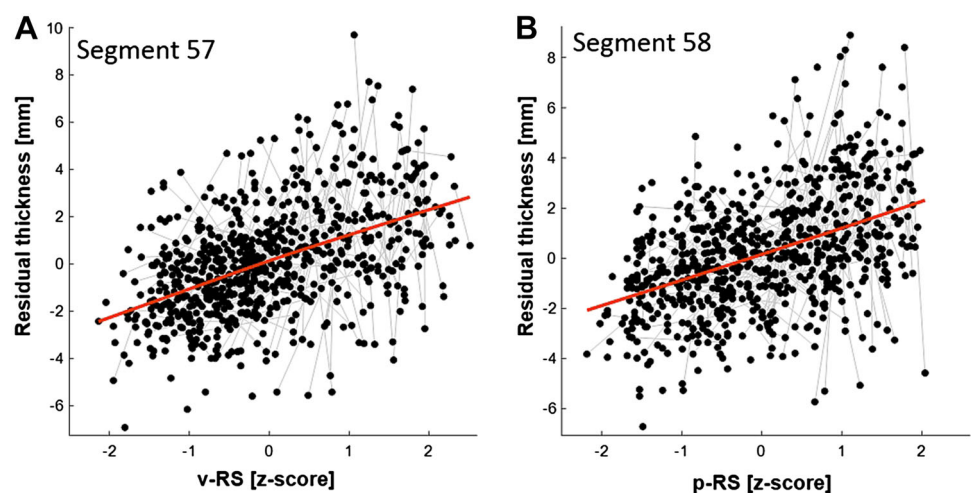


Fig. 5 Spaghetti plots at location of maximum effect of analysis step 2. Plot **a** illustrates the main effect of verbal raw (v-RS) test scores on residualized callosal thickness in segment 57 (red line depicting linear fit) after considering TBV differences. Likewise, plot **b** shows the main effect of performance raw (p-RS) test scores on residualized callosal thickness in segment 58 (red line depicting linear fit)



that speeded choice reactions were required, demanding a rapid exchange of information between the hemispheres. The performance in the four tasks which constitute the WASI test, on the other hand, is not relying on fast stimulus–response mapping in a similar manner. Also the

association between general white-matter tract fractional anisotropy and intelligence measures has been previously shown to be mediated by general information-processing speed (Penke et al. 2012). Thus, it is conceivable that in the present study no structure–function association was found

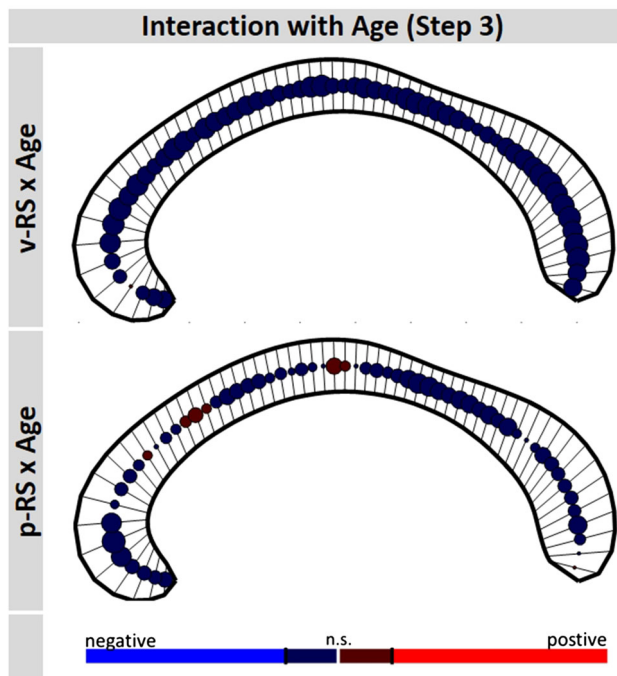


Fig. 6 Interaction of raw test scores (performance, p-RS; verbal, v-RS) and Age in predicting regional callosal thickness in analysis step 3. At each callosal segment, the direction and magnitude of the association is visualized by a *circle*. The size of the *circle* is proportional to empirical *t* value of the predictor coding the interaction; positive and negative associations are coded *red* and *blue*, respectively. Significance level is adjusted to a false-discovery-rate (FDR) of 0.05. The anterior corpus callosum is on the *left side* of each *panel*

since time-critical communication between the hemispheres was not required to the same extent in the WASI tests.

