COGNITIVE NEUROSCIENCE, 2012, i*First*, 1–9

Psychology Press
Taylor & Francis Group

Intelligence and the brain: A model-based approach

Rogier A. Kievit[1], Hilko van Rooijen[1], Jelte M. Wicherts[1], Lourens J. Waldorp[1], Kees-Jan Kan[2], H. Steven Scholte[1], and Denny Borsboom[1]

[1]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[2]Department of Biological Psychology, Free University, Amsterdam, The Netherlands

Various biological correlates of general intelligence (*g*) have been reported. Despite this, however, the relationship between neurological measurements and *g* is not fully clear. We use structural equation modeling to model the relationship between behavioral Wechsler Adult Intelligence Scale (WAIS) estimates of *g* and neurological measurements (voxel-based morphometry and diffusion tensor imaging of eight regions of interest). We discuss psychometric models that explicate the relationship between *g* and the brain in a manner in line with the scientific study of *g*. Fitting the proposed models to the data, we find that a MIMIC model (for multiple indicators, multiple causes), where the contributions of different brain regions to a unidimensional *g* are estimated separately, provides the best fit against the data.

Keywords: Intelligence; Cognitive neuroscience; *g*; Structural equation modeling; Neuro *g*; Psychometrics.

Although technological advances have expanded the possibilities for empirical research, the nature of the relationship between intelligence and the brain remains a contentious topic. This is partly an embarrassment of riches, as the list of properties that correlate with general intelligence is daunting. Research has shown that gray matter density, white matter integrity, skull volume, cortical thickness, uric acid levels, height, lower amplitude in the averaged evoked potential (a measure derived from the EEG signal), increased and decreased neural activity, functional efficiency, and nerve conduction velocity are associated with *g* (Deary, Penke, & Johnson, 2010; Jensen, 1998; Jung & Haier, 2007; and references therein). In this paper, we aim to structure these findings by proposing a new modeling framework for studying the relationship between behavioral measurements of general intelligence and neurological measurements of the brain.

The interpretation of the term "intelligence" is itself the topic of several long-standing debates (Mackintosh, 1998). For purposes of clarity, we focus on the interpretation of intelligence that is statistically most clearly defined, namely as the general factor of intelligence (also known as "*g*"). We consider *g* to represent a common source of variance in general cognitive ability. Studying the brain in relation to *g*, some researchers have coined the phrase "neuro *g*" (Haier et al., 2009). This suggests a strong thesis, namely that there is some fundamental biological substrate that acts as a common cause of individual variability in performance on a wide variety of cognitive tests. From this point of view, we could consider the search for neuro *g* a quest for "[t]he substrate of human intelligence" (Luders, Narr, Thompson, & Toga, 2009, p. 156), undertaken "[t]o capture the essence of a neural basis of intelligence" (Jung & Haier, 2007, p. 178). The above suggests that the aim of neuroscientific research on intelligence is not just to find some (neurological) property that *correlates* with a common factor extracted from a battery of IQ-test scores, but rather showing *g* to be a physical property of the brain.

Recent studies have extended the focus from single measures to integrating multiple lines of evidence. These approaches have combined structural and

Correspondence should be addressed to: Rogier A. Kievit, Department of Psychological Methods, University of Amsterdam, Roetersstraat 15, 1018WB Amsterdam, The Netherlands. E-mail: r.a.kievit@uva.nl

www.psypress.com/cognitiveneuroscience http://dx.doi.org/10.1080/17588928.2011.628383

functional data (Choi et al., 2008), structural and functional networks (Li et al., 2009; van den Heuvel, Stam, Kahn, & Hulshoff Pol, 2009), and genetic modeling (Chiang et al., 2009). These approaches show that the synthesis of different techniques may be necessary to get a handle on the nature of *g* and the brain. Despite many advances, however, the identification of a neural property that could be identified with *g* has not been forthcoming. This has led some scholars to question whether any simple identification should be expected. For instance, Bartholomew (2004) suggests that *g* is not itself a physical characteristic of the brain, but rather a function of a set of distinct brain characteristics.

The distinction is comparable to the distinction between the length (height) of a person and his or her physical fitness. If *g* is similar to "length" (albeit more difficult to measure), then what we have to discover is how to best measure *g* in the brain, with the least amount of measurement error. However, if *g* is more similar to something akin to "physical fitness," then searching for a single physical property is ill-advised: Although physical fitness depends on a range of physical properties (e.g., lung capacity, metabolic rates, energy efficiency), the construct of "physical fitness" is best seen as a composite function of these characteristics, and cannot be identified with any single characteristic. If the latter perspective is more accurate, then neuroscientific research on *g* should not be interpreted as an attempt to find out *what g is* (i.e., which neural property uniquely defines *g*), but as a search for a possibly large and heterogeneous set of properties that together *determine g*.

