Neuroscience and Biobehavioral Reviews 57 (2015) 411–432

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean?☆

Jakob Pietschnig[a,b,c,*], Lars Penke[d], Jelte M. Wicherts[e], Michael Zeiler[b], Martin Voracek[b,d,f]

[a] Department of Applied Psychology—Health, Development, Enhancement and Intervention, Faculty of Psychology, University of Vienna, Vienna, Austria
[b] Department of Basic Psychological Research and Research Methods, Faculty of Psychology, University of Vienna, Vienna, Austria
[c] Department of Psychology, School of Science and Technology, Middlesex University Dubai, Dubai, United Arab Emirates
[d] Georg Elias Müller Department of Psychology, Georg August University Göttingen, Göttingen, Germany
[e] Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands
[f] Department of Psychology, University of Zürich, Zürich, Switzerland

ARTICLE INFO

Article history:
Received 5 May 2015
Received in revised form 7 September 2015
Accepted 30 September 2015
Available online 9 October 2015

Keywords:
Intelligence
In vivo brain volume
Meta-analysis
Meta-regression
Reporting bias

ABSTRACT

Positive associations between human intelligence and brain size have been suspected for more than 150 years. Nowadays, modern non-invasive measures of in vivo brain volume (Magnetic Resonance Imaging) make it possible to reliably assess associations with IQ. By means of a systematic review of published studies and unpublished results obtained by personal communications with researchers, we identified 88 studies examining effect sizes of 148 healthy and clinical mixed-sex samples (>8000 individuals). Our results showed significant positive associations of brain volume and IQ ($r = .24$, $R^2 = .06$) that generalize over age (children vs. adults), IQ domain (full-scale, performance, and verbal IQ), and sex. Application of a number of methods for detection of publication bias indicates that strong and positive correlation coefficients have been reported frequently in the literature whilst small and non-significant associations appear to have been often omitted from reports. We show that the strength of the positive association of brain volume and IQ has been overestimated in the literature, but remains robust even when accounting for different types of dissemination bias, although reported effects have been declining over time. While it is tempting to interpret this association in the context of human cognitive evolution and species differences in brain size and cognitive ability, we show that it is not warranted to interpret brain size as an isomorphic proxy of human intelligence differences.

© 2015 Elsevier Ltd. All rights reserved.

Contents

1. Introduction 412
2. Surrogate measures of brain volume and intelligence 412
3. Meta-analysis of the association between in vivo brain volume and IQ 412
   3.1. Methods 420
      3.1.1. Literature search 420
      3.1.2. Inclusion criteria 420
      3.1.3. Coding 420
      3.1.4. Data analysis 420
      3.1.5. Final sample 423

☆ The authors would like to thank T. Haubner and S. Pavlovic for their help regarding literature acquisition.

* Corresponding author at: Department for Applied Psychology: Health, Development, Enhancement and Intervention, Faculty of Psychology, University of Vienna, Liebiggasse 5, A-1010 Vienna, Austria. Tel.:+43 1 4277 47237.
E-mail address: jakob.pietschnig@univie.ac.at (J. Pietschnig).

http://dx.doi.org/10.1016/j.neubiorev.2015.09.017
0149-7634/© 2015 Elsevier Ltd. All rights reserved.

3.2. Results 423
3.2.1. Brain volume and IQ 423
3.2.2. Subgroup analysis 424
3.2.3. Meta-regression 424
3.2.4. Publication bias 424
3.3. Discussion of the meta-analysis 426
4. Why is brain size associated with intelligence? 427
4.1. Species differences in brain size and cognitive capability 427
4.2. How can the brain size-IQ association be interpreted? 428
5. Conclusion 428
Appendix A. Supplementary data 429
References 429

1. Introduction

Associations between brain size and intelligence have been subject to investigation for more than a century. In 1836, Friedrich Tiedemann, a German anatomist and physiologist wrote: "There is undoubtedly a very close connexion between the absolute size of the brain and the intellectual powers and functions of the mind" (Tiedemann, 1836, p. 502). Thereafter, this assertion has been entertained by influential minds (e.g., Darwin, 1871; Lombroso, 1864; Broca, 1861, as cited in Rushton and Ankney, 2009) until today. It is extensively reflected in the literature (Deary et al., 2010; Jensen, 1982; McDaniel, 2005; Rushton and Ankney, 2009; Van Valen, 1974) and indeed in lay psychology and the common language, as in the proverbial "big brained" as a synonym for being smart. However, this alleged association has also been subjected to intense debate and controversy (e.g., Deary et al., 2010; Gould, 1981; Jensen, 1982) about its meaning and strength. In this article we will review the evidence on the strength of the linear association between brain size and measures of intelligence, with a particular focus on the most comprehensive and detailed meta-analysis of the relationship between human in vivo brain volume and IQ. We will then critically discuss how this association can be interpreted and how it relates to brain size differences between species.

2. Surrogate measures of brain volume and intelligence

Even though an association between brain volume and intelligence had been hypothesized early on, for long there was a lack of good in vivo measures of brain volume. As a first attempt to quantify the association between brain volume and intelligence, Galton (1888) used linear external head measures (height, breadth, depth) as a proxy for brain size and achievements at universities as a measure for cognitive abilities. The introduction of intelligence tests allowed assessment of cognitive abilities by means of standardized measures, but investigations still had to rely on crude markers of brain volume (e.g., head circumference; Murdoch and Sullivan, 1923). Such external measures have later been criticized as yielding inaccurate estimates of inner skull capacity (intracranial volume, ICV; Simmons, 1942). However, recent studies that compared head circumference with ICV assessed precisely in vivo using Magnetic Resonance Imaging (MRI) in large samples showed that head circumference provides a reasonable estimate of ICV, with correlations of .62 for men and .56 for women (Booth et al., 2015; Wolf et al., 2003). Head circumference is actually a commonly used surrogate for brain volume measurement, e.g., in epidemiological cohort studies. However, expectably the correlation between head circumference and IQ is weaker than the correlation between ICV and IQ (Booth et al., 2015; MacLullich et al., 2002), and even though it tends to be positive, it is not as reliable as some reviews suggest (Rushton and Ankney, 1996, 2000, 2009).

ICV itself is also often used as a surrogate for brain volume, especially in anthropological studies. But even ICV does not reflect brain volume perfectly, as atrophy results in brain shrinkage relative to the ICV, even in normal ageing (Royle et al., 2013). A thickening of the inner skull in response to atrophy can also bias ICV when it is used as an estimate of early-life or pre-morbid brain volume (Finby and Kraft, 1972; May et al., 2011). However, the advent of MRI made it possible to reliably and accurately measure in vivo brain volume non-invasively. These studies now make up most of the available evidence on an association between brain volume and IQ.

3. Meta-analysis of the association between in vivo brain volume and IQ

In the first study using MRI, Willerman et al. (1991) reported an association of $r = .51$ between brain volume and full-scale IQ of the revised Wechsler Adult Intelligence Scale (WAIS-R; Wechsler, 1981) in a sample of 40 healthy college students. Subsequently, as the use of MRI became more common, several replication attempts of this strong effect of MRI-based studies (a previous review of head circumference and IQ had found overall effects of only about $r = .30$; Van Valen, 1974) were published, but varied considerably in size and in some cases even in direction (e.g., Collinson et al., 2003). Several narrative reviews on associations of brain size measures and intelligence have been published since the mid-1990s (Gignac et al., 2003; Miller and Penke, 2007; Rushton and Ankney, 1996, 2000, 2009; Vernon et al., 2000), concluding that there was strong evidence for a positive relationship between these two variables. However, narrative reviews are limited in their capability to determine the strength of such associations or influences of moderating variables and are vulnerable to bias (e.g., Borenstein et al., 2009, pp. 301–302). Further insights on this topic may be gained from systematic quantitative reviews (i.e., meta-analyses), which provide a large bandwidth of tools to assess overall strength of effects, group differences, influences of moderating variables, and potential bias.

As a first step to shed light on the strength of this association in a subset of the general population (namely healthy individuals), a meta-analysis of in vivo brain volume and full-scale IQ was published in 2005 (McDaniel, 2005). For 24 studies comprising 37 samples of healthy men and women (>1500 individuals), a moderate significant association of brain volume and intelligence was reported ($r = .29$). McDaniel (2005) found stronger effects for women than men ($r = .36$ and $r = .30$) but similar correlations among adults and children ($r = .30$ and $r = .28$) although there were no formal significance tests reported. This paper attracted a lot of attention in the scientific community, quickly becoming one of the most highly cited articles of *Intelligence*, the leading journal in the field of intelligence research.

However, McDaniel (2005) directed attention towards some concerns about validity of results of his meta-analysis as he pointed out that reporting practice in the literature gave reason to surmise

Table 1

Details of included studies.

Author	year	Review coverage						Participants	Mean age	Sex	Reporting	IQ domain	Measure	Type of test	*n*	*r*
		R1	R2	V	G	M	R3									
Yeo et al.	1987	x	–	x	–	–	–	Clinical	38.40	Mixed	Reported	FSIQ	CT	WAIS	41	.07
Yeo et al.	1987	–	–	–	–	–	–	Clinical	38.40	Mixed	Reported	Performance	CT	WAIS	41	.06
Yeo et al.	1987	–	–	–	–	–	–	Clinical	38.40	Mixed	Reported	Verbal	CT	WAIS	41	.12
Willerman et al.	1991	x	x	x	x	x	x	Healthy	18.90	Women	Reported	FSIQ	MRI	WAIS-R	20	.33
Willerman et al.	1991	x	x	x	x	x	x	Healthy	18.90	Men	Reported	FSIQ	MRI	WAIS-R	20	.51*
Andreasen et al.	1993	x	–	x	x	x	x	Healthy	38.00	Women	Reported	FSIQ	MRI	WAIS-R	30	.44*
Andreasen et al.	1993	–	–	–	–	–	–	Healthy	38.00	Women	Reported	Performance	MRI	WAIS-R	30	.30
Andreasen et al.	1993	–	–	–	–	–	–	Healthy	38.00	Women	Reported	Verbal	MRI	WAIS-R	30	.43*
Andreasen et al.	1993	x	–	x	x	x	x	Healthy	38.00	Men	Reported	FSIQ	MRI	WAIS-R	37	.40*
Andreasen et al.	1993	–	–	–	–	–	–	Healthy	38.00	Men	Reported	Performance	MRI	WAIS-R	37	.43**
Andreasen et al.	1993	–	–	–	–	–	–	Healthy	38.00	Men	Reported	Verbal	MRI	WAIS-R	37	.33*
Raz et al.	1993	x	–	x	x	x	x	Healthy	43.80	Mixed	Reported	Fluid	MRI	CFIT	29	.43*
Raz et al.	1993	–	–	–	–	–	–	Healthy	43.80	Mixed	Reported	Verbal	MRI	V3	29	.10
Castellanos et al.	1994	–	–	x	–	x	x	Healthy	12.10	Men	Reported	FSIQ	MRI	WISC-R	46	.33*
Egan et al.	1994	x	–	x	x	x	x	Healthy	22.50	Men	Reported	Performance	MRI	WAIS-R	40	.24
Egan et al.	1994	–	–	–	–	–	–	Healthy	22.50	Men	Reported	Verbal	MRI	WAIS-R	40	.24
Wickett et al.	1994	x	–	x	x	x	x	Healthy	25.00	Women	Reported	FSIQ	MRI	MAB	40	.40*
Wickett et al.	1994	–	–	–	–	–	–	Healthy	25.00	Women	Reported	Performance	MRI	MAB	40	.28
Wickett et al.	1994	–	–	–	–	–	–	Healthy	25.00	Women	Reported	Verbal	MRI	MAB	40	.44**
Harvey et al.	1994	x	x	–	–	–	x	Clinical	35.60	Mixed	Reported	Verbal	MRI	NART	26	.38
Harvey et al.	1994	x	x	–	–	–	x	Clinical	31.10	Mixed	Reported	Verbal	MRI	NART	48	.24
Harvey et al.	1994	x	x	x	–	–	x	Healthy	31.60	Mixed	Reported	Verbal	MRI	NART	34	.69***
Jones et al.	1994	–	–	x	–	–	x	Healthy	31.70	Mixed	Reported	Verbal	CT	NART and WAIS-R	67	.30*
Egan et al.	1995	x	–	x	–	x	x	Healthy	22.50	Men	Reported	FSIQ	MRI	WAIS-R	40	.31
Egan et al.	1995	–	–	–	–	–	–	Healthy	22.50	Men	Reported	Performance	MRI	WAIS-R	40	.22
Egan et al.	1995	–	–	–	–	–	–	Healthy	22.50	Men	Reported	Verbal	MRI	WAIS-R	40	.21
Kareken et al.	1995	–	–	x	–	x	x	Healthy	27.66	Mixed	PC	FSIQ	MRI	WAIS-R	68	.30*
Haier et al.	1995	x	–	–	–	–	x	Clinical	26.52	Mixed	Reported	FSIQ	MRI	WAIS-R	26	.65***
Raz et al.	1995	–	–	–	–	–	–	Clinical	35.20	Mixed	Reported	FSIQ	MRI	BCS and WPPSI-R	11	−.24
Raz et al.	1995	–	–	–	–	–	–	Clinical	35.20	Mixed	Reported	Verbal	MRI	BCS	11	.90***
Bigler	1995	–	–	–	–	–	–	Clinical	29.54	Mixed	Reported	FSIQ	MRI	WAIS-R	72	−.03
Reiss et al.	1995	–	–	–	–	–	–	Clinical	10.80	Mixed	Reported	FSIQ	MRI	BS, SBIS and WISC-R	51	.25
Reiss et al.[a]	1995	–	–	–	–	–	–	Healthy	11.28	Mixed	PC	FSIQ	MRI	BS, SBIS and WISC-R	87	.00
Reiss et al.	1996	–	–	x	x	x	x	Healthy	10.60	Women	PC	FSIQ	MRI	Not reported	57	.37**
Reiss et al.	1996	–	–	x	x	x	x	Healthy	10.10	Men	PC	FSIQ	MRI	Not reported	12	.52
Blatter et al.	1997	–	–	–	–	–	–	Clinical	Not reported	Not reported	Reported	Performance	MRI	WAIS-R	21	.47*
Blatter et al.	1997	–	–	–	–	–	–	Clinical	Not reported	Not reported	Reported	Verbal	MRI	WAIS-R	22	.57**
Paradiso et al.	1997	–	–	–	–	–	–	Healthy	24.80	Mixed	Reported	FSIQ	MRI	WAIS-R	62	.38**
Paradiso et al.	1997	–	–	–	–	–	–	Healthy	24.80	Mixed	Reported	Performance	MRI	WAIS-R	62	.32*
Paradiso et al.	1997	–	–	–	–	–	–	Healthy	24.80	Mixed	Reported	Verbal	MRI	WAIS-R	62	.27*
Mori et al.	1997	–	–	–	–	–	–	Clinical	70.20	Mixed	Reported	FSIQ	MRI	WAIS-R	60	.40***
Mori et al.	1997	–	–	–	–	–	–	Clinical	70.20	Mixed	Reported	Performance	MRI	WAIS-R	60	.37**
Mori et al.	1997	–	–	–	–	–	–	Clinical	70.20	Mixed	Reported	Verbal	MRI	WAIS-R	60	.37**
Flashman et al.	1998	–	x	–	x	x	x	Healthy	27.00	Mixed	Reported	FSIQ	MRI	WAIS-R	90	.25*
Flashman et al.	1998	–	–	–	–	–	–	Healthy	27.00	Mixed	Reported	Performance	MRI	WAIS-R	90	.26*
Flashman et al.	1998	–	–	–	–	–	–	Healthy	27.00	Mixed	Reported	Verbal	MRI	WAIS-R	90	.16
Gur et al.	1999	–	x	–	x	x	x	Healthy	25.00	Women	Reported	FSIQ	MRI	VLT, JLOT, and WAIS-R	40	.40**
Gur et al.	1999	–	–	–	x	–	–	Healthy	25.00	Women	Reported	Verbal	MRI	WAIS-R and CVLT	40	.40**

Table 1 (*Continued*)