The present findings also need to be discussed in relation to previous studies correlating deviation IQ scores with measures of callosal anatomy in comparable developmental samples. Irrespective of whether interpreting the outcome of the analyses without (positive association) and with Age covariates (no association), the present findings do not align with a series of previous developmental studies reporting negative associations (Allin et al. 2007; Ganjavi et al. 2011; Hutchinson et al. 2009; Luders et al. 2011). As also argued in the introduction, deviation IQ scores reflect the relative position of an individual within a specific norm age group and no longer the absolute level of performance (Angoff 1984). This is, for example, reflected in the present sample as raw scores in the prediction of IQ scores explains at maximum 42% of the variance (see “Method” section). Furthermore, the conversion from raw measure to IQ score also removes overall differences in performance level between individuals that fall into different age groups (different conversion table are used) and set to the mean to the norm distribution’s mean (Wechsler 1999). The converted

IQ scores will consequently appear stable (Burgaleta et al. 2014) throughout childhood and adolescence while the absolute level of performance naturally rises continuously in this life period (see Fig. 1). At the same time, however, none of the previous studies converted (or adjusted to norm-data) callosal morphology measures. Rather, two of the above developmental studies (Allin et al. 2007; Luders et al. 2011) related deviation IQ scores to raw measures of midsagittal callosal area and regional callosal thickness, respectively, both without correcting for chronological age. Considering the properties of deviation IQ the reported findings can be reformulated to: the participants’ relative position in performance with respect to their age-specific norm group was negatively associated with the participants’ absolute measures of corpus-callosum size/thickness. In the present study, a positive correlation between absolute level of performance and absolute callosal thickness was detected (when not considering Age in the analysis). Of note, the attempt to replicate the negative associations of Luders et al. (2011) and Allin et al. (2007) using deviation IQ scores did not yield any significant associations in the present sample (see Supplementary Fig. 2). Two other studies (Ganjavi et al. 2011; Hutchinson et al. 2009) introduced chronological age as covariate into the statistical design while at the same time keeping the already age-adjusted deviation IQ score. While the covariate will remove age-related variance also from the callosal measures by implicitly treating these as deviation from the predicted mean for the respective age, it does not express the data relative to an external norm distribution as it is done for deviation IQ score, again making the interpretation difficult. Considering Age as covariate and in interaction with Test Scores in the present study, did not reveal any significant association between test performance and absolute thickness.

Finally, previous studies also report sex differences in the association of callosal and deviation IQ measures (Dunst et al. 2014; Luders et al. 2007; Tang et al. 2010), whereby in developmental samples a more negative associations in male compared to female subsamples has been reported (Ganjavi et al. 2011; Luders et al. 2011). In the present study, we find a Test score by Sex interaction in the genu of the corpus callosum for v-RS and p-RS in analysis step 1, which was in both cases driven by a positive association in females while no significant association was found in male participants (see Fig. 4). As axons located in the genu interconnect prefrontal cortices (Benedictis et al. 2016; Schmahmann and Pandya 2006) the findings might relate to the frontal components of intelligence (Jung and Haier 2007), whereby selectively female participants seem to benefit from a stronger structural connectivity via the corpus callosum. However, this dissociation only holds for “absolute” callosal measures; once TIV was accounted for the interaction did not survive FDR correction.