Clearly, the relationship between *g* and the brain can be studied from a wide range of assumptions and hypotheses. However, these assumptions are not always made explicit. In this paper, we show how different assumptions and interpretations can be translated into testable psychometric models. Structural equation models (SEM) have been used profitably in analyzing functional neural systems (e.g., McIntosh & Gonzalez-Lima, 1994). Excellent introductions in the application and interpretation of SEM include Bollen (1989) and Kline (2005). We use SEM in a new way by modeling various competing hypotheses for the relationship between observed and latent variables of *g* and the brain. We will examine the relationship between a set of representative regions of interest (ROIs) to study which models, with which implications, best represent the empirical data. We first discuss a number of candidate models for representing the relationship between neurological and psychological measurements.

METHODS

Models

We use SEM to fit the following candidate models representing different assumptions and interpretations against the data.

Neuro g 1: Same concept, new measures

The first psychometric interpretation of neuro *g* considers neurological measurements to depend on *the same property* as has traditionally been measured with psychometric instruments (i.e., *g,* as measured through IQ tests). The measurement model corresponding to this interpretation is shown in Figure 1.

From this perspective, one can consider neurological measurements (by means of brain scans) to be a new, potentially more precise way of measuring the same property (in this case, *g*). This model entails several statistical characteristics and conceptual implications explained in more detail in Kievit et al. (2011). This is the simplest model considered here, and is most in line with terminology stressing "the neural substrate of *g*"or the "biological essence of *g*."

Neuro g 2 and Neuro g 3: Different properties, different measures

Another conceptualization represents neuro *g* as a latent variable distinct from psychometric *g*. It can be seen as the biological "cousin" of psychometric *g*, and, as such, is a *different* property.

From this perspective, represented in Figure 2, neuro *g* is a latent variable that represents a

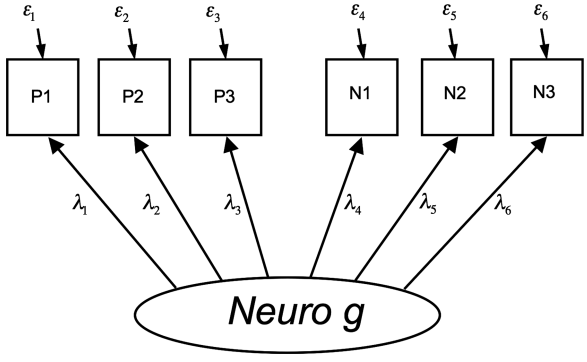

Figure 1. Neuro *g* 1. Reflective, unidimensional representation of neuro *g*. Neuro *g* is the same property as *g*, estimated by both psychological (P) and neurological (N) indicators, with factor loadings (lambda) and residual terms (epsilon) estimated empirically.

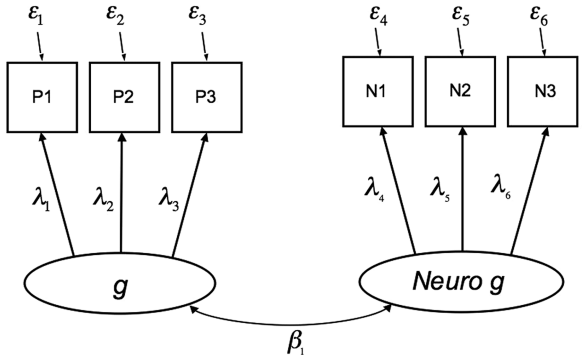

Figure 2. Neuro *g* 2. Neuro *g* as a separate latent variable, estimated by a unidimensional constellation of neural indicators. The correlation with psychometric *g* can be assessed empirically. If this correlation is 1, this model is equivalent to model 1.

unidimensional neurological factor, measured by a set of neurological variables. This neuro *g* can be seen as a "property of the brain" that is *relevant* with respect to *g*, but not *identical* to it.

Hence, this model offers a different interpretation of the term "neuro *g*": namely, as a property of the brain that can be estimated by neurological measurements, and that correlates with *g* to an extent assessable by empirical study. The purpose of empirical research then is to discover *which* neurological properties (for instance, "gray matter density in region X," or "white matter integrity in area Y") covary together in a population such that this "neuro *g*" correlates most highly with *g*. As such, it could be interpreted as a latent factor that might be called "brain fitness." This perspective can take two psychometrically similar but conceptually distinct routes. In the first route, a single latent variable is conceived of as the aggregate of "brain fitness" variables. Versions of this model were implemented by MacLullich et al. (2002) and Penke et al. (2010). In the second route, one utilizes multiple "neuro *g*" factors, attempting to caputre the dimensions along which people can vary neurologically. This conceptualization is represented in Figure 3 (neuro *g* 3) below, in this case with two latent neuro *g*'s, although it could be extended to include more factors if necessary.