Author	year	Review coverage						Participants	Mean age	Sex	Reporting	IQ domain	Measure	Type of test	*n*	*r*
		R1	R2	V	G	M	R3									
Gur et al.	1999	–	x	–	x	x	x	Healthy	27.00	Men	Reported	FSIQ	MRI	VLT, JLOT, and WAIS-R	40	.39*
Gur et al.[a]	1999	–	–	–	x	–	–	Healthy	27.00	Men	PC	Verbal	MRI	WAIS-R and CVLT	40	.00
Tan et al.	1999	–	x	–	x	x	x	Healthy	22.00	Women	Reported	Fluid	MRI	CFIT	54	.62***
Tan et al.	1999	–	x	–	x	x	x	Healthy	22.00	Men	Reported	Fluid	MRI	CFIT	49	.28
Warwick et al.[a]	1999	–	–	–	–	–	–	Clinical	21.60	Women	PC	Verbal	MRI	Quick	11	.00
Warwick et al.	1999	–	–	–	–	–	–	Clinical	21.55	Women	PC	Verbal	MRI	Quick	24	.53**
Warwick et al.[a]	1999	–	–	–	–	–	–	Healthy	21.50	Women	PC	Verbal	MRI	Quick	13	.00
Warwick et al.[a]	1999	–	–	–	–	–	–	Clinical	21.80	Men	PC	Verbal	MRI	Quick	10	.00
Warwick et al.[a]	1999	–	–	–	–	–	–	Clinical	21.80	Men	PC	Verbal	MRI	Quick	10	.00
Warwick et al.	1999	–	–	–	–	–	–	Clinical	21.63	Men	Reported	Verbal	MRI	Quick	45	.31*
Warwick et al.[a]	1999	–	–	–	–	–	–	Healthy	21.50	Men	Reported	Verbal	MRI	Quick	25	.00
Leonard et al.[a]	1999	–	–	–	–	–	–	Clinical	43.00	Men	PC	Performance	MRI	WAIS-R	37	.00
Leonard et al.[a]	1999	–	–	–	–	–	–	Clinical	43.00	Men	PC	Verbal	MRI	WAIS-R	37	.00
Leonard et al.[a]	1999	–	–	–	–	–	–	Healthy	42.00	Men	PC	Performance	MRI	WAIS-R	33	.00
Leonard et al.[a]	1999	–	–	–	–	–	–	Healthy	42.00	Men	PC	Verbal	MRI	WAIS-R	33	.00
Pennington et al.	2000	–	–	–	x	x	x	Healthy	19.06	Mixed	Reported	FSIQ	MRI	WISC-R and WAIS-R	36	.31
Pennington et al.	2000	–	–	–	x	–	x	Healthy	16.97	Mixed	Reported	FSIQ	MRI	WISC-R and WAIS-R	96	.42***
Wickett et al.	2000	–	x	x	x	x	x	Healthy	24.97	Men	Reported	FSIQ	MRI	MAB	68	.35**
Wickett et al.	2000	–	–	–	–	–	–	Healthy	24.97	Men	Reported	Performance	MRI	MAB	68	.31**
Wickett et al.	2000	–	–	–	–	–	–	Healthy	24.97	Men	Reported	Verbal	MRI	MAB	68	.33**
Garde et al.	2000	–	–	–	–	x	–	Healthy	80.70	Women	PC	FSIQ	MRI	WAIS	22	.22
Garde et al.	2000	–	–	–	–	x	–	Healthy	80.70	Men	PC	FSIQ	MRI	WAIS	46	.07
Schoenemann et al.	2000	–	–	–	x	x	x	Healthy	23.20	Women	PC	Fluid	MRI	RPM	72	.21
Schoenemann et al.	2000	–	–	–	–	x	x	Healthy	23.20	Women	Reported	Verbal	MRI	MAB	36	.12
Lawson et al.	2000	–	–	–	–	–	–	Clinical	Not reported	Not reported	Reported	FSIQ	MRI	WISC-III, WPPSI-R, DAS, SBIS, and GMDS	47	.43**
Kumra et al.[a]	2000	–	–	–	–	–	–	Clinical	12.30	Mixed	PC	FSIQ	MRI	WISC-III, WISC-R, and WAIS	27	.00
Kumra et al.[a]	2000	–	–	–	–	–	–	Clinical	14.40	Mixed	PC	FSIQ	MRI	WISC-III, WISC-R, and WAIS	44	.00
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	FSIQ	MRI	WISC-III	11	−.03
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	Performance	MRI	WISC-III	11	−.18
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	Verbal	MRI	WISC-III	11	−.04
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	FSIQ	MRI	WISC-III	8	.55
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	Performance	MRI	WISC-III	8	.35
Isaacs et al.	2000	–	–	–	–	–	–	Healthy	7.75	Mixed	PC	Verbal	MRI	WISC-III	8	.57
Castellanos et al.	2001	–	–	–	–	–	–	Clinical	9.70	Women	Reported	FSIQ	MRI	WISC-R and WISC-III	40	.36*
Coffey et al.	2001	–	–	–	–	–	–	Healthy	74.85	Mixed	Reported	Performance	MRI	WAIS-R	318	.06
Coffey et al.	2001	–	–	–	–	–	–	Healthy	74.85	Mixed	Reported	Verbal	MRI	Verbal fluency task	319	−.06
MacLullich et al.	2002	–	–	–	–	x	x	Healthy	67.80	Men	Reported	Fluid	MRI	SPM	97	.39***
MacLullich et al.	2002	–	–	–	–	–	–	Healthy	67.80	Men	Reported	Verbal	MRI	NART	97	.30**
Aylward et al.	2002	–	–	–	–	x	–	Healthy	Not reported	Men	PC	FSIQ	MRI	Not reported	46	−.13
Aylward et al.	2002	–	–	–	–	–	–	Clinical	18.80	Mixed	Reported	FSIQ	MRI	Not reported	67	.10
Aylward et al.	2002	–	–	–	–	–	–	Clinical	18.80	Mixed	Reported	Performance	MRI	Not reported	67	.10

Aylward et al.	2002	–	–	–	–	–	–	Clinical	18.80	Mixed	Reported	Verbal	MRI	Not reported	67	.08
Aylward et al.	2002	–	–	–	–	–	x	Healthy	18.90	Mixed	Reported	Performance	MRI	Not reported	83	.09
Aylward et al.	2002	–	–	–	–	–	x	Healthy	18.90	Mixed	Reported	Verbal	MRI	Not reported	83	−.01
Aylward et al.	2002	–	–	–	–	x	–	Healthy	Not reported	Not reported	PC	FSIQ	MRI	Not reported	30	.08
Nosarti et al.	2002	–	–	–	–	x	–	Healthy	14.90	Mixed	PC	FSIQ	MRI	Not reported	42	.37
Shapleske et al.	2002	–	–	–	–	x	–	Healthy	33.30	Men	PC	FSIQ	MRI	Not reported	3	−.86
Shapleske et al.	2002	–	–	–	–	x	–	Healthy	33.30	Men	PC	FSIQ	MRI	Not reported	23	.13
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Women	PC	FSIQ	Not reported	Not reported	8	.46
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Men	PC	FSIQ	Not reported	Not reported	7	.17
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Women	PC	FSIQ	Not reported	Not reported	7	−.67
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Men	PC	FSIQ	Not reported	Not reported	7	.67
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Women	PC	FSIQ	Not reported	Not reported	39	.34*
Giedd	2005	–	–	–	–	x	–	Healthy	Not reported	Men	PC	FSIQ	Not reported	Not reported	63	.27*
Kesler et al.	2003	–	–	–	–	–	–	Clinical	25.80	Mixed	Reported	FSIQ	MRI	WAIS-R	25	.47*
Kesler et al.	2003	–	–	–	–	–	–	Clinical	25.80	Mixed	Reported	Verbal	MRI	WAIS-R	25	.57**
Yurgelun-Todd et al.	2003	–	–	–	–	–	–	Healthy	14.60	Women	Reported	FSIQ	MRI	SILT	24	.20
Yurgelun-Todd et al.	2003	–	–	–	–	–	–	Healthy	14.60	Women	Reported	Verbal	MRI	SILT	24	.17
Yurgelun-Todd et al.	2003	–	–	–	–	–	–	Healthy	14.50	Men	Reported	FSIQ	MRI	SILT	13	.26
Yurgelun-Todd et al.	2003	–	–	–	–	–	–	Healthy	14.50	Men	Reported	Verbal	MRI	SILT	13	.19
Collinson et al.	2003	–	–	–	–	–	–	Clinical	16.80	Mixed	PC	FSIQ	MRI	WISC-R and WAIS-R	32	−.27
Collinson et al.	2003	–	–	–	–	–	–	Clinical	16.80	Mixed	PC	Performance	MRI	WISC-R and WAIS-R	32	−.19
Collinson et al.	2003	–	–	–	–	–	–	Clinical	16.80	Mixed	PC	Verbal	MRI	WISC-R and WAIS-R	32	−.28
Collinson et al.	2003	–	–	–	–	–	–	Healthy	16.40	Mixed	PC	FSIQ	MRI	WISC-R and WAIS-R	22	−.13
Collinson et al.	2003	–	–	–	–	–	–	Healthy	16.40	Mixed	PC	Performance	MRI	WISC-R and WAIS-R	22	−.17
Collinson et al.	2003	–	–	–	–	–	–	Healthy	16.40	Mixed	PC	Verbal	MRI	WISC-R and WAIS-R	22	−.09
Frangou et al.	2004	–	–	–	–	x	–	Healthy	15.05	Mixed	Reported	FSIQ	MRI	WISC-III and WAIS-III	40	.41**
Ivanovic et al.	2004a	–	–	–	–	–	–	Healthy	18.00	Women	Reported	Performance	MRI	WAIS-R	49	.38**
Ivanovic et al.	2004a	–	–	–	–	–	–	Healthy	18.00	Women	Reported	Verbal	MRI	WAIS-R	49	.33*
Ivanovic et al.	2004a	–	–	–	–	–	–	Healthy	18.00	Men	Reported	Performance	MRI	WAIS-R	47	.52***
Ivanovic et al.	2004a	–	–	–	–	–	–	Healthy	18.00	Men	Reported	Verbal	MRI	WAIS-R	47	.55***
Ivanovic et al.	2004b	–	–	–	–	x	x	Healthy	18.00	Mixed	Reported	FSIQ	MRI	WAIS-R	96	.44***
Toulopoulou et al.	2004	–	–	–	–	–	–	Clinical	42.23	Mixed	Reported	FSIQ	MRI	WAIS-R	201	.28***
Toulopoulou et al.	2004	–	–	–	–	–	–	Clinical	42.23	Mixed	Reported	Verbal	MRI	WAIS-R	201	.28***
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.90	Women	PC	FSIQ	MRI	WISC-R and WISC-III	38	.24
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.90	Women	PC	Performance	MRI	WISC-R and WISC-III	38	.21
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.60	Women	PC	Verbal	MRI	WISC-R and WISC-III	38	.20
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.90	Men	PC	FSIQ	MRI	WISC-R and WISC-III	38	.27
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.90	Men	PC	Performance	MRI	WISC-R and WISC-III	38	.15
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	15.90	Men	PC	Verbal	MRI	WISC-R and WISC-III	38	.33*
Isaacs et al.	2004	–	–	–	–	–	–	Healthy	14.86	Mixed	PC	FSIQ	MRI	WISC-R and WISC-III	16	.49

Table 1 (*Continued*)

Author	year	Review coverage						Participants	Mean age	Sex	Reporting	IQ domain	Measure	Type of test	*n*	*r*
		R1	R2	V	G	M	R3									
Waiter et al.	2004	–	–	–	–	–	–	Clinical	15.40	Men	PC	FSIQ	MRI	WISC-III and WAIS-IV	16	−.06
Waiter et al.	2004	–	–	–	–	–	–	Clinical	15.40	Men	PC	Performance	MRI	WISC-III and WAIS-IV	16	.10
Waiter et al.	2004	–	–	–	–	–	–	Clinical	15.40	Men	PC	Verbal	MRI	WISC-III and WAIS-IV	16	−.17
Waiter et al.	2004	–	–	–	–	–	–	Healthy	15.50	Men	PC	FSIQ	MRI	WISC-III and WAIS-IV	16	.13
Waiter et al.	2004	–	–	–	–	–	–	Healthy	15.50	Men	PC	Performance	MRI	WISC-III and WAIS-IV	16	.23
Waiter et al.	2004	–	–	–	–	–	–	Healthy	15.50	Men	PC	Verbal	MRI	WISC-III and WAIS-IV	16	.20
Rojas et al.	2004	–	–	–	–	–	–	Clinical	30.30	Mixed	PC	FSIQ	MRI	WAIS-R and WAIS-III	15	.07
Rojas et al.	2004	–	–	–	–	–	–	Clinical	30.30	Mixed	PC	Performance	MRI	WAIS-R and WAIS-III	15	.15
Rojas et al.	2004	–	–	–	–	–	–	Clinical	30.30	Mixed	PC	Verbal	MRI	WAIS-R and WAIS-III	15	.30
Rojas et al.	2004	–	–	–	–	–	–	Healthy	43.62	Mixed	PC	FSIQ	MRI	WAIS-R and WAIS-III	17	.31
Rojas et al.	2004	–	–	–	–	–	–	Healthy	43.62	Mixed	PC	Performance	MRI	WAIS-R and WAIS-III	17	.27
Rojas et al.	2004	–	–	–	–	–	–	Healthy	43.62	Mixed	PC	Verbal	MRI	WAIS-R and WAIS-III	17	.19
Thoma et al.	2005	–	–	–	–	–	–	Healthy	23.50	Men	Reported	FSIQ	MRI	APM, COWA, TMT, VKMRT, WAIS-R	19	.27
Antonova et al.	2005	–	–	–	–	–	–	Clinical	40.49	Mixed	PC	Verbal	MRI	WAIS-III	44	.16
Antonova et al.	2005	–	–	–	–	–	–	Healthy	33.72	Mixed	PC	Verbal	MRI	WAIS-III	43	.24
Lodygensky et al.	2005	–	–	–	–	–	–	Clinical	8.58	Mixed	PC	FSIQ	MRI	WISC-R	60	.35**
Lodygensky et al.	2005	–	–	–	–	–	–	Healthy	8.42	Mixed	PC	FSIQ	MRI	WISC-R	21	.46*
Witelson et al.	2006	–	–	–	–	–	x	Clinical	54.60	Women	Reported	Performance	WDM	WAIS	33	.32
Witelson et al.	2006	–	–	–	–	–	x	Clinical	54.60	Women	Reported	Verbal	WDM	WAIS	40	.59***
Witelson et al.	2006	–	–	–	–	–	x	Clinical	58.60	Men	Reported	Performance	WDM	WAIS	31	−.23
Witelson et al.	2006	–	–	–	–	–	x	Clinical	58.60	Men	Reported	Verbal	WDM	WAIS	20	−.27
Witelson et al.	2006	–	–	–	–	–	x	Clinical	58.60	Men	Reported	Verbal	WDM	WAIS	17	.62**
Debanné et al.	2006	–	–	–	–	–	–	Clinical	16.70	Mixed	PC	FSIQ	MRI	WISC-III and WAIS-III	43	.16
Debanné et al.	2006	–	–	–	–	–	–	Healthy	15.10	Mixed	PC	FSIQ	MRI	WISC-III and WAIS-III	41	.16
Staff et al.[b]	2006	–	–	–	–	x	x	Healthy	79.50	Mixed	PC	Fluid	MRI	SPM	102	−.10
Staff et al.	2006	–	–	–	–	–	–	Healthy	79.50	Mixed	PC	Verbal	MRI	NART	102	−.14
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.16	Men	PC	FSIQ	MRI	WISC-III	38	.02
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.16	Men	PC	Performance	MRI	WISC-III	38	−.02
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.16	Men	PC	Verbal	MRI	WISC-III	38	.08
Voelbel et al.	2006	–	–	–	–	–	–	Healthy	10.77	Men	PC	Performance	MRI	WISC-III	13	.06
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.08	Men	PC	FSIQ	MRI	WISC-III	12	−.14
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.08	Men	PC	Performance	MRI	WISC-III	12	−.48
Voelbel et al.	2006	–	–	–	–	–	–	Clinical	10.08	Men	PC	Verbal	MRI	WISC-III	12	.23
Voelbel et al.	2006	–	–	–	–	–	–	Healthy	10.77	Men	PC	Verbal	MRI	WISC-III	13	−.15
Voelbel et al.	2006	–	–	–	–	–	–	Healthy	10.77	Men	PC	FSIQ	MRI	WISC-III	13	−.11
Rojas et al.	2006	–	–	–	–	–	–	Clinical	20.79	Men	PC	FSIQ	MRI	WAIS-III and WISC-III	24	.30
Rojas et al.	2006	–	–	–	–	–	–	Clinical	20.79	Men	PC	Performance	MRI	WAIS-III and WISC-III	24	.31

Rojas et al.	2006	–	–	–	–	–	–	Clinical	20.79	Men	PC	Verbal	MRI	WAIS-III and WISC-III	24	.28
Rojas et al.	2006	–	–	–	–	–	–	Healthy	21.41	Men	PC	FSIQ	MRI	WAIS-III and WISC-III	23	.46*
Rojas et al.	2006	–	–	–	–	–	–	Healthy	21.41	Men	PC	Performance	MRI	WAIS-III and WISC-III	23	.09
Rojas et al.	2006	–	–	–	–	–	–	Healthy	21.41	Men	PC	Verbal	MRI	WAIS-III and WISC-III	23	.55**
Wozniak et al.	2006	–	–	–	–	–	–	Clinical	12.30	Mixed	PC	FSIQ	MRI	WISC-III and WISC-IV	14	.41
Wozniak et al.	2006	–	–	–	–	–	–	Healthy	12.30	Mixed	PC	FSIQ	MRI	WISC-III and WISC-IV	13	.59*
Luders et al.	2007	–	–	–	–	–	–	Healthy	28.48	Mixed	Reported	FSIQ	MRI	WAIS-R	62	.28*
Schumann et al.	2007	–	–	–	–	–	–	Healthy	13.10	Men	Reported	FSIQ	MRI	WASI	22	.41
Schumann et al.	2007	–	–	–	–	–	–	Healthy	13.10	Men	Reported	Performance	MRI	WASI	22	.25
Schumann et al.	2007	–	–	–	–	–	–	Healthy	13.10	Men	Reported	Verbal	MRI	WASI	22	.38
Chiang et al.	2007	–	–	–	–	–	–	Clinical	29.20	Mixed	Reported	Performance	MRI	WAIS	39	.10
Chiang et al.	2007	–	–	–	–	–	–	Clinical	29.20	Mixed	Reported	Verbal	MRI	WAIS	39	−.02
Chiang et al.	2007	–	–	–	–	–	–	Healthy	Not reported	Not reported	Reported	Performance	MRI	WAIS	16	.41
Chiang et al.	2007	–	–	–	–	–	–	Healthy	Not reported	Not reported	Reported	Verbal	MRI	WAIS	16	−.44
Nakamura et al.	2007	–	–	–	–	–	–	Clinical	40.60	Mixed	PC	FSIQ	MRI	WAIS-III	43	.32*
Nakamura et al.	2007	–	–	–	–	–	–	Clinical	40.60	Mixed	PC	Performance	MRI	WAIS-III	44	.34*
Nakamura et al.	2007	–	–	–	–	–	–	Clinical	40.60	Mixed	PC	Verbal	MRI	WAIS-III	44	.26
Nakamura et al.	2007	–	–	–	–	–	–	Healthy	40.80	Mixed	PC	FSIQ	MRI	WAIS-III	44	.38*
Nakamura et al.	2007	–	–	–	–	–	–	Healthy	40.80	Mixed	PC	Performance	MRI	WAIS-III	43	.29
Nakamura et al.	2007	–	–	–	–	–	–	Healthy	40.80	Mixed	PC	Verbal	MRI	WAIS-III	44	.40**
DeBoer et al.	2007	–	–	–	–	–	–	Clinical	10.75	Not reported	PC	FSIQ	MRI	WISC-III and WISC-IV	21	.25
DeBoer et al.	2007	–	–	–	–	–	–	Clinical	10.75	Not reported	PC	Performance	MRI	WISC-III and WISC-IV	21	.38*
DeBoer et al.	2007	–	–	–	–	–	–	Clinical	10.75	Not reported	PC	Verbal	MRI	WISC-III and WISC-IV	21	.30
DeBoer et al.	2007	–	–	–	–	–	–	Healthy	10.50	Not reported	PC	FSIQ	MRI	WISC-III and WISC-IV	20	−.55*
DeBoer et al.	2007	–	–	–	–	–	–	Healthy	10.50	Not reported	PC	Performance	MRI	WISC-III and WISC-IV	20	−.22
DeBoer et al.	2007	–	–	–	–	–	–	Healthy	10.50	Not reported	PC	Verbal	MRI	WISC-III and WISC-IV	20	−.20
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	40.96	Women	PC	FSIQ	MRI	WAIS-R	69	.34**
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	40.90	Women	PC	Performance	MRI	WAIS-R	68	.29*
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	40.90	Women	PC	Verbal	MRI	WAIS-R	68	.43***
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	34.32	Women	PC	FSIQ	MRI	WAIS-R	22	.60**
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	34.32	Women	PC	Performance	MRI	WAIS-R	22	.30
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	34.32	Women	PC	Verbal	MRI	WAIS-R	22	.54**
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	39.64	Men	PC	FSIQ	MRI	WAIS-R	205	.28***
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	39.65	Men	PC	Performance	MRI	WAIS-R	203	.17*
Schottenbauer et al.	2007	–	–	–	–	–	–	Clinical	39.66	Men	PC	Verbal	MRI	WAIS-R	202	.28***
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	37.77	Men	PC	FSIQ	MRI	WAIS-R	35	.33
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	37.77	Men	PC	Performance	MRI	WAIS-R	35	.17
Schottenbauer et al.	2007	–	–	–	–	–	–	Healthy	37.77	Men	PC	Verbal	MRI	WAIS-R	35	.38*
Fine et al.	2007	–	–	–	–	–	–	Healthy	40.10	Mixed	PC	FSIQ	MRI	WASI	44	−.11
Fine et al.	2007	–	–	–	–	–	–	Healthy	10.47	Mixed	PC	FSIQ	MRI	WASI	24	.23