In summary, intellectual abilities are supported by a large-scale bihemispheric network (Deary 2012; Jung and Haier 2007; Shaw 2007) suggesting functional relevance of inter-hemispheric coordination (Banich 2003; Davis and Cabeza 2015). The present study was able to confirm that a general structure–function correlation exists during development but only as long as the participants’ age was not considered. Thus, we did not find any association that cannot be explained by a temporal co-occurrence of overall developmental trends in intellectual development and structural callosal increase. However, we here examined the macrostructural development using regional thickness measures. It remains for future studies to determine whether these findings are confirmed when indices of microstructural differences, such as fractional anisotropy, are assessed. Future studies also have to establish, whether the present finding utilizing RS instead of deviation-IQ extend into adult age, as the structural development of the corpus callosum continuous at least to the end of the third life decade (e.g., Pujol et al. 1993).

Acknowledgements This research was supported by grants from the Norwegian Research Council K. B. W. and A. M. F., as well as by the Department of Psychology, University of Oslo.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allin M, Nosarti C, Narberhaus A, Walshe M, Frearson S, Kalpakidou A, Murray R (2007) Growth of the corpus callosum in adolescents born preterm. *Arch Pediatr Adolesc Med* 161(12):1183–1189. doi:10.1001/archpedi.161.12.1183
- Angoff WH (1984) Scales, norms, and equivalent scores. Educational testing service, Princeton
- Banich MT (2003) Interaction between the hemispheres and its implications for the processing capacity of the brain. In: Davidson RJ, Hugdahl K (eds) *The asymmetrical brain*. MIT Press, Cambridge, pp 261–302
- Belger A, Banich MT (1998) Costs and benefits of integrating information between the cerebral hemispheres: a computational perspective. *Neuropsychology* 12(3):380
- Benedictis A, Petit L, Descoteaux M, Marras CE, Barbareschi M, Corsini F, Sarubbo S (2016) New insights in the homotopic and heterotopic connectivity of the frontal portion of the human corpus callosum revealed by microdissection and diffusion tractography. *Hum Brain Mapp* 37(12):4718–4735
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29(4):1165–1188
- Burgaleta M, Johnson W, Waber DP, Colom R, Karama S (2014) Cognitive ability changes and dynamics of cortical thickness development in healthy children and adolescents. *Neuroimage* 84:810–819
- Campbell AL, Bogen JE, Smith A (1981) Disorganization and reorganization of cognitive and sensorimotor functions in cerebral commissurotomy. *Brain* 104(3):493–511
- Chen J, Liu J, Calhoun VD, Arias-Vasquez A, Zwiens MP, Gupta CN, Turner JA (2014) Exploration of scanning effects in multi-site structural MRI studies. *J Neurosci Methods* 230:37–50. doi:10.1016/j.jneumeth.2014.04.023
- Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C, Thompson PM (2009) Genetics of brain fiber architecture and intellectual performance. *J Neurosci* 29(7):2212–2224. doi:10.1523/JNEUROSCI.4184-08.2009
- Choi YY, Shamosh NA, Cho SH, DeYoung CG, Lee MJ, Lee J-M, Gray JR (2008) Multiple bases of human intelligence revealed by cortical thickness and neural activation. *J Neurosci* 28(41):10323–10329
- Davis SW, Cabeza R (2015) Cross-hemispheric collaboration and segregation associated with task difficulty as revealed by structural and functional connectivity. *J Neurosci* 35(21):8191–8200
- Deary IJ (2012) Intelligence. *Ann Rev Psychol* 63:453–482. doi:10.1146/annurev-psych-120710-100353
- Deary IJ, Penke L, Johnson W (2010) The neuroscience of human intelligence differences. *Nat Rev Neurosci* 11(3):201–211. doi:10.1038/nrn2793
- 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. doi:10.1016/j.bandc.2014.08.006
- Faul F, Erdfelder E, Buchner A, Lang A-G (2009) Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 41(4):1149–1160
- Ganjavi H, Lewis JD, Bellec P, MacDonald PA, Waber DP, Evans AC, Brain Development Cooperative, G (2011) Negative associations between corpus callosum midsagittal area and IQ in a representative sample of healthy children and adolescents. *PLoS One* 6(5):e19698. doi:10.1371/journal.pone.0019698
- Giedd JN, Rumsey JM, Castellanos FX, Rajapakse JC, Kaysen D, Vaituzis AC, Rapoport JL (1996) A quantitative MRI study of the corpus callosum in children and adolescents. *Dev Brain Res* 91(2):274–280
- Haier RJ, Jung RE, Yeo RA, Head K, Alkire MT (2004) Structural brain variation and general intelligence. *Neuroimage* 23(1):425–433
- Hansen TI, Brezova V, Eikenes L, Haberg A, Vangberg TR (2015) How does the accuracy of intracranial volume measurements affect normalized brain volumes? Sample size estimates based on 966 subjects from the HUNT MRI cohort. *AJNR Am J Neuroradiol* 36(8):1450–1456. doi:10.3174/ajnr.A4299
- Hofer SM, Flaherty BP, Hoffman L (2006) Cross-sectional analysis of time-dependent data: mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohort samples. *Multivar Behav Res* 41(2):165–187
- Hulshoff Pol HE, Schnack HG, Posthuma D, Mandl RC, Baare WF, van Oel C, Kahn RS (2006) Genetic contributions to human brain morphology and intelligence. *J Neurosci* 26(40):10235–10242. doi:10.1523/JNEUROSCI.1312-06.2006
- Huster RJ, Westerhausen R, Herrmann C (2011) Sex differences in cognitive control are associated with midcingulate and callosal morphology. *Brain Struct Funct* 215(3–4):225–235
- Hutchinson AD, Mathias JL, Jacobson BL, Ruzic L, Bond AN, Banich MT (2009) Relationship between intelligence and the size and composition of the corpus callosum. *Exp Brain Res* 192(3):455–464. doi:10.1007/s00221-008-1604-5