Allowing for multiple latent variables has two additional benefits: Firstly, this may be more neurologically plausible, as it does not require all neurological measurements to be monotonically related, as do models 1 and 2.

Secondly, such separate latent variables would allow for more substantive interpretation. It may be, for instance, that one neurological latent variable represents an estimate of "perceptual organization," another

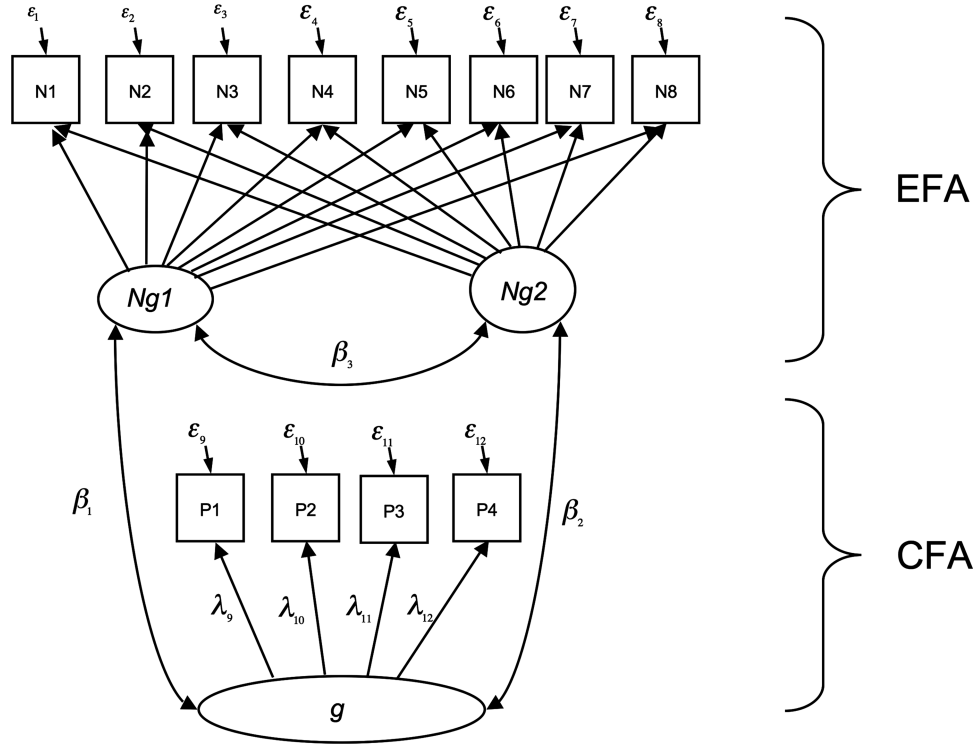

Figure 3. Neuro *g* 3. A combined EFA/CFA model. The dimensionality of the neurological indicators, in this case, eight measures, is estimated in an EFA. The resulting two factors are correlated with psychometric *g*. This model can be extended to include a broader sample of neurological indicators. For visual clarity, the lambdas for the EFA (top part) are not drawn. EFA, Exploratory factor analysis; CFA, Confirmatory factor analysis.

represents "processing speed," and yet another represents "verbal ability." Attempting to model such latent variables may allow researchers to get a grip on the structure of different neurological dimensions along which people differ, and this would in turn facilitate interpretation and comparability across studies. This method may also be particularly appropriate for studying lower order factors such as spatial or verbal ability. Previous research suggests that although cognitive abilities are positively correlated, they may rely on different neural subsystems. If this is the case, it is worth examining the psychometric structure of neurological variables at the level of these lower-order factors. Those neural indicators that covary positively with a lower order factor may be especially insightful. For instance, Henson (2005) discusses how neurological and behavioral measurements may *together* be able to distinguish between theories that propose either a unitary short-term memory factor or more than one lower-order memory factor, based on convergence of neuroimaging and behavioral data.

Of course, the structure of covariation between different neurological measurements and their relation to higher- or lower-order ability factors will depend on the nature of those indicators. For instance, that the development of white matter in different regions of the brain, at least in early life, is based on similar genetic mechanisms may suggest that there will be a higher degree of similarity between such measurements within an individual than between, say, white and gray matter values in the same brain region. Such considerations need to be taken into account when modeling. However, the structure of covariation of neurological measurements is ultimately an empirical question, and it is exactly the study of such covariance that may provide valuable insights regarding the neurophysiology of the brain and *g*.

Neuro g 4: The brain determines differences in g

A final psychometric possibility we consider here is offered by a so-called MIMIC model (for multiple indicators, multiple causes) (Jöreskog & Goldberger, 1975). This model represents an asymmetrical relationship between cause (formative) and effect (reflective) indicators (cf. Edwards & Bagozzi, 2000). From this perspective, represented in Figure 4, psychometric *g* is *determined* by a constellation of neurological properties, but *measured* by psychological variables (e.g. an IQ test).

