Intelligence 64 (2017) 18–29

Contents lists available at ScienceDirect

Intelligence

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

Brain volume and intelligence: The moderating role of intelligence measurement quality

Gilles E. Gignac[a,*], Timothy C. Bates[b]

[a] University of Western Australia, Australia
[b] The University of Edinburgh, Scotland

ARTICLE INFO

Keywords:
Meta-analysis
Meta-regression
Brain volume
Intelligence
P-curve

ABSTRACT

A substantial amount of empirical research has estimated the association between brain volume and intelligence. The most recent meta-analysis (Pietschnig, Penke, Wicherts, Zeiler, & Voracek, 2015) reported a correlation of .24 between brain volume and intelligence – notably lower than previous meta-analytic estimates. This headline meta-analytic result was based on a mixture of samples (healthy and clinical) and sample correlations not corrected for range restriction. Additionally, the role of IQ assessment quality was not considered. Finally, evidential value of the literature was not formally evaluated. Based on the results of our meta-analysis of the Pietschnig et al.'s sample data, the corrected correlation between brain volume and intelligence in healthy adult samples was $r = .31$ ($k = 32$; $N = 1758$). Furthermore, the quality of intelligence measurement was found to moderate the effect between brain volume and intelligence ($b = .08$, $p = .028$). Investigations that used 'fair', 'good', and 'excellent' measures of intelligence yielded corrected brain volume and intelligence correlations of .23 ($k = 9$; $N = 547$), .32 ($k = 10$; $N = 646$), and .39 ($k = 13$; $N = 565$), respectively. The Henmi/Copas adjusted confidence intervals, the p-uniform results, and the p-curve results failed to suggest evidence of publication bias and/or p-hacking. The results were interpreted to suggest that the association between in vivo brain volume and intelligence is arguably best characterised as $r \approx .40$. Researchers are encouraged to consider intelligence measurement quality in future meta-analyses, based on the guidelines provided in this investigation.

1. Introduction

The topic of brain size and its possible association with intelligence, both within and between species, has been the subject of a substantial amount of research and debate (Mackintosh, 2011). Recently, Pietschnig et al. (2015) reported a meta-analytic observed correlation between human brain volume and intelligence of $r = .24$, based on 120 sample correlations ($N = 6778$). A limitation associated with the Pietschnig et al. (2015) investigation is that it did not provide an estimate of the association between brain volume and intelligence corrected for range restriction. Additionally, Pietschnig et al. (2015) did not explore the possibility that quality of intelligence measurement may moderate the magnitude of the association between brain volume and intelligence. Finally, Pietschnig et al. (2015) did not formally evaluate the evidential value of the reported research via a p-curve analysis.

Consequently, the purpose of this investigation was to extend the Pietschnig et al. (2015) meta-analysis in three ways. First, to estimate the correlation between in vivo human brain volume and intelligence based on correlations associated with relatively few artefacts, i.e., correlations derived from healthy adult samples and corrected for range restriction. Secondly, to develop a guide to help classify the quality of general intelligence measurement, in order to test the hypothesis that there is a positive association between intelligence test measurement quality and the magnitude of effect sizes reported across empirical investigations. Finally, to conduct a p-curve analysis to evaluate the reported brain volume and intelligence statistically significant correlations for evidential value.

1.1. Brain volume and intelligence: quantitative reviews

The association between in vivo brain volume and intelligence has been reviewed quantitatively several times over the years. More than a decade ago, Gignac, Vernon, and Wickett (2003) estimated the observed correlation between brain volume and IQ based on 14 samples ($N = 858$), all of which were derived from peer reviewed publications. Gignac et al. (2003) reported an N-weighted mean correlation of .37 between brain volume and intelligence. In six of the 14 investigations included in the meta-analysis, the IQ score standard deviations were

---

[*] Corresponding author at: School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia.
E-mail address: gilles.gignac@uwa.edu.au (G.E. Gignac).

http://dx.doi.org/10.1016/j.intell.2017.06.004
Received 6 December 2016; Received in revised form 25 May 2017; Accepted 26 June 2017

available. Consequently, Gignac et al. (2003) also reported an *N*-weighted mean corrected correlation of .43 between brain volume and IQ.[1]

McDaniel (2005) revisited the in vivo brain volume and intelligence association by conducting a more comprehensive meta-analysis than that of Gignac et al. (2003). McDaniel's (2005) inclusion criteria were the following: clinically healthy samples; total brain volume measurement; and well-established measures of intelligence (Wechsler scales; Raven's; but not the National Adult Reading Test, for example). Based on the samples which met those criteria ($k = 37$; $N = 1530$), McDaniel (2005) reported an observed correlation of $r = .29$ between brain volume and global intelligence. Additionally, McDaniel (2005) reported a range restricted corrected correlation of $r = .33$. Thus, the corrected correlation reported by McDaniel (2005) was smaller than the corrected correlation reported by Gignac et al. (2003; $r = .43$).

It is noteworthy that McDaniel (2005) found that the mean correlation between brain volume and intelligence was larger for adults than for children. For example, the brain volume and intelligence corrected correlation for adult males was estimated at $r = .38$, whereas the same correlation for male children was estimated at $r = .22$. McDaniel (2005) did not speculate as to why the effects may have been larger for adults in comparison to children. It is suggested here that both incomplete neurophysiological maturation and individual differences in the rate of maturation explain some of the increase in the magnitude of the brain volume and intelligence correlation from childhood to adulthood. For example, there are individual differences in the neurophysiological maturation of the frontal lobes across childhood and adolescents (Nagy, Westerberg, & Klingberg, 2004; Segalowitz & Davies, 2004). Furthermore, several of the neurophysiological characteristics of maturation may be substantially independent of brain volume (e.g., pruning, intra-cortical myelination; Paus, 2005). Thus, until such neurophysiological characteristics are largely stabilised once maturation is complete (i.e., adulthood), the correlation between brain volume and intelligence may be expected to be attenuated. Stated alternatively, the correlation between brain volume and intelligence in children may not be a fully accurate reflection of the effect.