- Jäncke L, Preis S, Steinmetz H (1999) The relation between forebrain volume and midsagittal size of the corpus callosum in children. *Neuroreport* 10(14):2981–2985
- Jung RE, Haier RJ (2007) The parieto-frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci* 30(2):135–154. doi:10.1017/S0140525X07001185 **discussion 154–187**
- Karama S, Colom R, Johnson W, Deary IJ, Haier R, Waber DP, Evans AC (2011) Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6–18. *Neuroimage* 55(4):1443–1453
- Kompus K, Kalpouzos G, Westerhausen R (2011) The size of the anterior corpus callosum correlates with the strength of hemispheric encoding-retrieval asymmetry in the ventrolateral prefrontal cortex. *Brain Res* 1419:61–67
- Krogsrud SK, Fjell AM, Tamnes CK, Grydeland H, Mork L, Due-Tønnessen P, Walhovd KB (2016) Changes in white matter microstructure in the developing brain—A longitudinal diffusion tensor imaging study of children from 4 to 11 years of age. *Neuroimage* 124(1):473–486. doi:10.1016/j.neuroimage.2015.09.017
- Luders E, Narr KL, Thompson PM, Rex DE, Woods RP, Deluca H, Toga AW (2006) Gender effects on cortical thickness and the influence of scaling. *Hum Brain Mapp* 27(4):314–324. doi:10.1002/hbm.20187
- Luders E, Narr KL, Bilder RM, Thompson PM, Szeszko PR, Hamilton L, Toga AW (2007) Positive correlations between corpus callosum thickness and intelligence. *Neuroimage* 37(4):1457–1464. doi:10.1016/j.neuroimage.2007.06.028
- Luders E, Narr KL, Thompson PM, Toga AW (2009) Neuroanatomical correlates of intelligence. *Intelligence* 37(2):156–163. doi:10.1016/j.intell.2008.07.002
- Luders E, Cherbuin N, Thompson PM, Gutman B, Anstey KJ, Sachdev P, Toga AW (2010a) When more is less: associations between corpus callosum size and handedness lateralization. *Neuroimage* 52(1):43–49
- Luders E, Thompson PM, Toga AW (2010b) The development of the corpus callosum in the healthy human brain. *J Neurosci* 30(33):10985–10990. doi:10.1523/JNEUROSCI.5122-09.2010
- Luders E, Thompson PM, Narr KL, Zamanyan A, Chou YY, Gutman B, Toga AW (2011) The link between callosal thickness and intelligence in healthy children and adolescents. *Neuroimage* 54(3):1823–1830. doi:10.1016/j.neuroimage.2010.09.083
- Mamelak AN, Barbaro NM, Walker JA, Laxer KD (1993) Corpus callosotomy: a quantitative study of the extent of resection, seizure control, and neuropsychological outcome. *J Neurosurg* 79(5):688–695
- McDaniel MA (2005) Big-brained people are smarter: a meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence* 33(4):337–346
- Men W, Falk D, Sun T, Chen W, Li J, Yin D, Fan M (2014) The corpus callosum of Albert Einstein's brain: another clue to his high intelligence? *Brain* 137(4):e268
- Menary K, Collins PF, Porter JN, Muetzel R, Olson EA, Kumar V, Luciana M (2013) Associations between cortical thickness and general intelligence in children, adolescents and young adults. *Intelligence* 41(5):597–606
- Narr KL, Woods RP, Thompson PM, Szeszko P, Robinson D, Dimcheva T, Bilder RM (2007) Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cereb Cortex* 17(9):2163–2171
- Navas-Sanchez FJ, Aleman-Gomez Y, Sanchez-Gonzalez J, Guzman-De-Villoria JA, Franco C, Robles O, Desco M (2014) White matter microstructure correlates of mathematical giftedness and intelligence quotient. *Hum Brain Mapp* 35(6):2619–2631. doi:10.1002/hbm.22355
- Neisser U (1997) Rising scores on intelligence tests: test scores are certainly going up all over the world, but whether intelligence itself has risen remains controversial. *Am Sci* 85(5):440–447