Table 1 (*Continued*)

Author	year	Review coverage						Participants	Mean age	Sex	Reporting	IQ domain	Measure	Type of test	*n*	*r*
		R1	R2	V	G	M	R3									
Amat et al.	2008	–	–	–	–	–	–	Healthy	31.50	Mixed	PC	FSIQ	MRI	WAIS-R	27	−.11
Amat et al.	2008	–	–	–	–	–	–	Healthy	31.50	Mixed	PC	Performance	MRI	WAIS-R	27	.18
Amat et al.	2008	–	–	–	–	–	–	Healthy	31.50	Mixed	PC	Verbal	MRI	WAIS-R	27	−.29
Raz et al.	2008	–	–	–	–	–	–	Clinical	59.75	Mixed	PC	Fluid	MRI	CFIT	32	−.02
Raz et al.	2008	–	–	–	–	–	–	Clinical	59.75	Mixed	PC	Verbal	MRI	V2 and V3	31	.15
Raz et al.	2008	–	–	–	–	–	–	Healthy	51.11	Mixed	PC	Fluid	MRI	CFIT	55	.18
Raz et al.	2008	–	–	–	–	–	–	Healthy	51.11	Mixed	PC	Verbal	MRI	V2 and V3	55	.13
Ebner et al.	2008	–	–	–	–	–	–	Clinical	34.52	Mixed	PC	Verbal	MRI	MWT	44	.15
Ebner et al.	2008	–	–	–	–	–	–	Healthy	32.45	Mixed	PC	Verbal	MRI	MWT	37	−.13
Zeegers et al.	2009	–	–	–	–	–	–	Clinical	3.72	Mixed	Reported	FSIQ	MRI	MSEL	21	.06
Zeegers et al.	2009	–	–	–	–	–	–	Clinical	3.44	Mixed	Reported	FSIQ	MRI	MSEL	10	.73*
Miller et al.	2009	–	–	–	–	–	–	Clinical	16.53	Mixed	Reported	FSIQ	MRI	WJ-III	16	−.30
Miller et al.	2009	–	–	–	–	–	–	Healthy	9.25	Mixed	Reported	FSIQ	MRI	WJ-III	12	.23
Miller et al.	2009	–	–	–	–	–	–	Healthy	12.08	Not reported	Reported	Fluid	MRI	WJ-III	11	−.11
Miller et al.	2009	–	–	–	–	–	–	Healthy	12.08	Not reported	Reported	Verbal	MRI	WJ-III	11	−.65*
Miller et al.	2009	–	–	–	–	–	–	Clinical	9.25	Not reported	Reported	Verbal	MRI	WJ-III	5	.84
Miller et al.	2009	–	–	–	–	–	–	Clinical	16.53	Not reported	Reported	Verbal	MRI	WJ-III	6	.76
Van Leeuwen et al.	2008	–	–	–	–	–	–	Healthy	9.10	Mixed	Reported	Fluid	MRI	SPM	214	.20**
Van Leeuwen et al.	2008	–	–	–	–	–	–	Healthy	9.10	Mixed	Reported	Performance	MRI	WISC-III	214	.28***
Van Leeuwen et al.	2008	–	–	–	–	–	–	Healthy	9.10	Mixed	Reported	Verbal	MRI	WISC-III	214	.33***
Shenkin et al.	2009	–	–	–	–	–	–	Healthy	78.40	Mixed	Reported	FSIQ	MRI	MHT, SPM, and WMS	99	.21*
Shenkin et al.	2009	–	–	–	–	–	–	Healthy	78.40	Mixed	Reported	Verbal	MRI	CWAT	107	.13
Qiu et al.	2009	–	–	–	–	–	–	Clinical	10.40	Mixed	PC	FSIQ	MRI	WISC-III and WISC-IV	47	.26
Qiu et al.	2009	–	–	–	–	–	–	Clinical	10.40	Mixed	PC	Performance	MRI	WISC-III and WISC-IV	47	.20
Qiu et al.	2009	–	–	–	–	–	–	Clinical	10.40	Mixed	PC	Verbal	MRI	WISC-III and WISC-IV	47	.21
Qiu et al.	2009	–	–	–	–	–	–	Healthy	10.50	Mixed	PC	FSIQ	MRI	WISC-III and WISC-IV	66	.26*
Qiu et al.	2009	–	–	–	–	–	–	Healthy	10.50	Mixed	PC	Performance	MRI	WISC-III and WISC-IV	66	.12
Qiu et al.	2009	–	–	–	–	–	–	Healthy	10.50	Mixed	PC	Verbal	MRI	WISC-III and WISC-IV	66	.35**
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	Performance	MRI	HAWIE-R	13	.16
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	Verbal	MRI	HAWIE-R	10	−.17
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	FSIQ	MRI	HAWIE-R	10	.02
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	Performance	MRI	HAWIE-R	10	.23
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	Verbal	MRI	HAWIE-R	13	.35
Weniger et al.	2009	–	–	–	–	–	–	Healthy	33.00	Women	PC	FSIQ	MRI	HAWIE-R	25	.15
Weniger et al.	2009	–	–	–	–	–	–	Healthy	33.00	Women	PC	Performance	MRI	HAWIE-R	25	.24
Weniger et al.	2009	–	–	–	–	–	–	Healthy	33.00	Women	PC	Verbal	MRI	HAWIE-R	25	.00
Weniger et al.	2009	–	–	–	–	–	–	Clinical	32.00	Women	PC	FSIQ	MRI	HAWIE-R	13	.27
Castro-Fornieles et al.	2009	–	–	–	–	–	–	Clinical	14.50	Mixed	PC	Performance	MRI	WISC-R	12	.38
Castro-Fornieles et al.	2009	–	–	–	–	–	–	Clinical	14.50	Mixed	PC	Verbal	MRI	WISC-R	12	.11
Castro-Fornieles et al.	2009	–	–	–	–	–	–	Healthy	14.60	Mixed	PC	Performance	MRI	WISC-R	9	.55
Castro-Fornieles et al.	2009	–	–	–	–	–	–	Healthy	14.60	Mixed	PC	Verbal	MRI	WISC-R	9	.43
Isaacs et al.[a]	2010	–	–	–	–	–	–	Healthy	15.75	Women	PC	FSIQ	MRI	WISC-III and WAIS-III	24	.00
Isaacs et al.[a]	2010	–	–	–	–	–	–	Healthy	15.75	Women	PC	Performance	MRI	WISC-III and WAIS-III	24	.00
Isaacs et al.[a]	2010	–	–	–	–	–	–	Healthy	15.75	Women	PC	Verbal	MRI	WISC-III and WAIS-III	24	.00

Isaacs et al.	2010	–	–	–	–	–	–	Healthy	15.75	Men	Reported	FSIQ	MRI	WISC-III and WAIS-III	26	.36
Isaacs et al.	2010	–	–	–	–	–	–	Healthy	15.75	Men	Reported	Performance	MRI	WISC-III and WAIS-III	26	.19
Isaacs et al.	2010	–	–	–	–	–	–	Healthy	15.75	Men	Reported	Verbal	MRI	WISC-III and WAIS-III	26	.48[**]
Betjemann et al.	2010	–	–	–	–	–	–	Healthy	11.40	Mixed	Reported	Performance	MRI	WISC-R	142	.42[***]
Betjemann et al.	2010	–	–	–	–	–	–	Healthy	11.40	Mixed	Reported	Verbal	MRI	WISC-R	142	.14
Lange et al.	2010	–	–	–	–	–	–	Healthy	10.88	Women	Reported	FSIQ	MRI	WASI	166	.22[**]
Lange et al.	2010	–	–	–	–	–	–	Healthy	10.88	Women	Reported	Performance	MRI	WASI	155	.20[**]
Lange et al.[a]	2010	–	–	–	–	–	–	Healthy	10.88	Women	PC	Verbal	MRI	WASI	155	.00
Lange et al.	2010	–	–	–	–	–	–	Healthy	10.95	Men	Reported	FSIQ	MRI	WASI	143	.23[**]
Lange et al.	2010	–	–	–	–	–	–	Healthy	10.95	Men	Reported	Performance	MRI	WASI	130	.28[***]
Lange et al.[a]	2010	–	–	–	–	–	–	Healthy	10.95	Men	PC	Verbal	MRI	WASI	130	.00
Hogan et al.	2010	–	–	–	–	–	–	Healthy	68.69	Mixed	PC	Fluid	MRI	SPM	234	.11
Hogan et al.	2010	–	–	–	–	–	–	Healthy	68.69	Mixed	PC	Verbal	MRI	NART	235	<.01
Hermann et al.[c]	2002	–	–	–	–	–	–	Clinical	36.09	Mixed	PC	FSIQ	MRI	WAIS-III	77	.21
Hermann et al.[c]	2002	–	–	–	–	–	–	Clinical	36.09	Mixed	PC	Performance	MRI	WAIS-III	77	.09
Hermann et al.[c]	2002	–	–	–	–	–	–	Clinical	36.09	Mixed	PC	Verbal	MRI	WAIS-III	77	.28[*]
Hermann et al.[c]	2002	–	–	–	–	–	–	Healthy	33.34	Mixed	PC	FSIQ	MRI	WAIS-III	67	.31[*]
Hermann et al.[c]	2002	–	–	–	–	–	–	Healthy	33.34	Mixed	PC	Performance	MRI	WAIS-III	67	.33[**]
Hermann et al.[c]	2002	–	–	–	–	–	–	Healthy	33.34	Mixed	PC	Verbal	MRI	WAIS-III	67	.23
Wallace et al.	2010	–	–	–	–	–	–	Healthy	11.80	Mixed	Reported	FSIQ	MRI	WASI	649	.14[***]
Wallace et al.	2010	–	–	–	–	–	–	Healthy	11.80	Mixed	Reported	Performance	MRI	WASI	649	.14[***]
Wallace et al.	2010	–	–	–	–	–	–	Healthy	11.80	Mixed	Reported	Verbal	MRI	WASI	649	.13[***]
Ashtari et al.	2011	–	–	–	–	–	–	Healthy	18.50	Men	Reported	FSIQ	MRI	WRAT-III	14	.57[*]
Ashtari et al.	2011	–	–	–	–	–	–	Clinical	19.30	Men	Reported	FSIQ	MRI	WRAT-III	14	.29
Kievit et al.	2011	–	–	–	–	–	–	Healthy	21.10	Mixed	PC	Performance	MRI	WAIS-III	80	.29
Kievit et al.	2011	–	–	–	–	–	–	Healthy	21.10	Mixed	PC	Verbal	MRI	WAIS-III	80	.23
Tate et al.[a]	2011	–	–	–	–	–	–	Clinical	81.70	Mixed	Reported	FSIQ	MRI	Shipley Scale	194	.00
Royle et al.	2013	–	–	–	–	–	–	Healthy	73.00	Men	Reported	FSIQ	MRI	WAIS-III	327	.27[***]
Royle et al.	2013	–	–	–	–	–	–	Healthy	73.00	Women	Reported	FSIQ	MRI	WAIS-III	293	.26[***]
Burgaleta et al.	2012	–	–	–	–	–	–	Healthy	19.88	Mixed	Reported	FSIQ	MRI	APM, DAT, and PMA	100	.17
Aydin et al.	2012	–	–	–	–	–	–	Healthy	15.10	Men	Reported	FSIQ	MRI	WISC-R	30	.40[*]
Aydin et al.	2012	–	–	–	–	–	–	Healthy	15.10	Men	Reported	Performance	MRI	WISC-R	30	.34
Aydin et al.	2012	–	–	–	–	–	–	Healthy	15.10	Men	Reported	Verbal	MRI	WISC-R	30	.26

Note. Review Coverage: x indicates that study had been included in previous review, –indicates that study had not been included in previous Review; PC = Personal communication; FSIQ = Full-scale IQ; CT = X-ray Computed Tomography, MRI = Magnetic Resonance Imaging, WDM = Water Displacement Method; APM = Raven's Advanced Progressive Matrices, BCS = Bracken Basic Concepts Scale, BS = Bayley Scales of Infant Development, CFIT = Cattell's Culture Fair Intelligence Test, COWA = Controlled Oral Word Association, CVLT = California Verbal Learning Test, CWAT = Controlled Word Association Test, DAS = Differential Ability Scale, DAT = Differential Aptitude Test, GMDS = Griffiths Mental Developmental Scales, HAWIE-R = Revised German Version of the Wechsler Adult Intelligence Scale, JLOT = Judgment of line orientation test, MAB = Multidimensional Aptitude Battery, MHT = Moray House Test, MWT = Multiple-choice Vocabulary Test, MSEL = Mullen Scales of Early Learning, NART = New Adult Reading Test, PMA = Primary Mental Abilities, Quick = Quick IQ Test, RPM = Raven's Progressive Matrices, SBIS = Stanford Binet Intelligence Scale, SILT = Shipley Institute of Living Test, SPM = Raven's Standard Progressive Matrices, TMT = Trail Making Test, V2 = Vocabulary Test, V3 = Extended Vocabulary Test, VKMRT = Vandenberg and Kuse Mental Rotation Test, WAIS = Wechsler Adult Intelligence Scale, WAIS-R = Wechsler Adult Intelligence Scale Revised, WAIS-IV = Wechsler Adult Intelligence Scale, 4th edition, WASI = Wechsler Abbreviated Scale of Intelligence, WISC-R = Wechsler Intelligence Scale for Children Revised, WISC-III = Wechsler Intelligence Scale for Children, 3rd edition, WISC-IV = Wechsler Intelligence Scale for Children, 4th edition, WJ-III = Woodcock-Johnson-Test of Cognitive Abilities, 3rd edition, WMS = Wechsler Memory Scale, WPPSI-R = Wechsler Preschool and Primary Scales of Intelligence Revised, WRAT-III = Wide Range Achievement Test, 3rd edition.

[a] No numerical value reported for nonsignificant effect size, thus set to zero.

[b] parts of data of Staff et al. (2006) was reported in Staff (2002) and included in McDaniel (2005).

[c] Study author was contacted regarding a paper published in 2002 but provided more recent data from 2010; Review coverage = study included in the review of Rushton and Ankney (1996; R1), Rushton and Ankney (2000; R2); Vernon et al. (2000; V), Gignac et al. (2003; G), McDaniel (2005; M), Rushton and Ankney (2009; R3).

[*] $p < .05$.

[**] $p < .01$.

[***] $p < .001$.

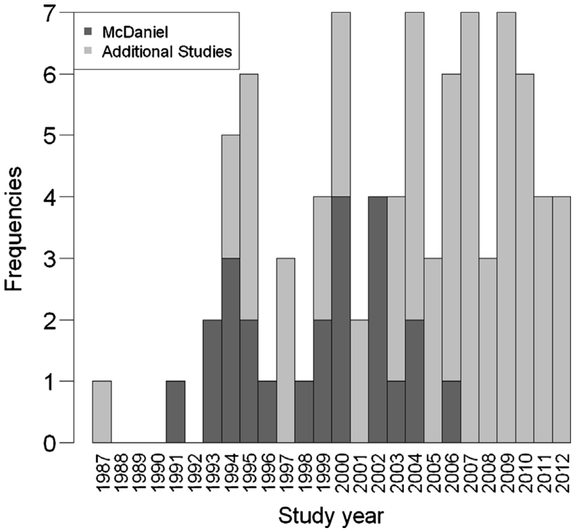

Fig. 1. Study year of included studies.

confounding publication bias. Publication bias refers to the tendency of researchers and journals to publish significant findings and strong effects more often, quicker, and more prominently (e.g., Rothstein et al., 2005) which was not assessed due to the relatively small number of correlations in McDaniel's meta-analysis. Additionally, only a relatively small number of moderating variables was accounted for (age and sex), and differences were assessed using a rather crude method (i.e., subgroup analysis). Indeed, the pattern of observed results of investigations examining associations of brain volume and IQ make it evident that a positive association of these variables is to be expected, yet the strength of this effect and potentially moderating variables remain unclear. Moreover, because scientific productivity on this particular topic has been increasing recently, resulting in large numbers of recent investigations addressing relationships of in vivo brain volume and IQ (see Table 1), results of the preceding meta-analysis need to be updated (Fig. 1).

Differential associations could arise for full-scale, performance, and verbal IQ. Specifically according to *g* theory (Jensen, 1998), full-scale IQ is more strongly associated with the general factor of intelligence (*g*) than lower-order factors such as performance or verbal IQ (single domains of intelligence naturally display lower *g*-loadings than full-scale IQ, because *g* should consist of all relevant domains of mental ability; Jensen, 1998, pp. 73–81), and hence full-scale IQ (as most highly loaded on *g*) should display the strongest correlations.

In order to update previous analyses, and to shed more light on moderating variables as well as generalizability to the general population and further intelligence domains, we present here the most comprehensive meta-analysis on this subject thus far. Associations of full-scale, performance, and verbal IQ with in vivo brain size of healthy and clinical samples (more than three times more samples and five times more participants than McDaniel, 2005) were investigated over a time-span of 25 years.

In the present meta-analysis, effect sizes were based on results published in the literature and obtained through personal communications with researchers of this field, thus making it possible to directly assess influences of selective reporting. In addition to more conventional methods such as subgroup analyses, we calculated hierarchical weighted multiple linear meta-regressions including several moderating variables. Furthermore, several methods for detection of publication bias were applied, as there is ample evidence of excessive bias due to selective publication and reporting in studies investigating associations of brain volume with other variables (Ioannidis, 2011).