McDaniel (2005) noted the difficulties associated with conducting a comprehensive meta-analysis, as many empirical investigations did not include standard deviation or internal consistency reliability estimates associated with the test scores. In fact, McDaniel (2005) was required to use standard deviation artefact distribution imputation for 21 of the sample correlations, as only 16 of the 37 brain volume and intelligence studies reported the standard deviation associated with intelligence test scores. Thus, the key brain volume and intelligence correlation ($r = .33$) reported by McDaniel (2005) rests upon the assumption that the imputation method worked in a valid manner.

More recently, Pietschnig et al. (2015) conducted a meta-analysis on the brain volume and intelligence empirical literature. In contrast to Gignac et al. (2003) and McDaniel (2005), Pietschnig et al. (2015) obtained a substantial number of personal communications relevant to the association between brain volume and intelligence across a variety of studies and samples. Based on 120 sample correlations derived from a mix of healthy and clinical samples ($N = 6778$), Pietschnig et al. (2015) reported a meta-analytic correlation of $r = .24$ between brain volume and global measures of intelligence (e.g., FSIQ). Thus, Pietschnig et al. (2015) reported an effect notably smaller than the meta-analytic estimates reported by McDaniel (2005; $r = .33$) and Gignac et al. (2003; $r = .43$). Pietschnig et al. (2015) suggested that the correlations reported in previous meta-analyses were likely over-estimates, as the published literature was likely affected by selective reporting (i.e., statistically non-significant effects were not reported). In support of such an argument, the meta-analytic correlation between brain volume and general intelligence based on published results was reported by Pietschnig et al. (2015) at $r = .30$ ($k = 53$; $N = 3956$). By contrast, the corresponding meta-analytic correlation in non-published work was estimated at just $r = .17$ ($k = 67$; $N = 2822$).

It should be noted, however, that both Gignac et al. (2003) and McDaniel (2005) restricted their meta-analyses to healthy samples, whereas Pietschnig et al.'s headline correlation of .24 included both healthy and clinically mixed samples. Arguably, intelligence test scores obtained from individuals suffering from various clinical conditions should not be considered optimally valid indicators of intellectual functioning. For this reason, it is commonly recommended that individuals "…should not be assessed [for intelligence] unless they appear suitably healthy and well rested." (Reschly, Myers, & Hartel, 2002, p. 101). From a statistical perspective, a correlation between intelligence and a criterion would be expected to be suppressed in clinical samples, because it is unreasonable to assume that all of the examinees suffer from the exact same condition to the same degree. Such individual differences in the clinical condition would be expected to affect the rank ordering in measurement of intelligence, in comparison to "true" intelligence, which is a threat to validity, in this context.

If we wish to estimate the population correlation accurately, sample ascertainment is critical. Whereas a sample restricted to healthy adult individuals will, allowing for sampling error, approximate the true population estimate, mixtures of samples, with non-random inclusion criteria, are likely to show considerable bias. This is true not only in extreme cases (imagine a sample of people "administered" the Raven after the consumption of 10 standard drinks of alcohol) but is likely to hold in general.

Consider, for example, the report of a relatively low correlation of $r = .07$ between brain volume and intelligence, based on a sample of 41 neurological patients (Yeo, Turkheimer, Raz, & Bigler, 1987). Nineteen of these patients presented with headache complaints, while 7 presented with symptoms of problems in concentration and memory. The validity of the brain volume and IQ correlation is not established in such a combination of groups. That is, arguably, patients suffering from concentration and memory problems will produce IQ scores which are lower than their natural maximal capacity. By contrast, migraine patients completing the IQ testing may be expected to show substantial variability, depending on, for instance, the varying level of migraine experienced during their testing, from none at all to severe. However, in both groups of cases, brain volumes likely remained stable. Consequently, rank ordering of the IQ scores in this mixed clinical sample was likely affected adversely by the heterogeneity of the clinical conditions between the patients. Such an adverse impact on rank ordering of IQ scores would also affect adversely the estimated correlation between brain volume and intelligence. In light of the above, it is our view that the best sample estimate of the true association between brain volume and intelligence, as well as tests of hypothesized moderator effects, is obtained by aggregating studies of generally healthy adult samples.

Additionally, it is important to note that Pietschnig et al. (2015) did not correct any of the correlations (published or non-published) for range restriction. By contrast, both Gignac et al. (2003) and McDaniel (2005) did take range restriction into consideration. Pietschnig et al. acknowledged the issue of range restriction in their meta-analysis, however, they did not apply a correction to their analysis, because "…a majority of the included samples' standard deviations for test performance were not reported" (p. 426–427). However, based on our review, nearly all of the studies associated with the healthy adult samples ($k = 32$) did report standard deviations for the intelligence test scores. The importance of correcting observed correlations for range restriction to obtain a more accurate estimate of the effect in the population has been well established (Le & Schmidt, 2006). For example, based on the results of a simulation investigation, Duan and Dunlap (1997) found that when the population correlation was .30 and the selection ratio was .90 (i.e., the sample standard deviation was 10% smaller than the

---

[1] For an introduction to the problem of range restriction and the estimation of correlations in the population, consult Wiberg and Sundström (2009). More advanced treatments can be found in Sackett and Yang (2000) and Hunter, Schmidt, and Le (2006).