- Nosarti C, Rushe TM, Woodruff PW, Stewart AL, Rifkin L, Murray RM (2004) Corpus callosum size and very preterm birth: relationship to neuropsychological outcome. *Brain* 127(9):2080–2089
- Oguni H, Olivier A, Andermann F, Comair J (1991) Anterior callosotomy in the treatment of medically intractable epilepsies: a study of 43 patients with a mean follow-up of 39 months. *Ann Neurol* 30(3):357–364
- Penke L, Maniega SM, Bastin M, Hernandez MV, Murray C, Royle N, Deary I (2012) Brain white matter tract integrity as a neural foundation for general intelligence. *Mol Psychiatry* 17(10):1026–1030
- Peterson BS, Feineigle PA, Staib LH, Gore JC (2001) Automated measurement of latent morphological features in the human corpus callosum. *Hum Brain Mapp* 12(4):232–245
- Pietschnig J, Penke L, Wicherts JM, Zeiler M, Voracek M (2015) Meta-analysis of associations between human brain volume and intelligence differences: how strong are they and what do they mean? *Neurosci Biobehav Rev* 57:411–432
- Pujol J, Vendrell P, Junqué C, Martí-Vilalta JL, Capdevila A (1993) When does human brain development end? Evidence of corpus callosum growth up to adulthood. *Ann Neurol* 34(1):71–75
- Putnam MC, Steven MS, Doron KW, Riggall AC, Gazzaniga MS (2010) Cortical projection topography of the human splenium: hemispheric asymmetry and individual differences. *J Cogn Neurosci* 22(8):1662–1669
- Salthouse TA (2011) Neuroanatomical substrates of age-related cognitive decline. *Psychol Bull* 137(5):753
- Schmahmann JD, Pandya DN (2006) Fiber pathways of the brain. Oxford University Press, Oxford
- Schulte T, Müller-Oehring E, Salo R, Pfefferbaum A, Sullivan E (2006) Callosal involvement in a lateralized stroop task in alcoholic and healthy subjects. *Neuropsychology* 20(6):727
- Shaw P (2007) Intelligence and the developing human brain. *Bioessays* 29(10):962–973
- Shaw P, Greenstein D, Lerch J, Clasen L, Lenroot R, Gogtay N, Giedd J (2006) Intellectual ability and cortical development in children and adolescents. *Nature* 440(7084):676–679
- Spitzka EA (1907) A study of the brains of six eminent scientists and scholars belonging to the American Anthropometric Society, together with a description of the skull of Professor E. D. Cope. *Trans Am Philos Soc* 21(4):175–308
- Strauss E, Wada J, Hunter M (1994) Callosal morphology and performance on intelligence tests. *J Clin Exp Neuropsychol* 16(1):79–83. doi:10.1080/01688639408402618
- Tamnes CK, Østby Y, Walhovd KB, Westlye LT, Due-Tønnessen P, Fjell AM (2010) Intellectual abilities and white matter microstructure in development: a diffusion tensor imaging study. *Hum Brain Mapp* 31(10):1609–1625
- Tamnes CK, Walhovd KB, Dale AM, Østby Y, Grydeland H, Richardson G, Due-Tønnessen P (2013) Brain development and aging: overlapping and unique patterns of change. *Neuroimage* 68:63–74
- Tang C, Eaves E, Ng J, Carpenter D, Mai X, Schroeder D, Haier R (2010) Brain networks for working memory and factors of intelligence assessed in males and females with fMRI and DTI. *Intelligence* 38(3):293–303
- Tanriverdi T, Olivier A, Poulin N, Andermann F, Dubeau F (2009) Long-term seizure outcome after corpus callosotomy: a retrospective analysis of 95 patients: clinical article. *J Neurosurg* 110(2):332–342
- Verbeke G, Molenberghs G (2009) Linear mixed models for longitudinal data. Springer Science & Business Media, Heidelberg