3.1. *Methods*

3.1.1. *Literature search*

First, we screened reference lists of six early reviews on associations of brain size and IQ to obtain potentially relevant studies (Gignac et al., 2003; McDaniel, 2005; Rushton and Ankney, 1996, 2000, 2009; Vernon et al., 2000). Second, we searched ISI Web of Science for all studies citing at least one of these six reviews. Third, we screened reference lists of retrieved studies for additional potentially eligible studies. Fourth, we entered search terms "brain volume AND intelligence", "brain volume AND IQ", "brain size AND intelligence", and "brain size AND IQ" in three scientific databases (ISI Web of Science, PubMed, Scopus) and assessed titles for relevance. Finally, we screened abstracts of 444 possibly relevant articles for eligibility (Fig. 2). Relevant literature was searched until May 2012.

3.1.2. *Inclusion criteria*

In order for studies to be included in the present analysis, they needed to fulfill three criteria. First, brain volume had to be assessed in individuals. Partial assessment of the brain (e.g., grey-matter only; Thompson et al., 2001; correlations of brain macro-structure with IQ only; Posthuma et al., 2002) was insufficient for inclusion in the analysis. Rather, assessment of whole brain or intracranial volume by X-ray Computed Tomography (2 studies; Jones et al., 1994; Yeo et al., 1987), Magnetic Resonance Imaging (MRI), or the Water Displacement Method (1 study; Witelson et al., 2006) had to be reported. Second, measures of full-scale IQ, performance IQ, or verbal IQ had to have been completed by participants. Third, reported data had to be independent of data of any other study included in the present meta-analysis. In cases where these criteria were met, but correlation coefficients were not reported, corresponding authors were personally contacted by email and asked to provide the relevant results. In our analyses, we used reported vs. non-reported coefficients as a moderating variable because non-reported coefficients may be expected to show lower values due to underreporting (i.e., due to publication bias).

3.1.3. *Coding*

Two experienced researchers (J.P., M.Z.) independently coded studies into categories (aim of study, inclusion in previous meta-analysis, sample type, type of psychometric test instrument, type of volumetric measure) and recorded relevant variables as well as sample characteristics. Additionally, the number of statistical corrections in the form of covariates used to calculate the effect sizes (e.g., height, weight) were assessed in order to allow consideration in meta-regressions as outlined in Section 2.6. Inconsistencies in coding were resolved by discussion.

In a number of studies, correlation coefficients of non-significant associations of IQ and brain volume were not reported. Whenever this was the case, corresponding authors of the respective articles were contacted and correlation coefficients were obtained through personal communications. Otherwise, following a conservative approach as described by Pigott (2009, pp. 408–409), non-significant effect sizes were set to zero (5, 11, and 3 effects for full-scale, verbal, and performance IQ, respectively).

3.1.4. *Data analysis*

Overall strength of associations of brain volume and IQ was estimated using random-effects models. As a descriptive measure of variability we computed the index I^2, which reflects the percentage of variability between effects due to true heterogeneity (i.e., bigger values of I^2, reflect more heterogeneity). Associations were

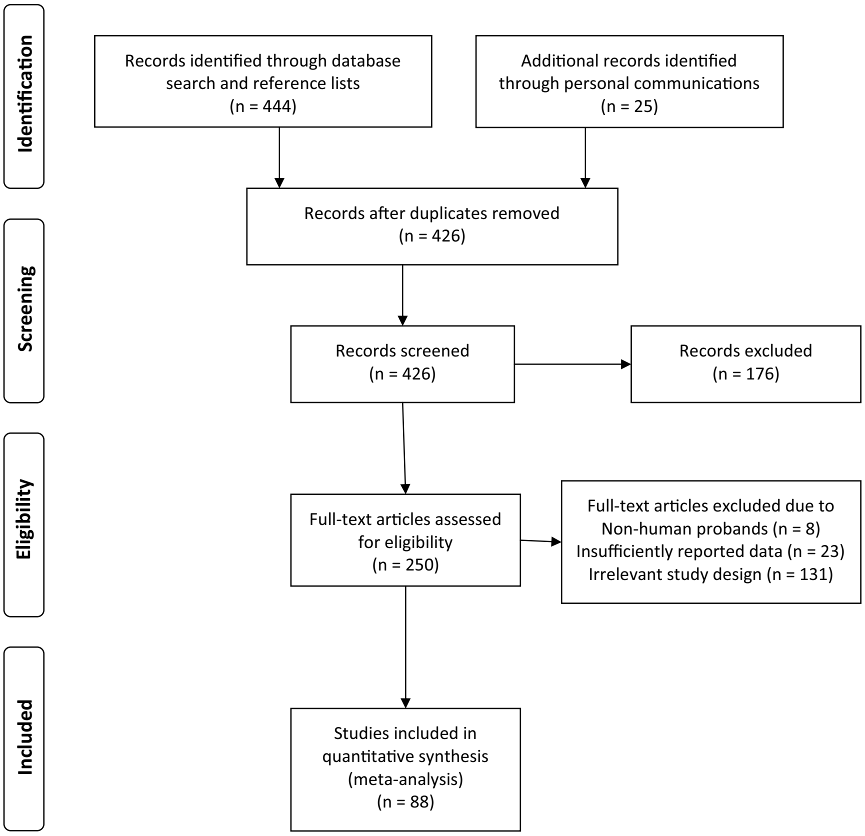
Identification
Screening
Eligibility
Included
Records identified through database search and reference lists (n = 444)
Additional records identified through personal communications (n = 25)
Records after duplicates removed (n = 426)
Records screened (n = 426)
Records excluded (n = 176)
Full-text articles assessed for eligibility (n = 250)
Full-text articles excluded due to
Non-human probands (n = 8)
Insufficiently reported data (n = 23)
Irrelevant study design (n = 131)
Studies included in quantitative synthesis (meta-analysis) (n = 88)

Fig. 2. Flowchart of study inclusion.

meta-analyzed independently for full-scale IQ, performance IQ, and verbal IQ, so that single studies could feature in each of the three analyses.

Additionally, we performed sensitivity analyses by omitting one effect size in each run of the overall analysis, respectively, in order to assess potentially distorting effects of individual effect sizes. To avoid well-known unfavorable effects of using the correlation coefficient *r* for overall effect size estimations (e.g., underestimation of effects), effect sizes were transformed to Fisher's *Z* prior calculation, following standard meta-analytical protocols (e.g., Borenstein, 2009, p. 231). We used $1/(n-3)$ as sampling variance. For ease of interpretation, we report results after back-transformation (i.e., in the *r* metric). In all calculations, studies were weighted according to study precision (inverse standard error of effect sizes). All analyses were performed using CMA (Comprehensive Meta-Analysis v2.2.030), the open-source software R 3.1.1 (R Development Core Team, 2014), and the packages metafor (Viechtbauer, 2010) and pwr (Champely et al., 2012) for R.

3.1.4.1. Subgroup analysis. To assess possible influences of moderating variables, we performed a series of subgroup analyses. Effect sizes were grouped according to effect reporting (correlation coefficient reported in publication or not), sample type (clinical vs. healthy samples), and sex (men-only vs. women-only sample). Calculations were based on mixed-effects models (i.e., within-subgroup estimates are based on random-effects but across-subgroup calculations on a fixed-effect analysis).

3.1.4.2. Meta-regression. In order to allow more fine-grained analyses of moderating variables, we applied weighted linear meta-regressions. First, for assessment of effect strength development over time, meta-regressions for study year on study effects were calculated for effect sizes of all healthy samples. Second, percentage of men in samples was regressed on study effects of healthy, clinical, and all samples, because previous findings indicated significant differences regarding sex (McDaniel, 2005). Third, theory-guided hierarchical weighted multiple mixed-effects meta-regressions were calculated. Study year was used as a single predictor in the initial model, as a time trend was hypothesized (i.e., a decline in effect strength; Schooler, 2011). In a second block, sample age (children vs. adults) and male percentage (percentage of men in samples) were included, because previous results indicated influences of these predictors on the strength of associations (McDaniel, 2005). In a third block, aim of study (main study goal was assessment of IQ brain size correlation vs. different main goal), effect reporting (correlation coefficient reported in publication or not), number of included covariates in primary study, sample type (clinical vs. healthy samples), and type of test (Wechsler-type test or not) were added as predictor variables. In the final model, inverse sample variances (as it can be considered to be indicative of confounding publication bias) were included as a predictor. Goodness of model fit was examined by changes in explained variance (R^2). Finally, interactions of significant predictors were explored by means of weighted multiple moderated meta-regressions (Jaccard and Turrisi, 2003).

Results of all meta-regressions are reported in units of Fisher's *Z*, although in order to assess stability of results, calculations were repeated using correlation coefficients. This approach was chosen due to the skewed distributional characteristics of *r*.

3.1.4.3. Publication bias. To further clarify the pattern and to shed light on influences on strength of effects depending on whether or not the correlation coefficient had been reported in a research paper, several methods for assessment of publication bias were applied. Application of a relatively large number of different approaches seems appropriate, as ramifications of publication bias

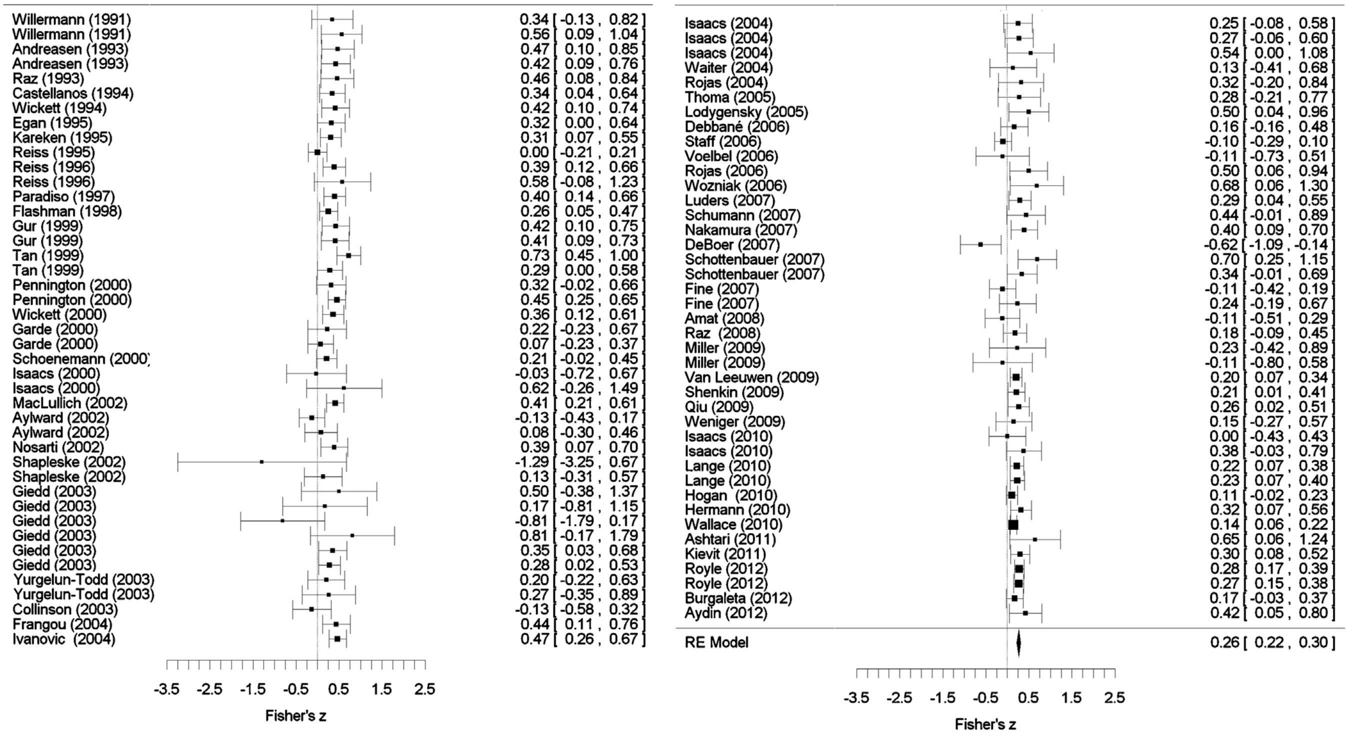

Fig. 3. Forest Plot for Meta-Analysis of Associations of Full-scale IQ and In Vivo Brain Size of Healthy Samples. *Note.* Overall effect size calculations are based on a random effects model; diamond represents overall effect size; symbol size is varied according to relative study weight within analysis; numbers in brackets are 95% confidence intervals of point estimation.

Table 2
Overall and subgroup-specific effect sizes for full-scale, performance, and verbal IQ.

	Full-scale IQ						Performance IQ						Verbal IQ					
	k	*n*	I^2	*r*	LCI	UCI	*k*	*n*	I^2	*r*	LCI	UCI	*k*	*n*	I^2	*r*	LCI	UCI
All samples	120	6778	38.85	.24***	.21	.27	64	3806	17.27	.21***	.17	.24	99	5458	55.93	.21***	.16	.25
Reported	53	3956	33.39	.30***	.25	.34	28	2580	41.35	.24***	.19	.30	47	3205	66.12	.28***	.21	.34
Personal communication	67	2822	34.41	.17***	.13	.23	36	1226	<0.01	.16***	.10	.21	52	2253	33.90	.14***	.08	.19
Healthy samples	84	5040	37.06	.26***	.22	.29	41	2845	17.03	.22***	.18	.27	60	3943	58.55	.18***	.13	.24
Reported	38	3254	22.14	.30***	.26	.34	21	2288	42.03	.26***	.20	.32	30	2508	65.44	.24***	.17	.31
Personal communication	46	1786	38.92	.19***	.12	.25	20	557	<0.01	.16***	.08	.25	30	1435	40.59	.11**	.03	.18
Clinical samples	36	1738	42.23	.20***	.13	.26	23	961	16.11	.16***	.08	.23	39	1515	43.48	.25***	.18	.33
Reported	15	702	53.83	.25***	.13	.37	7	292	45.28	.17*	.01	.32	17	697	62.62	.36***	.23	.48
Personal communication	21	1036	25.42	.16***	.08	.24	16	669	1.34	.15***	.07	.23	22	818	<0.01	.21***	.14	.28

Note. I^2 = percentage of variability between effects due to true heterogeneity; LCI = lower bound of 95% confidence interval; UCI = upper bound of 95% confidence interval.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

have been frequently demonstrated (e.g., Pietschnig et al., 2010). Importance of this matter is reflected by growing awareness of this issue in the scientific community and the development of new methods to account for it. The different approaches of these methods allow a comprehensive assessment of publication bias by providing differential perspectives on funnel plot asymmetry as outlined below.

First, funnel plots were visually inspected for evidence of asymmetry (Light and Pillemer, 1984). Second, Begg and Mazumdar's rank correlational method was employed (Begg and Mazumdar, 1994). This method is based on a significance test examining whether or not there is an indication of an association between study effect sizes and study precision. In absence of publication bias, there should be no such association observable. It should be noted that in presence of less than 25 samples, this method possesses only moderate power to detect publication bias (Sterne et al., 2000). Third, we applied Sterne and Egger's (2005) mixed effect regression method. In this method, study precision is regressed on the standard normal deviate of effect sizes (i.e., effect sizes divided by their standard errors) which should not lead to an intercept differing from zero in case of absence of publication bias. Fourth, we used Ioannidis and Trikalinos' (2007) test for excess significance. This test compares the number of observed significant effect sizes with the number of expected significant effects based on the cumulative power of studies (overall effect size estimates were used in power calculations). Analyses for excess significance were performed for all effect sizes as well as for reported effects only. Fifth, we used Trim-and-fill analysis (Duval and Tweedie, 2000) which detects funnel plot asymmetry on one tail of the effect size distribution (i.e., typically effects smaller in strength than the overall effect) and fills in missing studies to correct for the estimated bias. Moreover, Trim-and-fill provides an estimate of the overall effect based on all observed and imputed studies, although the authors of this method caution against interpretation of this effect other

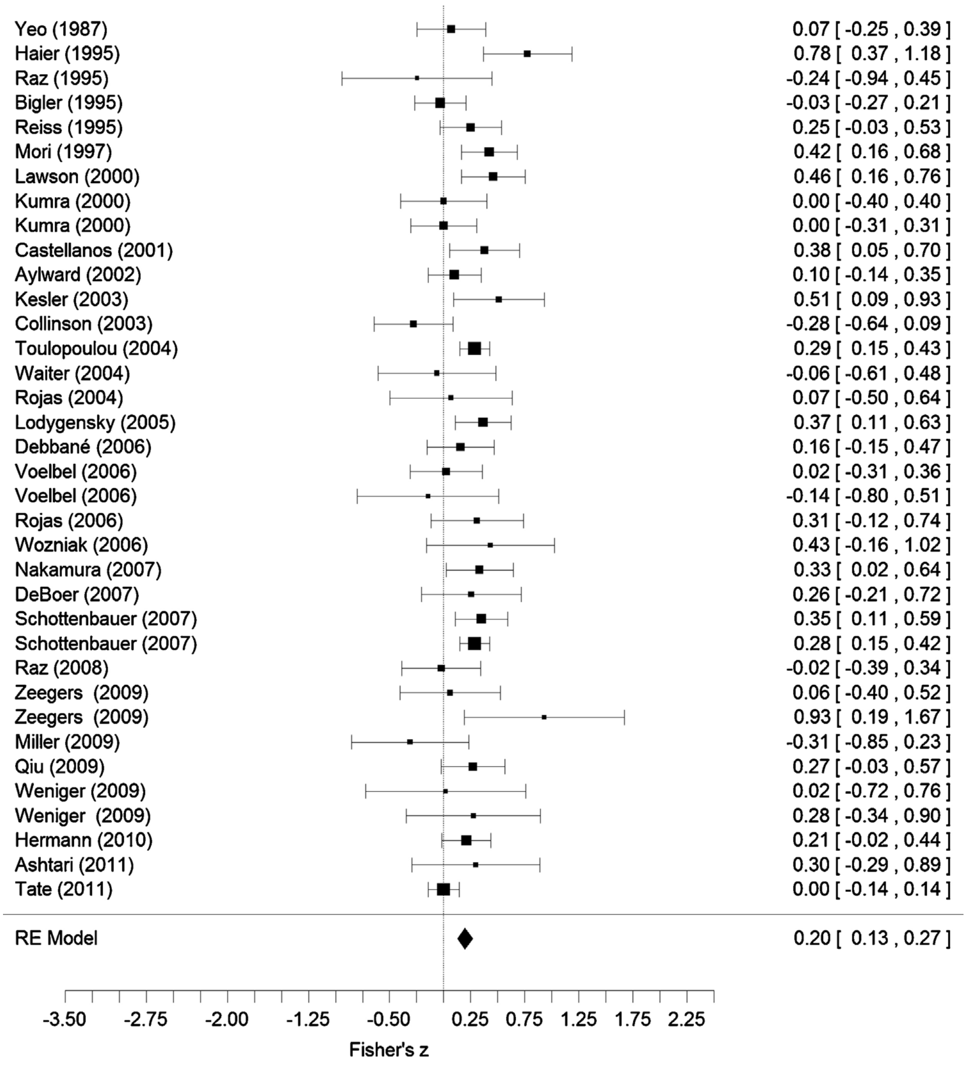

Fig. 4. Forest Plot for Meta-Analysis of Associations of Full-scale IQ and In Vivo Brain Size of Patient Samples. *Note.* See Fig. 3.

than a sensitivity analysis. Sixth, sensitivity analyses using different selection models were performed. We used four different study weight functions as specified by Vevea and Woods (2005) assuming either moderate or severe and one- or two-tailed selection of effect sizes. In absence of publication bias, overall effect estimates should not substantially differ between the uncorrected estimate and corrected estimates based on either of the specified selection models. All calculations for publication bias were performed only for effect sizes that had been reported (i.e., no fixed effect sizes or such that had been obtained through personal communications were included in these calculations) except for the excess significance method.

3.1.5. *Final sample*

In all, 88 studies comprising 148 independent healthy and clinical mixed-sex samples (8036 individuals) were included in the data analyses (McDaniel, 2005, covered 21.1% of included studies). Clinical samples were defined as samples comprising individuals with conditions likely to affect cognitive processing (autism, schizophrenia, traumatic brain injury). Such conditions are likely to affect associations between IQ and brain volume. For 28 of all samples, no associations of full-scale IQ with brain volume were available, so the main analysis was based on 120 effect sizes (54 reported, 66 obtained through personal communications or set to zero; Table 1). Assessments of dependent variables of these studies were based on 39 different but mainly Wechsler-type IQ test measures. All data and R codes are available from the supplementary material.

3.2. *Results*

3.2.1. *Brain volume and IQ*

The 120 correlation coefficients yielded a highly significant overall effect of $r = .24$ ($p < .001$; 95% CI [.21, .27]) for full-scale IQ. Forest plots (depicting single study effects, confidence intervals, and overall effects) for healthy and clinical samples are depicted in Figs. 3 and 4, respectively. The effect generalized to performance IQ ($r = .21$) and verbal IQ ($r = .21$), although effect sizes were somewhat lower for these intelligence domains (no formal significance test was carried out to assess these differences due to data dependencies). Of note, in all intelligence domains effects seemed to be stronger (i) for healthy than for clinical samples and (ii) for reported correlations than for such that had been obtained through personal communications (except for stronger effects of clinical samples for verbal IQ; Table 2). Sensitivity analyses revealed that omitting single studies had negligible influences on strength of effects for all three intelligence domains, thus demonstrating stability of results (i.e., no threats to validity of results due to leverage points). Although results invariably yielded significant positive associations of brain volume and intelligence, visual inspection of strength

Table 3
Subgroup analyses for full-scale IQ according to age, sample type, sex, and publication status.

Comparison	*k*	*Q*	*p*
Healthy vs clinical samples	120	2.57	.109
Adults vs children	120	0.41	.521
Reported *r* vs personal communication	120	13.77	<.001
Healthy	84	10.70	.001
Clinical	36	2.01	.157
Men vs women	57	0.81	.369
Healthy	47	0.34	.563
Clinical	10	0.95	.330

Note. Threshold for classification as adult was being of age 19 or older; exclusively female-only and male-only samples were used to calculate subgroup analyses for men and women; *Q* = weighted sum of squared differences between individual study effects and pooled study effect (Cochrans'*Q*); *df* = 1 for all analyses.

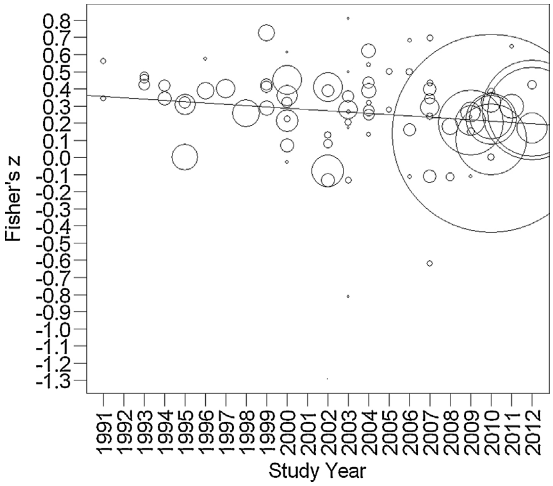

Fig. 5. Cross-Temporal Meta-Regression on Fisher's *z* for Healthy Samples. *Note.* Symbol size is varied according to relative study weight within analysis.

of effects indicated differentiated outcomes regarding moderating variables (Table 2).

We acknowledge that some readers might feel more comfortable in interpreting overall effects without inclusion of effect sizes that have been set to zero (i.e., thus providing a more liberal estimate, instead of a conservative one). Removal of these assumed zero effects yielded $r = .25$ ($p < .001$; 95% CI [.22, .29]; $k = 115$) for full-scale, $r = .21$ ($p < .001$; 95% CI [.17, .25]; $k = 61$) for performance, and $r = .23$ ($p < .001$; 95% CI [.18, .28]; $k = 88$) for verbal IQ. However, we suggest to interpret these effects cautiously because they provide necessarily a somewhat inflated estimate.

3.2.2. *Subgroup analysis*

Table 3 summarizes results for subgroup analyses for full-scale IQ of dichotomous moderator variables. Observed associations between brain volume and full-scale IQ were significantly higher in healthy samples and for correlation coefficients that had been reported in published articles. No significant differences between associations regarding sex or age of participants (children vs. adults) could be shown. For performance IQ and verbal IQ, the patterns of these results were virtually identical, although for verbal IQ subgroup analyses for healthy and clinical samples failed to reach significance (results omitted for brevity).

3.2.3. *Meta-regression*

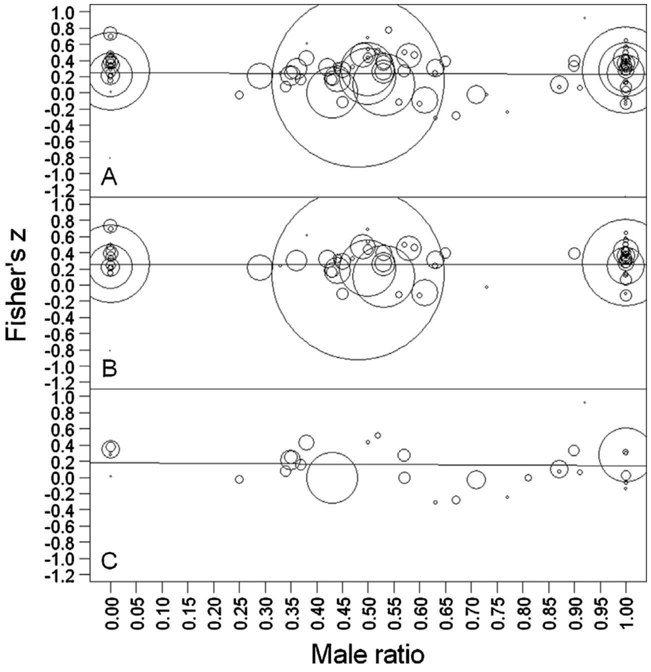

Fig. 6. Meta-Regression of Percentage of Men in Samples on Fisher's *z*. *Note.* Symbol size is varied according to relative study weight within analysis; (A) overall samples, (B) healthy samples, (C) clinical samples.

First, when study year was regressed on all healthy samples a significant decrease of effects over time was observed (slope = −0.008, $p = .02$; Fig. 5). Second, meta-regressions of percentage of men within samples on effect sizes did not yield significant influences of sex in healthy, clinical, or overall samples (Fig. 6). Indeed, signs of slopes were negative in all three regressions, indicating slightly stronger effects for women (slopes = −0.001, −0.041, and −0.015, respectively).

Third, the initial model of the hierarchical weighted meta-regression showed significant influences of study year as the single predictor on the overall effect, indicating that effects decreased in strength over time ($R^2 = .10$). When sample age (children vs. adults) and male percentage (percentage of men within samples) were added as predictors, all included predictors failed to reach significance and explained variance decreased ($R^2 = .04$) indicating worse model fit. In the third model, inclusion of study aim, effect reporting (correlation coefficient reported in publication or not), number of included covariates in primary study, sample type (clinical vs. healthy samples), and type of test (Wechsler-type test or not) as predictors increased explained variance again ($R^2 = .15$). In this model, two significant predictors emerged (effect reporting and sample type). Slopes of these predictors showed significant stronger effects for reported coefficients and healthy samples. Our fourth model explained the highest amount of explained variance ($R^2 = .30$), thus indicating the best model fit. Four significant predictors emerged, indicating stronger effects for reported coefficients, healthy samples, small samples, as well as Wechsler-type tests (for summary statistics of regression models, see Table 4).

Finally, weighted moderated regressions of significant model predictors on effect sizes showed no meaningful significant first or second order interactions. In order to assess robustness of these results, all regression analyses were repeated using the correlation coefficient *r* as dependent variable. Results were virtually identical to analyses based on Fisher's *z* (i.e., same predictors emerged as meaningful in hierarchical regressions, moderated regressions showed no meaningful significant interactions; numerical results omitted for brevity).

3.2.4. *Publication bias*

Visual inspection of funnel plots indicated slight asymmetry to the left of the overall effect for full-scale IQ. Begg and

Table 4
Hierarchical weighted meta-regression on effect sizes.

	Coefficient *b*	SE	LBCI	UBCI	*p*	R^2
Initial model						.10
Study year	−0.006	0.003	−0.012	>−.001	.046	
Second block						.04
Study year	−0.006	0.003	−0.012	<.001	.052	
Children (0) vs adult sample (1)	0.013	0.036	−0.058	0.084	.713	
Male percentage in sample	−0.031	0.052	−.132	0.070	.553	
Third block						.15
Study year	−0.005	0.003	−0.011	0.001	.091	
Children (0) vs adult sample (1)	0.020	0.036	−0.051	0.092	.579	
Male percentage in sample	−0.030	0.051	−0.129	0.070	.558	
Assessment of association was not (0) or was (1) main goal of study	−0.011	0.048	−0.105	0.084	.828	
Effect obtained through personal communication or set to zero (0) vs reported effects (1)	0.095	0.045	0.008	0.183	.032	
Healthy (0) vs clinical sample (1)	−0.088	0.044	−0.174	−0.002	.045	
Non-WAIS-type tests (0) vs WAIS-type tests (1)	0.063	0.039	−0.014	0.139	.107	
Number of included covariates	−0.006	0.022	−0.049	0.037	.780	
Final model						0.30
Study year	−0.002	0.003	−0.008	0.005	.644	
Children (0) vs adult sample (1)	0.017	0.035	−0.052	0.086	.634	
Male percentage in sample	−0.031	0.049	−.128	0.065	.523	
Assessment of association was not (0) or was (1) main goal of study	0.010	0.047	−0.083	0.103	.832	
Effect obtained through personal communication or set to zero (0) vs reported effects (1)	0.114	0.044	0.028	0.201	.010	
Healthy (0) vs clinical sample (1)	−0.087	0.043	−0.171	−0.003	.042	
Non-WAIS-type tests (0) vs WAIS-type tests (1)	0.076	0.038	0.002	0.150	.045	
Number of included covariates	−0.006	0.021	−0.048	0.036	.784	
Inverse variance of samples	>−0.001	<0.001	−0.001	>−0.001	.025	

Note. $k = 115$; LBCI = lower bound of 95% confidence interval; UBCI = upper bound of 95% confidence interval; Studies were weighted according to sampling variance ($1/(n-3)$); b = unstandardized regression coefficient; calculations are based on mixed-effects models and Fisher's z transformation.

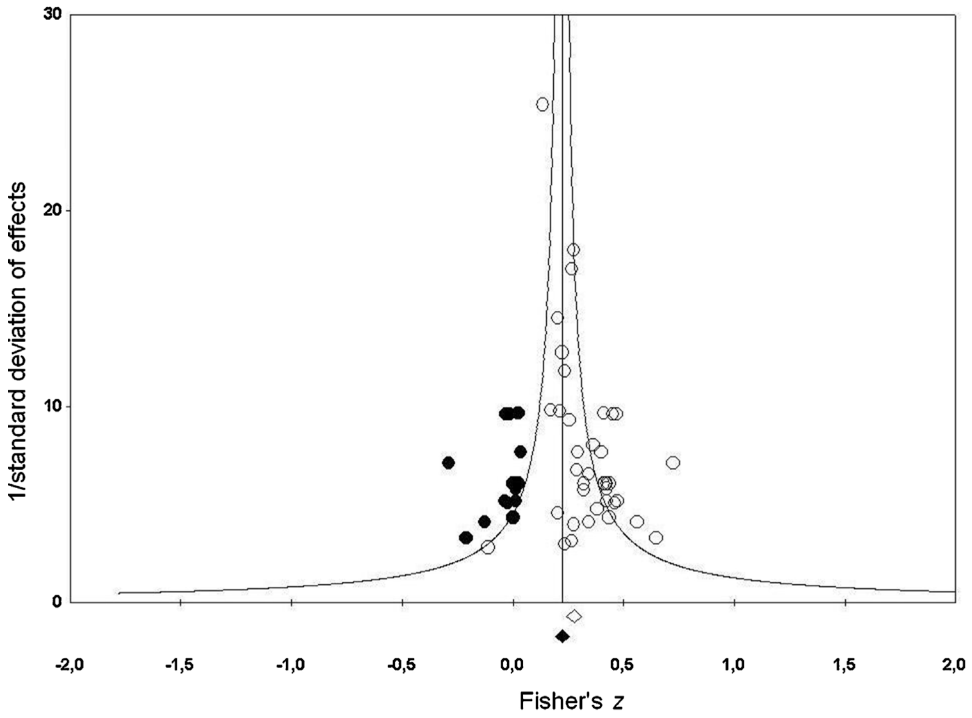

Fig. 7. Funnel Plot for Reported Healthy Samples. *Note.* Circles represent reported effect sizes; solid circles represent missing studies according to Trim-and-fill analysis for publication bias; studies' weights are displayed on the ordinate according to study precision (1/standard deviation of effects); hollow diamond represents observed overall effect size; solid diamond represents adjusted effect size (Fisher's z).

Mazumdar's rank correlation ($r_s = .08$; $p = .19$) did not reach significance, thus indicating no evidence for publication bias, but Sterne and Egger's regression method did ($p = .03$). There was no clear indication of excess significance for the overall effect ($p = .15$ and .06 for reported and all coefficients, respectively). Trim-and-fill analysis yielded 14 missing studies to the left of the overall effect, thus indicating considerable inflation of the overall effect and necessity for effect adjustment (Fig. 7). Results from our selection model analyses corroborated the above evidence for effect size inflation (for a summary of results see first column of Table 5).

When calculations of these methods were performed separately for effect sizes of healthy and clinical samples, different patterns emerged for these two groups. Begg and Mazumdar's method again did not reach significance, but Sterne and Egger's regression yielded strong evidence for publication bias for healthy samples. Similarly, reported effect sizes of healthy but not clinical samples showed

Table 5
Results of five different indicators for publication bias for full-scale IQ effect sizes.

		Overall ($k = 53$)	Healthy samples ($k = 38$)	Clinical samples ($k = 15$)
Begg & Mazumdar	p Value	.19	.12	.35
Egger	p Value	0.03	0.001	.90
Excess significance (based on reported coefficients only)	$\chi^2(1)/p$ value	2.08/.15	5.29/0.02	1.05/.31
Excess significance (based on all coefficients)	$\chi^2(1)/p$ value	3.57/0.06	2.63/.10	0.57/0.45
Trim-and-fill	Observed r	.30	.30	.25
	Adjusted r	.24	.23	.25
	Added studies	14	16	0
Selection models (Vevea and Woods, 2005)	No selection	.30	.30	.25
	Moderate one-tailed selection	.28	.29	.22
	Severe one-tailed selection	.26	.28	.15
	Moderate two-tailed selection	.28	.29	.24
	Severe two-tailed selection	.27	.28	.22

Note. Only published studies were used to calculate measures for publication bias except for excess significance tests; all calculations were based on random effects models; p values for both Begg & Mazumdar's and Egger's test are 1-sided.

excess of significant results. Trim-and-fill analysis indicated sixteen missing effects to the left of the observed effect for healthy samples (more than a third of observed effects) but no missing studies for clinical samples. Interestingly, selection model approaches showed evidence for effect size inflation in both healthy and clinical samples (second and third column of Table 5). This pattern was virtually identical for performance IQ and verbal IQ, although evidence for effect inflation was strongest for verbal IQ and was observed for clinical samples as well (results omitted for brevity).

3.3. *Discussion of the meta-analysis*

In all, our findings demonstrate a moderate positive association of in vivo brain volume and intelligence that generalizes over full-scale IQ, performance IQ, and verbal IQ, but is differentiated in respect to test and sample type and is possibly inflated by selective reporting of significant effects. Our results indicate substantially weaker associations of brain volume and IQ than previous estimates.

Our findings raise several points of interest. First, brain volume was significantly positively associated with all three investigated intelligence domains (full-scale IQ, performance IQ, verbal IQ). In all, 6%, 4%, and 4% of variance, respectively, were attributable to these associations, thus yielding a moderate effect (Cohen, 1988). As hypothesized, these associations were stronger for full-scale IQ than for performance IQ and verbal IQ. This result was to be expected, because associations of full-scale IQ should display stronger effects than single components of the general construct due to higher loadings on psychometric *g*. These differences should be interpreted while keeping in mind that assessment of statistical significance of differences of effect strength was not possible because of data dependency (i.e., most studies provided correlation coefficients for all three intelligence domains based on the same sample).

Second, when inclusion criteria were specified to the criteria used by McDaniel (2005; i.e., inclusion of healthy samples only) study year turned out to be a significant predictor, thus possibly indicating a lag of publication of weaker and non-significant effects. This phenomenon is well-known and has been extensively discussed in the literature (Ioannidis, 1998; Schooler, 2011).