- Wechsler D (1999) Wechsler Abbreviated Scale of Intelligence (WASI). The Psychological Corporation, San Antonio
- Welcome SE, Chiarello C (2008) How dynamic is interhemispheric interaction? Effects of task switching on the across-hemisphere advantage. *Brain Cogn* 67(1):69–75
- Westerhausen R, Kreuder F, Sequeira SDS, Walter C, Woerner W, Wittling RA, Wittling W (2004) Effects of handedness and gender on macro-and microstructure of the corpus callosum and its subregions: a combined high-resolution and diffusion-tensor MRI study. *Cogn Brain Res* 21(3):418–426
- Westerhausen R, Gruner R, Specht K, Hugdahl K (2009) Functional relevance of interindividual differences in temporal lobe callosal pathways: a DTI tractography study. *Cereb Cortex* 19(6):1322–1329. doi:[10.1093/cercor/bhn173](https://doi.org/10.1093/cercor/bhn173)
- Westerhausen R, Luders E, Specht K, Ofte SH, Toga AW, Thompson PM, Hugdahl K (2011) Structural and functional reorganization of the corpus callosum between the age of 6 and 8 years. *Cereb Cortex* 21(5):1012–1017
- Westerhausen R, Fjell AM, Krogsrud SK, Rohani DA, Skranes JS, Haberg AK, Walhovd KB (2016) Selective increase in posterior corpus callosum thickness between the age of 4 and 11 years. *Neuroimage* 139:17–25. doi:[10.1016/j.neuroimage.2016.06.008](https://doi.org/10.1016/j.neuroimage.2016.06.008)