Third, effects were stronger when correlation coefficients had been reported in a publication, than when they had been obtained through personal communications as indicated in subgroup analysis. These results were consistent with the finding of stronger effects for reported (vs. non-reported) effect sizes and small samples in our regression analysis. This suggests reporting bias (i.e., more detailed, faster, and more visible reporting of results in the research literature in case of significant and strong effects; e.g., Hahn et al., 2002) as a conceivable source of effect inflation and validates the concerns raised by McDaniel (2005) in respect to validity of the observed overall effect due to selective reporting in previous meta-analyses. Additionally, analyses for publication bias indicated missing effect sizes at the lower tail of the effect size distribution (i.e., to the left of the observed mean effect) for published effects. This is consistent with previous findings of selective reporting in published literature investigating associations of brain volume with other variables (Ioannidis, 2011) and further corroborates stronger effects of published results. A correction using the Trim-and-fill method lowered the mean estimates to $r = .24$, $r = .21$, and $r = .17$ for full-scale IQ, verbal IQ, and performance IQ, respectively. This further illustrates effect inflation in all IQ domains due to publication bias.

Fourth, IQ as measured by Wechsler-type intelligence tests was more strongly associated with brain volume than other intelligence test measures. This finding may be explained by the relative narrowness of other employed test instruments and their smaller saturation in *g*. Although it cannot be completely ruled out that this effect might have been due to systematically higher reliabilities of Wechsler-type intelligence test, the typically high reliabilities of the included intelligence test measures render this alternative explanation less likely than the proposed effect of *g*.

Fifth, subgroup analysis showed stronger effects for healthy than clinical samples. This finding is not surprising, because individuals in included clinical samples suffered from a variety of conditions affecting cognitive processing (i.e., autism, schizophrenia, traumatic brain injury) which are likely to blur associations. In fact, all effect sizes of clinical samples were numerically lower than effect sizes of healthy samples except for associations with verbal IQ.

Sixth, in contrast to findings reported by McDaniel (2005), we did not observe significant sex differences. Noticeably, this non-significant result was not only shown by subgroup analysis for overall, healthy, and clinical samples, but also emerged in more fine-grained analyses (i.e., meta-regressions), which should be more sensitive in detection of differences.

Finally, no significant effects of sample age could be found. This emphasizes the robustness of the association as effect strength does not seem to be influenced by brain growth, but remains stable over age. Admittedly, it would have been preferable to include age as a continuous rather than dichotomous variable as this would have provided a more detailed assessment. However, because inclusion as continuous predictor would have led to a considerable loss of includable effect sizes due to infrequent reporting of mean age, inclusion as a dichotomous variable was preferred.

Of note, we did not apply range restriction corrections for sample attenuation in the present meta-analysis which might have led to a slight underestimation of overall effects. We decided not to apply corrections because for a majority of the included samples standard deviations for test performance were not reported. Therefore, correcting for range restriction would have required us to interpolate estimates for these studies based on a comparatively small number of reported parameters, thus introducing further uncertainty rather than allowing us to assess a hypothesized true value. Considering this, the present estimate based on the actual observed values was deemed a more reliable estimation of the overall association. Similarly, no corrections for measurement error of intelligence tests and volumetric measures were applied because both intelligence tests (Hunt, 2011) as well as volumetric measurements (MacLaren et al., 2014) have been typically observed to be highly reliable.

In all, the present study clearly demonstrates a positive moderate association of in vivo brain volume with intelligence. Furthermore, we could show that this effect is observable in healthy individuals as well as (albeit smaller) in clinical samples. Although the association is confounded by reporting bias and therefore smaller than presumed according to previous investigations, it is robust as it generalizes over age, intelligence domain, and sex.

4. Why is brain size associated with intelligence?

In vivo MRI studies clearly confirm earlier findings of an association between intelligence and brain size based on surrogate measures such as head dimensions or inner skull volume. Interestingly, the association generalizes across sex despite marked sexual dimorphism in brain volume, and the relationship with intelligence even seems to hold when correcting for height and body mass (Rushton and Ankney, 2009). Furthermore, post-mortem studies of brain weight also show an association with intellectual achievement (Broca, 1861, as cited in Rushton and Ankney, 2009) and IQ (Witelson et al., 2006). This is surprising, because the relationships of intelligence could very well be different with absolute and relative brain volume, within and across sexes, with internal and external volume measures, and with brain volume and brain weight (Cairo, 2011). Yet no matter how it is measured, larger brains show a small but reliable association with higher IQ.

Early on this finding was linked to evidence that brain volume differs between species. It is sometimes used to conclude that there are intelligence differences between species, human populations, and men and women as well (e.g., Rushton and Ankney, 2009). But are these conclusions valid? And do, in turn, species differences tell us anything about why brain size is robustly associated with intelligence among humans?

4.1. Species differences in brain size and cognitive capability

Historically, interest in a relationship between brain size and intelligence among humans was certainly sparked by the common understanding humans are the most intelligent species on earth (see Brancucci, 2012; Cairo, 2011), often defined in the comparative literature as being cognitively or behaviorally flexible when dealing with the environment (Roth and Dicke, 2005), combined with the belief that humans have exceptionally large brains. Implicitly or explicitly, this assumed cross-species relationship seems to be used as the main explanation why brain size and IQ are robustly correlated within our species. Cross-species studies of brain size and cognitive ability have long been plagued by problems how to define and measure both variables in a comparatively meaningful way, as well as by a lack of good data (Cairo, 2011; Healey and Rowe, 2007; Roth and Dicke, 2005). However, an impressive collaborative endeavor recently showed a robust association of, in particular, absolute brain size with comparable tests of cognitive self-control across 36 species, including birds, rodents, carnivores, elephants and primates (MacLean et al., 2014). It has to be kept in mind though that humans are not the species with the largest absolute brain size, as elephants and some cetaceans have multiple times larger brains. Nor do humans have the largest brains relative to body mass or size (a measure that was less well correlated with cognitive ability in MacLean et al., 2014), or an exceptionally large or neuron-rich cortex or frontal lobe relative to brain size (Barton and Venditti, 2012; Cairo, 2011; Herculano-Houzel, 2012).

The brain index most famously placing humans and primates on top is the encephalization quotient (EQ, Jerison, 1973). The EQ standardizes brain size (or sometimes weight) by body weight according to the formula:

$$\mathrm{EQ} = \frac{\text{brain size}\left(\text{cm}^3\right)}{.12 \times (\text{body weight in grams})^{.67}}$$

In this formula, 0.12 and 0.67 were empirically derived parameters that fit the results best to the cognitive differences between species. It is thus not surprising that it confirms the *a priori* assumption that humans should have the highest EQ of all species. However, more recent studies, based on better assessments of both brain size and cognitive ability, provide clear evidence that absolute brain size outperforms the EQ, as well as any other relative measure of brain size, in predicting cognitive differences between species (Deaner et al., 2007; MacLean et al., 2014). Therefore the EQ is increasingly disregarded as a useful index.

The meaningfulness of between-species comparisons of brain size relies to a certain extent on the assumption that the brains of different species are mostly up-scaled or down-scaled versions of the same brain architecture (Herculano-Houzel, 2012), implying that different brain structures always evolve together. This concerted evolution hypothesis postulates that the development of a given brain structure is constrained by the development of other brain structures, meaning that they should develop predominantly as a whole, and there is some evidence supporting this model (Finlay and Darlington, 1995).

In the last decade, the isotropic fractionator method, which provides accurate counts of different cell types in defined brain regions, allowed major progress in comparative brain studies (Herculano-Houzel and Lent, 2005). Applied to over 30 species from three mammalian orders, it revealed different brain architectures and scaling rules for different orders. Human brains exhibit precisely the same compact brain architecture as all other primates, which allows for densely packed neurons. Among primate brains, the human brain is merely an up-scaled version, with exactly the amount of neurons (about 86 billion) that can be expected from its absolute size. Still humans have the largest brain of all primates, which they could probably afford due to their nutrition-rich, cooked diet. The absolute human brain size combined with the compact primate brain architecture makes humans to the best of our knowledge the species with the largest number of neurons, which most likely plays an important role in their exceptional cognitive functions (Herculano-Houzel, 2012).

However, more neurons are unlikely the sole explanation of humans' exceptional cognitive capacities, as there is also considerable evidence for brain re-organization even within the primate order (Smaers and Soligo, 2013). Brain structures are not completely constrained to evolve in concert. The mosaic evolution hypothesis postulates that different brain structures can independently evolve in response to specific environmental demands (Barton and Harvey, 2000). Indeed, genetic correlations between different brain macrostructures in mice appear to be modest, indicating that there are no major constraints on their independent evolution (Hager et al., 2012). Based on phylogenetic analyses of brain structure data, Smaers and Soligo (2013) argue that across more than 40 million years of anthropoid primate evolution, mosaic changes contribute more to explaining neural diversity than changes in relative brain size, and different mosaic patterns are differentially selected for when brains increase or decrease in size.

Overall, absolute brain size, more so than brain size relative to body mass or EQ, is a reasonable rough indicator of a species' cognitive capabilities. Furthermore, while many aspects of the human brain are not as exceptional as previously thought and brain reorganization appeared to have played an at least as important role as size increases in the anthropoid lineage, its large absolute size combined with the compact basic primate brain architecture that let it host, as far as we know, the largest number of neurons in nature. This way, absolute brain size certainly plays some role in the evolution of the exceptional cognitive capabilities of our species. However, does this immediately translate to an explanation for the robust brain size-IQ association among humans?

4.2. *How can the brain size-IQ association be interpreted?*

While increases in brain size, as a proxy for higher neuron numbers, apparently played some causal role in the evolution of higher cognitive abilities within the primate lineage, deducing a similar causal role of brain size differences among humans for individual differences in intelligence does not seem to be warranted. For one, how much individual differences in brain size among humans are due to differences in neuron number is not undisputed. Although evidence from an older study using stereological methods suggests that neuron number is substantially positively related to the size of various cerebral macrostructures in humans (Pakkenberg and Gundersen, 1997), to our knowledge no estimate is available to date for total brain size. This would be possible using more modern methods such as the isotropic fractionator (Herculano-Houzel and Lent, 2005), which has profoundly revised our knowledge about brain cell numbers (Herculano-Houzel, 2009, 2012). Interestingly, recent unpublished data by Morterá and Herculano-Houzel (cited in Herculano-Houzel, 2009, p. 9) showed no association of brain volume with neuron counts among mice of the same age. In the light of the inconsistent pattern and comparatively small amount of evidence for this association, the strength of the contribution of cerebral neuron count must be considered inconclusive.

The problems also become apparent when looking at the largest brain size differences within the human species. Megalencephaly syndromes are rare disorders characterized by markedly enlarged brain sizes of 2.5 SD or more above the population average. Despite their unusually large brains, individuals with primary megalencephaly tend to show decreased IQ and are at risk of mental retardation (Petersson et al., 1999). In the normal range, the largest human brain size differences are between men and women. While sex differences in height-adjusted brain size have been criticized as potential statistical artifacts (Forstmeier, 2011), highlighting one reason why body size-adjusted indices of brain size (like the EQ) are even more problematic for within-species comparisons (Cairo, 2011), the sex difference for absolute brain size is clear. On average, men have a 10.8% larger total brain volume than women, a difference of 2.1 standard deviations or 131 ml. Differences are even more pronounced for intracranial and cerebrum volumes (Ruigrok et al., 2014). Some evidence suggests that these sex differences are due to higher cerebral neuron numbers in men compared to women (Pakkenberg and Gundersen, 1997; Pelvig et al., 2008). Still, despite some reports to the contrary (e.g., Jackson and Rushton, 2006; Nyborg, 2005), careful analyses of datasets not limited by range restriction clearly indicate the absence of sex differences in IQ (Dykiert et al., 2009; Flynn, 2012; Johnson et al., 2009). Thus large brains and neuron numbers do not need to translate into higher intelligence among humans.

The brains of men and women differ not only in size, but also in structure. For example, women show more white matter connections between the hemispheres and more complex cortical gyrification (Luders et al., 2004). Imaging studies indicate that men and women use their structurally different brains differently in order to reach comparable results in intelligence tests (Deary et al., 2010). For example, higher intelligence is more associated with fronto-parietal grey matter volume and temporal-occipital cortical thickness in men, and with white matter volume, grey matter volume in Broca's area, and frontal cortical thickness in women (Haier et al., 2005; Narr et al., 2007).

While the categorical nature of sex differences makes them arguably the most obvious individual differences among humans, it is reasonable to believe that a similar logic applies to brain and intelligence differences within the sexes (Deary et al., 2010). Neuroscientific studies have identified several structural and functional correlates of individual differences in intelligence beyond and independent of mere brain size, including functional parieto-frontal neuronal networks (Langeslag et al., 2014; Vakhtin et al., 2014), neuronal efficiency (Neubauer and Fink, 2009), and white matter integrity (Penke et al., 2012; Valdes Hernandez et al., 2013). Intelligence is also robustly associated with developmental stability, as approximated by body fluctuating asymmetry (Banks et al., 2010), a relationship that appears to be independent of brain size as well (Bates, 2007). Critically, many of these factors have effects on IQ that are incremental and compensatory to those of brain size, indicating that none of these factors seems to be necessary or sufficient for intelligence, with supervenience ('many-to-one'), not isomorphism (one-to-one), best describing their relationship (Kievit et al., 2011; Ritchie et al., 2015). Thus, while increases in neuron numbers, and as a consequence brain size, appeared to be one important causal factor in the evolution of human intelligence, differences in brain size among humans are only one of many interchangeable and compensatory correlates of intelligence differences—and, as suggested by the current meta-analysis, a modest one.

In addition, the direction of causality between individual differences brain size and intelligence is not completely straightforward. Of course the most intuitive interpretation is that brain size, just as neuroanatomy in general, precedes cognitive development and is thus assumed to cause intelligence differences. Indeed neuroanatomy is highly heritable and strongly genetically correlated with IQ (Posthuma et al., 2002; see also Deary et al., 2010, for a review). However, even high heritabilities do not indicate that a trait is innate or genetically determined, and even strong genetic correlations do not necessarily indicate shared underlying genetic variants in any biologically meaningful sense, but might simply indicate a role of one variable in the development of the other (Johnson et al., 2011; Solovieff et al., 2013). Also, a causal link between brain size and IQ might partly go in the opposite direction, as practice and experience can lead to volume increases in relevant brain areas (Brown et al., 2003; Steffener and Stern, 2012). Maybe as a consequence, cortical thickness in old age is predicted by childhood IQ, with childhood IQ also fully accounting for the correlation between old-age IQ and cortical thickness (Karama et al., 2014). Similarly, a famous study by Shaw et al. (2006) found that IQ was not related to cortical thickness per se, but to the plasticity of cortical thickness during childhood. So even though it is plausible that brain size is at least partially a causal factor for IQ, more research is necessary to fully unravel the interplay between genes, environment, brain anatomy and cognitive development.

5. Conclusion

In conclusion, we could show a robust, albeit modest, association between IQ and brain size in humans. Surprisingly, this association remains robust across age, intelligence domain, and sex of participants and even holds when accounting for effect inflation due to publication bias.

However, invoking the literature on cross-species comparisons and primate cognitive evolution to argue for brain size as an isomorphic proxy for human intelligence differences is not warranted. Such assumptions are often made in studies on human population differences in brain size and IQ (Rushton and Ankney, 2009), a literature that is furthermore constrained by limited data quality (Wicherts et al., 2010a,b).

Instead, brain size, likely a proxy for neuron number, is one of many neuronal factors associated with individual differences in intelligence, alongside parieto-frontal neuronal networks, neuronal efficiency, white matter integrity, cortical gyrification, overall developmental stability, and probably others. These factors seem to influence intelligence interchangeably, leading to heterogeneity in how much each factor plays a role in each individual's IQ level. The functional implications and interplay of these factors should be the focus of future research on the neuronal foundations of human intelligence differences.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neubiorev.2015.09.017.

References

Amat, J.A., Bansal, R., Whiteman, R., Haggerty, R., Royal, J., Peterson, B.S., 2008. Correlates of intellectual ability with morphology of the hippocampus and amygdala in healthy adults. Brain Cogn. 66, 105–114 (Personal communication: J.A. Amat to J.P., July 28, 2010).
This study is included in the meta-analysis.

Andreasen, N.C., Flaum, M., Swayze II, V., O'Leary, D.S., Alliger, R., Cohen, G., Ehrhardt, J., Yuh, W.T.C., 1993. Intelligence and brain structure in normal individuals. Am. J. Psychiatry 150, 130–134.
This study is included in the meta-analysis.

Antonova, E., Kumari, V., Morris, R., Halari, R., Anilkumar, A., Mehrotra, R., Sharma, T., 2005. The relationship of structural alterations to cognitive deficits in schizophrenia: A voxel-based morphometry study. Biol. Psychiatry 58, 457–467 (Personal communication: E. Antonova to M.Z., July 15, 2010).
This study is included in the meta-analysis.

Ashtari, M., Avants, B., Cyckowski, L., Cervellione, K.L., Roofeh, D., Cook, P., Gee, J., Sevy, S., Kumra, S., 2011. Medial temporal structures and memory functions in adolescents with heavy cannabis use. J. Psychiatr. Res. 45, 1055–1066.
This study is included in the meta-analysis.

Aydin, K., Seda, U., Yakut, A., Emiroglu, B., Yilmaz, F., 2012. N-acetylaspartate concentration in corpus callosum is positively correlated with intelligence in adolescents. NeuroImage 59, 1058–1064.
This study is included in the meta-analysis.

Aylward, E.H., Minshew, N.J., Field, K., Sparks, B.F., Singh, N., 2002. Effects of age on brain volume and head circumference in autism. Neurology 59, 175–183.
This study is included in the meta-analysis.

Banks, G.C., Batchelor, J.H., McDaniel, M.A., 2010. Smarter people are (a bit) more symmetrical: a meta-analysis of the relationship between intelligence and fluctuating asymmetry. Intelligence 38, 393–401.

Barton, R.A., Harvey, P.H., 2000. Mosaic evolution of brain structure in mammals. Nature 405, 1055–1058.

Barton, R.A., Venditti, C., 2012. Human frontal lobes are not relatively large. Proc. Natl. Acad. Sci. U.S.A. 110, 9001–9006.

Bates, T.C., 2007. Fluctuating asymmetry and intelligence. Intelligence 35, 41–46.

Begg, C.B., Mazumdar, M., 1994. Operating characteristics of a rank correlation test for publication bias. Biometrics 50, 1088–1101.

Betjemann, R.S., Johnson, E.P., Barnard, H., Boada, R., Filley, C.M., Filipek, P.A., et al., 2010. Genetic covariation between brain volumes and IQ, reading performance, and processing speed. Behav. Genet. 40, 135–145.
This study is included in the meta-analysis.

Bigler, E.D., 1995. Brain morphology and intelligence. Dev. Neuropsychol. 11, 377–403.
This study is included in the meta-analysis.

Blatter, D.D., Bigler, E.D., Gale, S.D., Johnson, S.C., Anderson, C.V., Burnett, B.M., et al., 1997. MR-based brain and cerebrospinal fluid measurement after traumatic brain injury: Correlation with neuropsychological outcome. Am. J. Neuroradiol. 18, 1–10.
This study is included in the meta-analysis.

Booth, T., Royle, N.A., Valdes Hernandez, M., Murray, C., Gow, A.J., Corley, J., 2015. Maximal Brain Size and Cognitive Ability Across the Life Course. University of Edinburgh, UK (Unpublished manuscript).

Borenstein, M., 2009. Effect sizes for continuous data. In: Cooper, H.M., Hedges, L.V., Valentine, J.C. (Eds.), The Handbook of Research Synthesis and Meta-Analysis. , second revised ed. Sage, New York, NY, pp. 221–235.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. Introduction to Meta-Analysis. Wiley, Chichester, UK.

Brancucci, A., 2012. Neural correlates of cognitive ability. J. Neurosci. Res. 90, 1299–1309.

Brown, J., Cooper-Kuhn, C.M., Kempermann, G., Van Praag, H., Winkler, J., Gage, F.H., Kuhn, G., 2003. Enriched environment and physical activity stimulate hippocampal but not olfactory bulb neurogenesis. Eur. J. Neurosci. 17, 2042–2046.

Burgaleta, M., Head, K., Alvarez-Linera, J., Martinez, K., Escorial, S., Haier, R., Colom, R., 2012. Sex differences in brain volume are related to specific skills, not to general intelligence. Intelligence 40, 60–68.
This study is included in the meta-analysis.

Cairo, O., 2011. External measures of cognition. Front. Hum. Neurosci. 5, 108.

Castellanos, F.X., Giedd, J.N., Berquin, P.C., Walter, J.M., Sharp, W., Tran, T., et al., 2001. Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. Arch. Gen. Psychiatry 58, 289–295.
This study is included in the meta-analysis.

Castellanos, F.X., Giedd, J.N., Eckburg, P., Marsh, W.L., Vaituzis, A.C., Kaysen, D., Hamburger, S.D., Rapoport, J.L., 1994. Quantitative morphology of the caudate nucleus in attention deficit hyperactivity disorder. Am. J. Psychiatry 151, 1791–1796.
This study is included in the meta-analysis.

Castro-Fornieles, J., Bargalló, N., Lázaro, L., Andrés, S., Falcon, C., Plana, M.T., Junqué, C., 2009. A cross-sectional and follow-up voxel-based morphometric MRI study in adolescent anorexia nervosa. J. Psychiatr. Res. 43, 331–340 (Personal communication: J. Castro-Fornieles to J.P., July 19, 2010).
This study is included in the meta-analysis.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., DeRosario, H., 2012. pwr: Basic functions for power analysis. In: R Package Version 1.1., pp. 1.

Chiang, M.-C., Reiss, A.L., Lee, A.D., Bellugi, U., Galaburda, A.M., Korenberg, J.R., et al., 2007. 3D pattern of brain abnormalities in Williams syndrome visualized using tensor-based morphometry. NeuroImage 36, 1096–1109.
This study is included in the meta-analysis.

Coffey, C.E., Ratcliff, G., Saxton, J.A., Bryan, R.N., Fried, L.P., Lucke, J.F., 2001. Cognitive correlates of human brain aging: a quantitative magnetic resonance imaging investigation. J. Neuropsychiatry Clin. Neurosci. 13, 471–485.
This study is included in the meta-analysis.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Lawrence Erlbaum, Hillsdale, NJ.

Collinson, S.L., Mackay, C.E., James, A.C., Quested, D.J., Phillips, T., Roberts, N., Crow, T.J., 2003. Brain volume, asymmetry and intellectual impairment in relation to sex in early-onset schizophrenia. Br. J. Psychiatry 183, 114–120 (Personal communication: S.L. Collinson to J.P., August 04, 2010).
This study is included in the meta-analysis.

Darwin, C., 1871. The Descent of Man. Murray, London, UK.

Deaner, R.O., Isler, K., Burkart, J., van Schaik, C., 2007. Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. Brain Behav. Evol. 70, 115–124.

Deary, I.J., Penke, L., Johnson, W., 2010. The neuroscience of human intelligence differences. Nat. Rev. Neurosci. 11, 201–211.

Debanné, M., Schaer, M., Farhoumand, R., Glaser, B., Eliez, S., 2006. Hippocampal volume reduction in 22q11.2 deletion syndrome. Neuropsychologia 44, 2360–2365 (Personal communication: M. Debanné to M.Z., July 20, 2010).
This study is included in the meta-analysis.

DeBoer, T., Wu, Z., Lee, A., Simon, T.J., 2007. Hippocampal volume reduction in children with chromosome 22q11.2 deletion syndrome is associated with cognitive impairment. Behav. Brain Funct. 3, 54 (Personal communication: T. DeBoer to M.Z., July 23, 2010).
This study is included in the meta-analysis.

Duval, S., Tweedie, R., 2000. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. J. Am. Stat. Assoc. 95, 89–98.

Dykiert, D., Gale, C.R., Deary, I.J., 2009. Are apparent sex differences in mean IQ scores created in part by sample restriction and increased male variance? Intelligence 37, 42–47.

Ebner, F., Tepest, R., Dani, I., Pfeiffer, U., Schulze, T.G., Rietschel, M., et al., 2008. The hippocampus in families with schizophrenia in relation to obstetric complications. Schizophr. Res. 104, 71–78 (Personal communication: T. Schneider-Axmann to J.P., July 28, 2010).
This study is included in the meta-analysis.

Egan, V., Chiswick, A., Santosh, C., Naidu, K., Rimmington, J.E., Best, J.J.K., 1994. Size isn't everything: a study of brain volume, intelligence and auditory evoked potentials. Pers. Individ. Dif. 17, 357–367.
This study is included in the meta-analysis.

Egan, V., Wickett, J.C., Vernon, P.A., 1995. Brain size and intelligence: erratum, addendum, and correction. Pers. Individ. Dif. 19, 113–115.
This study is included in the meta-analysis.

Finby, N., Kraft, E., 1972. The aging skull: comparative roentgen study. Clin. Radiol. 23, 410–414.

Fine, J.G., Semrud-Clikeman, M., Keith, T.Z., Stapleton, L.M., Hynd, G.W., 2007. Reading and the corpus callosum: an MRI family study of volume and area. Neuropsychology 21, 235–241 (Personal communication: J.G. Fine to M.Z., October 8, 2010).
This study is included in the meta-analysis.

Finlay, B.L., Darlington, R.B., 1995. Linked regularities in the development and evolution of mammalian brains. Science 268, 1578–1584.

Flashman, L.A., Andreasen, N.C., Flaum, M., Swayze II, V.W., 1998. Intelligence and regional brain volumes in normal controls. Intelligence 25, 149–160.
This study is included in the meta-analysis.

Flynn, J.R., 2012. Are we Getting Smarter? Rising IQ in the Twenty-First Century. Cambridge University Press, Cambridge, UK, pp. 132–158.

Forstmeier, W., 2011. Women have relatively larger brains than men: a comment on the misuse of general linear models in the study of sexual dimorphism. Anat. Rec. 294, 1856–1863.

FrangouF S., Chitins, X., Williams, C.R., 2004. Mapping IQ and gray matter density in healthy young people. NeuroImage 23, 800–805.
This study is included in the meta-analysis.

Galton, F., 1888. Head growth in university students at the University of Cambridge. Nature 38, 14–15.

Garde, E., Mortensen, E.L., Krabbe, K., Rostrup, E., Larsson, H.B.W., 2000. Relation between age-related decline in intelligence and cerebral white-matter hyperintensities in healthy octogenarians: a longitudinal study. Lancet 356, 628–634 (Personal communication to M.A. McDaniel, obtained through the meta-analysis of McDaniel (2005)).
This study is included in the meta-analysis.

Giedd, J.N., 2003. (Personal communication to M.A. McDaniel, obtained through the meta-analysis of McDaniel (2005).
This study is included in the meta-analysis.

Gignac, G., Vernon, P.A., Wickett, J.C., 2003. Factors influencing the relationship between brain size and intelligence. In: Nyborg, H. (Ed.), The Scientific Study of General Intelligence: Tribute to Arthur R. Jensen. Elsevier, Oxford, UK, pp. 93–106.

Gould, S.J., 1981. The Mismeasure of Man. Norton, New York, NY.

Gur, R.C., Turetsky, B.I., Matsui, M., Yan, M., Bilker, W., Hughett, P., Gur, R.E., 1999. Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. J. Neurosci. 19, 4065–4072.
This study is included in the meta-analysis.

Hager, R., Lu, L., Rosen, G.D., Williams, R.W., 2012. Genetic architecture supports mosaic brain evolution and independent brain-body size regulation. Nat. Commun. 3, 1079.

Hahn, S., Williamson, P.R., Hutton, J.L., 2002. Investigation of within-study selective reporting in clinical research: follow-up applications submitted to local research. J. Eval. Clin. Pract. 8, 353–359.

Haier, R.J., Chueh, D., Touchette, P., Lott, I., Buchsbaum, M.S., MacMillan, D., et al., 1995. Brain size and cerebral glucose metabolic rate in nonspecific mental retardation and down syndrome. Intelligence 20, 191–210.
This study is included in the meta-analysis.

Haier, R.J., Jung, R.E., Yeo, R.A., Head, K., Alkire, M.T., 2005. The neuroanatomy of general intelligence: sex matters. NeuroImage 25, 320–327.

Harvey, I., Persaud, R., Ron, M.A., Baker, G., Murray, R.M., 1994. Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls. Psychol. Med. 24, 689–699.
This study is included in the meta-analysis.

Healey, S.D., Rowe, C., 2007. A critique of comparative studies of brain size. Proc. R. Soc., B 274, 453–464.

Herculano-Houzel, S., 2009. The human brain in numbers: a linearly scaled-up primate brain. Front. Hum. Neurosci. 3, 31.

Herculano-Houzel, S., 2012. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. Proc. Natl. Acad. Sci. U.S.A. 109, 10661–10668.

Herculano-Houzel, S., Lent, R., 2005. Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. J. Neurosci. 25, 2518–2521.

Hermann, B.P., Seidenberg, M., Bell, B., 2002. The neurodevelopmental impact of childhood onset temporal lobe epilepsy on brain structure and function and the risk of progressive cognitive effects. Prog. Brain Res. 135, 429–439 (Personal communication: B.P. Hermann to J.P., September 15, 2010).
This study is included in the meta-analysis.

Hogan, M.J., Staff, R.T., Bunting, B.P., Murray, A.D., Ahearn, T.S., Deary, I.J., Whalley, L.J., 2010. Cerebellar brain volume accounts for variance in cognitive performance in older adults. Cortex 47, 441–450 (Personal communication: R.T. Staff to J.P., July 27, 2010).
This study is included in the meta-analysis.

Hunt, E., 2011. Human Intelligence. Cambridge University Press, New York, NY, pp. 154.

Ioannidis, J.P.A., 1998. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. JAMA 279, 281–286.

Ioannidis, J.P.A., 2011. Excess significance bias in the literature on brain volume abnormalities. Arch. Gen. Psychiatry 68, 773–780.

Ioannidis, J.P.A., Trikalinos, T.A., 2007. An exploratory test for an excess of significant findings. Clin. Trials 4, 245–253.

Isaacs, E.B., Edmonds, C.J., Chong, W.K., Lucas, A., Morley, R., Gadian, D.G., 2004. Brain morphometry and IQ measurements in preterm children. Brain 127, 2595–2607 (Personal communication: E.B. Isaacs to M.Z., July 22, 2010).
This study is included in the meta-analysis.

Isaacs, E.B., Fischl, B.R., Quinn, B.T., Chong, W.K., Gadian, D.G., Lucas, A., 2010. Impact of breast milk on intelligence quotient, brain size, and white matter development. Pediatr. Res. 67, 357–362.
This study is included in the meta-analysis.

Isaacs, E.B., Lucas, A., Chong, W.K., Wood, S.J., Jacobsen, C.L., Marshall, C., Vargha-Khadem, F., Gadian, D.G., 2000. Hippocampal volume and everyday memory in children of very low birth weight. Pediatr. Res. 47, 713–720 (Personal communication: E.B. Isaacs to M.Z. July 22, 2010).
This study is included in the meta-analysis.

Ivanovic, D.M., Leiva, B.P., Castro, C.G., Olivares, M.G., Jansana, J.M.M., Castro, V.G., et al., 2004a. Brain development parameters and intelligence in Chilean high school graduates. Intelligence 32, 461–479.
This study is included in the meta-analysis.

Ivanovic, D.M., Leiva, B.P., Pérez, H.T., Olivares, M.G., Díaz, N.S., Urrutia, M.S.C., et al., 2004b. Head size and intelligence, learning, nutritional status and brain development: head, IQ, learning, nutrition and brain. Neuropsychologia 42, 1118–1131.
This study is included in the meta-analysis.

Jaccard, J., Turrisi, R., 2003. Interaction Effects in Multiple Regression, second ed. Sage, Thousand Oaks, CA.

Jackson, D.N., Rushton, J.P., 2006. Males have greater *g*: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test. Intelligence 34, 479–486.

Jensen, A.R., 1982. The debunking of scientific fossils and straw persons. Contemp. Educ. Rev. 1, 121–135.

Jensen, A.R., 1998. The *g* Factor: The Science of Mental Ability. Praeger, Westport, CT.

Jerison, H.J., 1973. Evolution of the Brain and Intelligence. Academic Press, New York, NY.

Jones, P.B., Harvey, I., Lewis, S.W., Toone, B.K., van Os, J., Williams, M., Murray, R.M., 1994. Cerebral ventricle dimensions as risk factors for schizophrenia and affective psychosis: an epidemiological approach to analysis. Psychol. Med. 24, 995–1011.
This study is included in the meta-analysis.

Johnson, W., Carothers, A., Deary, I.J., 2009. A role for the *X* chromosome in sex differences in variability in general intelligence? Perspect. Psychol. Sci. 4, 598–611.

Johnson, W., Penke, L., Spinath, F.M., 2011. Heritability in the era of molecular genetics: some thoughts for understanding genetic influences on behavioral traits. Eur. J. Personal. 25, 254–266.

Karama, S., Bastin, M.E., Murray, C., Royle, N.A., Penke, L., Muñoz Maniega, S., et al., 2014. Childhood IQ explains associations between IQ and brain cortical thickness in old age. Mol. Psychiatry 19, 555–559.

Kareken, D.A., Gur, R.C., Mozley, P.D., Mozley, L.H., Saykin, A.J., Shtasel, D.L., Gur, R.E., 1995. Cognitive functioning and neuroanatomic volume measures in schizophrenia. Neuropsychology 9, 211–219.
This study is included in the meta-analysis.

Kesler, S.R., Adams, H.F., Blasey, C.M., Bigler, E.D., 2003. Premorbid intellectual functioning, education, and brain size in traumatic brain injury: An investigation of the cognitive reserve hypothesis. Appl. Neuropsychol. 10, 153–162.
This study is included in the meta-analysis.

Kievit, R.A., Romeijn, J.W., Waldorp, L.J., Wicherts, J.M., Scholte, H.S., Borsboom, D., 2011. Mind the Gap: a psychometric approach to the reduction problem. Psychol. Inq. 22, 76–87 (Personal communication: R.A. Kievit to J.P., December 12, 2011).
This study is included in the meta-analysis.

Kumra, S., Giedd, J.N., Vaituzis, A.C., Jacobsen, L.K., McKenna, K., Bedwell, J., et al., 2000. Childhood-onset psychotic disorders: magnetic resonance imaging of volumetric differences in brain structure. Am. J. Psychiatry 157, 1467–1474.
This study is included in the meta-analysis.

Lange, N., Froimowitz, M.P., Bigler, E.D., Lainhart, J.E., Brain Development Cooperative Group, 2010. Associations between IQ, total and regional brain volumes, and demography in a large normative sample of healthy children and adolescents. Dev. Neuropsychol. 35, 296–317.
This study is included in the meta-analysis.

Langeslag, S.J.E., Schmidt, M., Ghassabian, A., Jaddoe, V.W., Hofman, A., van der Lugt, A., et al., 2014. Functional connectivity between parietal and frontal brain regions and intelligence in young children: the Generation R study. Hum. Brain Mapp. 34, 3299–3307.

LawsonF J.A., Vogrin, S., Bleasel, A.F., Cook, M.J., Bye, A.M.E., 2000. Cerebral and cerebellar volume reduction in children with intractable epilepsy. Epilepsia 41, 1456–1462.
This study is included in the meta-analysis.

Leonard, C.M., Kuldau, J.M., Breier, J.I., Zuffante, P.A., Gautier, E.R., Heron, D.-C., et al., 1999. Cumulative effect of anatomical risk factors for schizophrenia: an MRI study. Biol. Psychiatry 46, 374–382.
This study is included in the meta-analysis.

Light, R.J., Pillemer, D.B., 1984. Summing Up: The Science of Reviewing Research. Harvard University Press, Cambridge, MA.

Lodygensky, G.A., Rademaker, K., Zimine, S., Gex-Fabry, M., Lieftink, A.F., Lazeyras, F., et al., 2005. Structural and functional brain development after hydrocortisone treatment for neonatal chronic lung disease. Pediatrics 116, 1–7 (Personal communication: G.A. Lodygensky to M.Z., September 1, 2010).
This study is included in the meta-analysis.

Lombroso, C., 1864. Genio e follia, in rapporto alla medicina legale, alla critica ed alla storia (Genius and Insanity in their Relation to Law, Critique and History). Giuseppe Chiusi, Milano, Italy.

Luders, E., Narr, K.L., Bilder, R.M., Thompson, P.M., Szeszko, P.R., Hamilton, L., Togaa, A.W., 2007. Positive correlations between corpus callosum thickness and intelligence. NeuroImage 37, 1457–1464.
This study is included in the meta-analysis.

Luders, E., et al., 2004. Gender differences in cortical complexity. Nature Neurosci. 7, 799–800.

MacLaren, J., Han, Z., Vos, S.B., Fischbein, N., Bammer, R., 2014. Reliability of brain volume measurements: a test–retest dataset. Sci. Data 1, 140037.

MacLullich, A.M.J., Ferguson, K.J., Deary, I.J., Seckl, J.R., Starr, J.M., Wardlaw, J.M., 2002. Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. Neurology 59, 169–174.
This study is included in the meta-analysis.

May, H., Peled, N., Dar, G., Abbas, J., Hershkovitz, I., 2011. Hyperostosis frontalis interna: what does it tell us about our health? Am. J. Hum. Biol. 23, 392–397.

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

MacLean, E.L., Hare, B., Nunn, C.L., Addessi, E., Amici, F., Anderson, R.C., et al., 2014. The evolution of self-control. Proc. Nat. Acad. Sci. U.S.A. 111, E2140–E2148.

Miller, G.F., Penke, L., 2007. The evolution of human intelligence and the coefficient of additive genetic variance in human brain size. Intelligence 35, 97–114.

Miller, J.F., Couch, J., Schwenk, K., Long, M., Towler, S., Theriaque, D.W., et al., 2009. Early childhood obesity is associated with compromised cerebellar development. Dev. Neuropsychol. 34, 272–283.
This study is included in the meta-analysis.

Mori, E., Hirono, N., Yamashita, H., Imamura, T., Ikejiri, Y., Ikeda, M., et al., 1997. Premorbid brain size as a determinant of reserve capacity against intellectual decline in Alzheimer's disease. Am. J. Psychiatry 154, 18–24.
This study is included in the meta-analysis.

Murdoch, K., Sullivan, L.R., 1923. A contribution to the study of mental and physical measurements in normal children. Am. Phys. Educ. Rev. 28, 209–330.

Nakamura, M., Nestor, P.G., McCarley, R.W., Levitt, J.J., Hsu, L., Kawashima, T., Niznikiewicz, M., Shenton, M.E., 2007. Altered orbitofrontal sulcogyral pattern in schizophrenia. Brain 130, 693–707 (Personal communication: M. Nakamura to J.P., July 20, 2010).
This study is included in the meta-analysis.

Narr, K.L., Woods, R.P., Thompson, P.M., Szeszko, P., Robinson, D., Dimtcheva, T., et al., 2007. Relationships between IQ and regional cortical gray matter thickness in healthy adults. Cereb. Cortex 17, 2163–2171.

Neubauer, A.C., Fink, A., 2009. Intelligence and neural efficiency. Neurosci. Biobehav. Rev. 33, 1004–1023.

Nosarti, C., Al-Asady, M.H.S., Frangou, S., Stewart, A.L., Rifkin, L., Murray, R.M., 2002. Adolescents who were born very preterm have decreased brain volumes. Brain 125, 1616–1623.
This study is included in the meta-analysis.

Nyborg, H., 2005. Sex-related differences in general intelligence g, brain size, and social status. Pers. Individ. Dif. 39, 497–509.

Pakkenberg, B., Gundersen, H.J.G., 1997. Neocortical neuron number in humans: effect of sex and age. J. Comp. Neurol. 384, 312–320.

Paradiso, S., Andreasen, N.C., O'Leary, D.S., Arndt, S., Robinson, R.G., 1997. Cerebellar size and cognition: correlations with IQ, verbal memory and motor dexterity. Neuropsychiatry Neuropsychol. Behav. Neurol. 10, 1–8.
This study is included in the meta-analysis.

Pelvig, D.P., Pakkenberg, H., Stark, A.K., Pakkenberg, B., 2008. Neocortical glial cell numbers in human brains. Neurobiol. Aging 29, 1754–1762.

Penke, L., Muñoz Maniega, S., Bastin, M.E., Valdés Hernandéz, M.C., Murray, C., Royle, N.A., et al., 2012. Brain matter tract integrity as a neural foundation of general intelligence. Mol. Psychiatry 17, 1026–1030.

Pennington, B.F., Filipek, P.A., Lefty, D., Chhabildas, N., Kennedy, D.N., Simon, J.H., et al., 2000. A twin MRI study of size variations in the human brain. J. Cogn. Neurosci. 12, 223–232.
This study is included in the meta-analysis.

Petersson, S., Pedersen, N.L., Schalling, M., Lavebratt, C., 1999. Primary megalencephaly at birth and low intelligence level. Neurology 53, 1254–1259.

Pietschnig, J., Voracek, M., Formann, A.K., 2010. Mozart effect—Shmozart effect: a meta-analysis. Intelligence 38, 314–323.

Pigott, T.D., 2009. Handling missing data. In: Cooper, H.M., Hedges, L.V., Valentine, J.C. (Eds.), The Handbook of Research Synthesis and Meta-Analysis. , second revised ed. Sage, New York, NY, pp. 399–416.

Posthuma, D., De Geus, E.J.C., Baaré, W.F.C., Hulshoff Pol, H.E., Kahn, R.S., Boomsma, D.I., 2002. The association between brain volume and intelligence is of genetic origin. Nat. Neurosci. 5, 83–84.

Qiu, A., Crocetti, D., Adler, M., Mahone, E.M., Denckla, M.B., Miller, M.I., Mostofsky, S.H., 2009. Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. Am. J. Psychiatry 166, 74–82 (Personal communication: A. Qiu to M.Z., July 16, 2010).
This study is included in the meta-analysis.

R Development Core Team, 2014. R: A Language and Environment for Statistical Computing (Version 3.1.1) [Computer Software]. R Foundation for Statistical Computing, Vienna, Austria.

Raz, N., Lindenberger, U., Ghisletta, P., Rodrigue, K.M., Kennedy, K.M., Acker, J.D., 2008. Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors. Cereb. Cortex 18, 718–726 (Personal communication: N. Raz to M.Z., July 15, 2010).
This study is included in the meta-analysis.

Raz, N., Torres, I.J., Briggs, S.D., Spencer, W.D., Thornton, A.E., Loken, W.J., et al., 1995. Selective neuroanatomic abnormalities in Down's syndrome and their cognitive correlates: evidence from MRI morphometry. Neurology 45, 357–366.
This study is included in the meta-analysis.

Raz, N., Torres, I.J., Spencer, W.D., Millman, D., Baertschi, J.C., Sarpel, G., 1993. Neuroanatomical correlates of age-sensitive and age-invariant cognitive abilities: an in vivo MRI investigation. Intelligence 17, 407–422.
This study is included in the meta-analysis.

Reiss, A.L., Abrams, M.T., Greenlaw, R., Freund, L., Denckla, M.B., 1995. Neurodevelopmental effects of the FMR-1 full mutation in humans. Nat. Med. 1, 159–167.
This study is included in the meta-analysis.

Reiss, A.L., Abrams, M.T., Singer, H.S., Ross, J.L., Denckla, M.B., 1996. Brain development, gender and IQ in children: a volumetric imaging study. Brain 119, 1763–1774.
This study is included in the meta-analysis.

Ritchie, S.J., Booth, T., Valdés Hernández, M.D.C., Corley, J., Muñoz Maniega, S., Gow, A.J., et al., 2015. Beyond a bigger brain: multivariable structural brain imaging and intelligence. Intelligence 51, 47–56.

Rojas, D.C., Peterson, E., Winterrowd, E., Reite, M.L., Rogers, S.J., Tregellas, J.R., 2006. Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. BMC Psychiatry 6, 56 (Personal communication: D.C. Rojas to M.Z., September 2, 2010).
This study is included in the meta-analysis.

Rojas, D.C., Smith, J.A., Benkers, T.L., Camou, S.L., Reite, M.L., Rogers, S.J., 2004. Hippocampus and amygdala volumes in parents of children with autistic disorder. Am. J. Psychiatry 161, 2038–2044 (Personal communication: D.C. Rojas to M.Z., September 2, 2010).
This study is included in the meta-analysis.

Roth, G., Dicke, U., 2005. Evolution of the brain and intelligence. Trends Cogn. Sci. 9, 250–257.

Rothstein, H.R., Sutton, A.J., Borenstein, M., 2005. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Wiley, Chichester, UK.

Royle, N.A., Booth, T., Valdes Hernandez, M.C., Penke, L., Murray, C., Gow, A.J., et al., 2013. Estimated maximal and current brain volume predict cognitive ability in old age. Neurobiol. Aging 34, 2726–2733.
This study is included in the meta-analysis.

Ruigrok, A.N., Salimi-Khorshidi, G., Lai, M.C., Baron-Cohen, S., Lombardo, M.V., Tait, R.J., Suckling, J., 2014. A meta-analysis of sex differences in human brain structure. Neurosci. Biobehav. Rev. 39, 34–50.

Rushton, J.P., Ankney, C.D., 1996. Brain size and cognitive ability: correlations with age, sex, social class, and race. Psychon. Bull. Rev. 3, 21–36.

Rushton, J.P., Ankney, C.D., 2000. Size matters: a review and new analyses of racial differences in cranial capacity and intelligence that refute Kamin and Omari. Pers. Individ. Dif. 29, 591–620.

Rushton, J.P., Ankney, C.D., 2009. Whole brain size and general mental ability: a review. Int. J. Neurosci. 119, 691–731.

Schoenemann, P.T., Budinger, T.F., Sarich, V.M., Wang, W.S.-Y., 2000. Brain size does not predict general cognitive ability within families. Proc. Natl. Acad. Sci. U.S.A. 97, 4932–4937.
This study is included in the meta-analysis.

Schooler, J., 2011. Unpublished results hide the decline effect. Nature 470, 437.

Schottenbauer, M.A., Momenan, R., Kerick, M., Hommer, D.W., 2007. Relationships among aging, IQ, and intracranial volume in alcoholics and control subjects. Neuropsychology 21, 337–345 (Personal communication: D.W. Hommer to M.Z., July 15, 2010).
This study is included in the meta-analysis.

Schumann, C.M., Hamstra, J., Goodlin-Jones, B.L., Kwon, H., Reiss, A.L., Amaral, D.G., 2007. Hippocampal size positively correlates with verbal IQ in male children. Hippocampus 17, 486–493.
This study is included in the meta-analysis.

Shapleske, J., Rossell, S.L., Chitnis, X.A., Suckling, J., Simmons, A., Bullmore, E.T., Woodruff, P.W.R., David, A.S., 2002. A computational morphometric MRI study of schizophrenia: effects of hallucinations. Cereb. Cortex 12, 1331–1341.
This study is included in the meta-analysis.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N.E.E.A., et al., 2006. Intellectual ability and cortical development in children and adolescents. Nature 440, 676–679.

Shenkin, S.D., Rivers, C.S., Deary, I.J., Starr, J.M., Wardlaw, J.M., 2009. Maximum (prior) brain size, not atrophy, correlates with cognition in community-dwelling older people: a cross-sectional neuroimaging study. BMC Geriatr. 9, 12.
This study is included in the meta-analysis.

Simmons, K., 1942. Cranial capacities by both plastic and water techniques with linear measurements of the reserve collection; white and negro. Hum. Biol. 14, 473–498.

Smaers, J.B., Soligo, C., 2013. Brain reorganization, not relative brain size, primarily characterizes anthropoid brain evolution. Proc. Biol. Sci. 280, 20130269.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., Smoller, J.W., 2013. Pleiotropy in complex traits: Challenges and strategies. Nat. Rev. Genet. 14, 483–495.

Staff R.T., 2002. Personal communication to Michael A. McDaniel on November 11, 2002, cited in McDaniel, 2005.

Staff, R.T., Murray, A.D., Deary, I.J., Whalley, L.J., 2006. Generality and specificity in cognitive aging: a volumetric brain analysis. NeuroImage 30, 1433–1440 (Personal communication: R.T. Staff to M.Z., July 27, 2010).
This study is included in the meta-analysis.

Steffener, J., Stern, Y., 2012. Exploring the neural basis of cognitive reserve in aging. Biochim. Biophys. Acta 1822, 467–473.

Sterne, J.A.C., Egger, M., 2005. Regression methods to detect publication and other bias in meta-analysis. In: Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Wiley, New York, NY, pp. 99–110.

Sterne, J.A.C., Gavaghan, D., Egger, M., 2000. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J. Clin. Epidemiol. 53, 1119–1129.

Tan, Ü., Tan, M., Polat, P., Ceylan, Y., Suma, S., Okur, A., 1999. Magnetic resonance imaging brain size/IQ relations in Turkish university students. Intelligence 27, 83–92.
This study is included in the meta-analysis.

Tate, D.F., Neeley, E.S., Norton, M.C., Tschanz, J.T., Miller, M.J., Wolfson, L., et al., 2011. Intracranial volume and dementia: some evidence in support of the cerebral reserve hypothesis. Brain Res. 1385, 151–162.
This study is included in the meta-analysis.

Thoma, R.J., Yeo, R.A., Gangestad, S.W., Halgren, E., Sanchez, N.M., Lewine, J.D., 2005. Cortical volume and developmental instability are independent predictors of general intellectual ability. Intelligence 33, 27–38.
This study is included in the meta-analysis.

Thompson, P.M., Cannon, T.D., Narr, K.L., Erp, T.V., Poutanen, V.-P., Huttunen, M., et al., 2001. Genetic influences on brain structure. Nat. Neurosci. 4, 1253–1258.

Tiedemann, F., 1836. On the brain of the Negro, compared with that of the European and the orang-outang. Philos. Trans. 196, 497–527.

Toulopoulou, T., Grech, A., Morris, R.G., Schulze, K., McDonald, C., Chapple, B., Rabe-Hesketh, S., Murray, R.M., 2004. The relationship between volumetric brain changes and cognitive function: a family study on schizophrenia. Biol. Psychiatry 56, 447–453.
This study is included in the meta-analysis.

Vakhtin, A.A., Ryman, S.G., Flores, R.A., Jung, R.E., 2014. Functional brain networks contributing to the Parieto-Frontal Integration Theory of intelligence. NeuroImage 103, 349–354.

Valdes Hernandez, M., Booth, T., Murray, C., Gow, A.J., Penke, L., Morris, Z., et al., 2013. Brain white matter damage in aging and cognitive ability in youth and old age. Neurobiol. Aging 34, 2740–2747.

Van Valen, L., 1974. Brain size and intelligence in man. Am. J. Phys. Anthropol. 40, 417–424.

Van Leeuwen, M., Peper, J.S., VandenBerg, S.M., Brouwer, R.M., Hulshoff-Pol, H.E., Kahn, R.S., Boomsma, D.I., 2008. A genetic analysis of brain volumes and IQ in children. Intelligence 37, 181–191.
This study is included in the meta-analysis.

Vernon, P.A., Wickett, J.C., Bazana, P.G., Stelmack, R.M., 2000. The neuropsychology and psychophysiology of human intelligence. In: Sternberg, R.J. (Ed.), Handbook of Intelligence. Cambridge University Press, New York, NY, pp. 245–264.

Vevea, J.L., Woods, C.M., 2005. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. Psychol. Methods 10, 428–443.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36, 1–48.

Voelbel, G.T., Bates, M.E., Buckman, J.F., Pandina, G., Hendren, R.L., 2006. Caudate nucleus volume and cognitive performance: are they related in childhood psychopathology? Biol. Psychiatry 60, 942–950 (Personal communication: M.E. Bates to M.Z., July 28, 2010).
This study is included in the meta-analysis.

Waiter, G.D., Williams, J.H.G., Murray, A.D., Gilchrist, A., Perrett, D.I., Whiten, A., 2004. A voxel-based investigation of brain structure in male adolescents with autistic spectrum disorder. NeuroImage 22, 619–625 (Personal communication: G.D. Waiter to M.Z., August 30, 2010).
This study is included in the meta-analysis.

Wallace, G.L., Lee, N.R., Prom-Wormley, E.C., Medland, S.E., Lenroot, R.K., Clasen, L.S., et al., 2010. A bivariate twin study of regional brain volumes and verbal and nonverbal intellectual skills during childhood and adolescence. Behav. Genet. 40, 125–134.
This study is included in the meta-analysis.

Warwick, M.M., Doody, G.A., Lawrie, S.M., Kestelman, J.N., Best, J.J.K., Johnstone, E.C., 1999. Volumetric magnetic resonance imaging study of the brain in subjects with sex chromosome aneuploidies. J. Neurol. Neurosurg. Psychiatry 66, 628–632.
This study is included in the meta-analysis.

Wechsler, D., 1981. Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R). The Psychological Corporation, San Antonio, TX.

Weniger, G., Lange, C., Sachsse, U., Irle, E., 2009. Reduced amygdala and hippocampus size in trauma-exposed women with borderline personality disorder and without posttraumatic stress disorder. J. Psychiatry Neurosci. 34, 383–388 (Personal communication: E. Irle to M.Z., July 25, 2010).
This study is included in the meta-analysis.

Wicherts, J.M., Borsboom, D., Dolan, C.V., 2010a. Evolution, brain size, and the national IQ of peoples around 3000 years B.C. Pers. Individ. Dif. 48, 104–106.

Wicherts, J.M., Dolan, C.V., van der Maas, H.L.J., 2010b. A systematic literature review of the average IQ of Sub-Saharan Africans. Intelligence 38, 1–20.

Wickett, J.C., Vernon, P.A., Lee, D.H., 1994. In vivo brain size, head perimeter, and intelligence in a sample of healthy adult females. Pers. Individ. Dif. 16, 831–838.

Wickett, J.C., Vernon, P.A., Lee, D.H., 2000. Relationships between factors of intelligence and brain volume. Pers. Individ. Dif. 29, 1095–1122.
This study is included in the meta-analysis.

Willerman, L., Schultz, R., Rutledge, J.N., Bigler, E.D., 1991. In vivo brain size and intelligence. Intelligence 15, 223–228.
This study is included in the meta-analysis.

Witelson, S.F., Beresh, H., Kigar, D.L., 2006. Intelligence and brain size in 100 postmortem brains: sex, lateralization and age factors. Brain 126, 386–398.
This study is included in the meta-analysis.

Wolf, H., Kruggel, F., Hensel, A., Wahlund, L.O., Arendt, T., Gertz, H.J., 2003. The relationship between head size and intracranial volume in elderly subjects. Brain Res. 973, 74–80.

Wozniak, J.R., Mueller, B.A., Chang, P.-N., Muetzel, R.L., Caros, L., Lim, K.O., 2006. Diffusion tensor imaging in children with fetal alcohol spectrum disorders. Alcohol Clin. Exp. Res. 30, 1799–1806 (Personal communication: J.R. Wozniak to M.Z., September 2, 2010).
This study is included in the meta-analysis.

Yeo, R.A., Turkheimer, E., Raz, N., Bigler, E.D., 1987. Volumetric asymmetries of the human brain: Intellectual correlates. Brain Cogn. 6, 15–23.
This study is included in the meta-analysis.

Yurgelun-Todd, D.A., Killgore, W.D.S., Cintron, C.B., 2003. Cognitive correlates of medial temporal lobe development across adolescence: a magnetic resonance imaging study. Percept. Mot. Skills 96, 3–17.
This study is included in the meta-analysis.

Zeegers, M., Hulshoff-Pol, H., Durston, S., Nederveen, H., Schnack, H., VanDaalen, E., et al., 2009. No differences in MR-based volumetry between 2- and 7-year-old children with autism spectrum disorder and developmental delay. Brain Dev. 21, 725–730.
This study is included in the meta-analysis.