Here, the neurological properties do not *measure* but together form or determine an individual's score on *g,* akin to the "physical fitness" example discussed previously. This is in line with conceptualizations

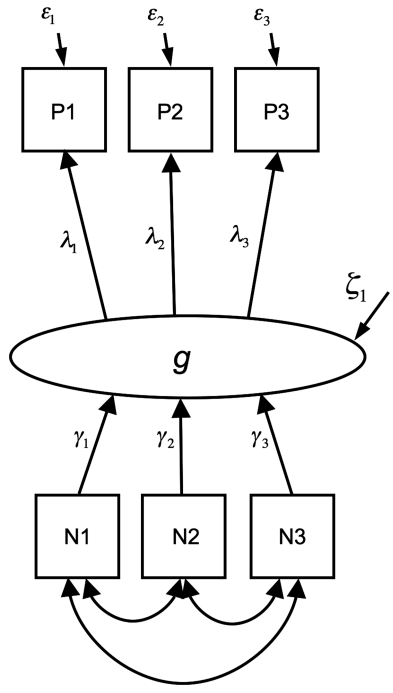

Figure 4. Neuro *g* 4. A MIMIC model representation of the relationship between *g* and the brain. The neurological indicators together determine psychometric *g*. The relative weights of the gammas are estimated based on the *g* indicators—in this case, the WAIS sum scores. MIMIC, Multiple Indicators, Multiple Causes; WAIS, Wechsler Adult Intelligence Scale.

where *g* is seen as something that is determined by a constellation of brain properties (Bartholomew, 2004). The model is also biologically less restrictive, as it allows for covariance between neurological measurements beyond those explaining variance in *g*; the neurological part of the model does not have to be unidimensional. The MIMIC model therefore may more naturally accommodate "brute facts" about the physiology of humans with respect to *g*. However, the model does assume that the neurological indicators have been measured without error, and this may be unrealistic.

Sample

The data consisted of a sample of 80 participants (29 males, 51 females) who completed the Dutch version of the WAIS-III—a fully validated, translated version of the original WAIS (Wechsler Adult Intelligence Scale) (cf. Wechsler, 2005). Participants, who ranged from 18 to 29 years old, with a mean age of 21.1 years ($SD = 2.55$), received either a financial reward or course credits for their participation. The participants

were tested in accordance with the ethical guidelines of the American Psychological Association, and the study was approved by the University of Amsterdam Ethics Committee. The behavioral measurements consisted of four domain indices (Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed) of the WAIS-III. In addition to the completion of the WAIS-III, all participants were scanned in a 3 Tesla Philips Intera MRI scanner to estimate white matter, gray matter density, and brain volume measurements in eight ROIs.

Image acquisition

Participants were scanned on a 3-T Philips Intera scanner, and all data were analyzed using FSL (Smith et al., 2004), MATLAB (Version 7.10.0, The Mathworks, Inc., Natick, MA, USA), and Mplus (Muthén & Muthén, 1998–2007). A structural MRI scan of each participant was acquired by using a T1-weighted 3D sequence (Turbo Field Echo, TE 4.6 ms, TR 9.6 ms, FA 8°, 182 sagittal slices of 1.2 mm, FOV 250^2 mm, reconstruction matrix 256^2).

Image analysis

The structural data were analyzed with voxel-based morphometry (VBM) carried out with FSL (Smith et al., 2004). First, structural images were brain-extracted (Smith, 2002). Next, tissue-type segmentation was carried out with FAST4 (Zhang, 2001). The so obtained gray-matter partial volumes were then aligned to MNI 152 standard space using the affine registration. The resulting images were averaged to create a study-specific template, to which the native gray-matter images were then nonlinearly reregistered with a method that uses a b-spline representation of the registration warp field (Andersson, Jenkinson, & Smith, 2007; Rueckert et al. 1999). The registered partial volume images were modulated (to correct for local expansion or contraction) by dividing by the Jacobian of the warp field. The modulated segmented images were smoothed with an isotropic Gaussian kernel with a sigma of 4 mm. In addition, a DWI (diffusion weighted imaging) scan was run (echo planar imaging, TE 94 ms, TR 7.720 ms, FA 90°, 40 slices, FOV 224^2, reconstruction matrix 128^2). The diffusion tensor imaging (DTI) data were also analyzed with FSL, now using the TBSS (Tract Based Spatial Statistics) package (Smith et al., 2006). The above procedure was applied to the first and second T1 and DWI scans separately, creating two data sets that are independent from the perspective of noise. The first data set was used to identify ROIs. Threshold-free cluster enhancement was applied to all data sets (Smith & Nichols, 2009). Data were further thresholded at a value of $p < .01$ (minimum cluster size 50 mm^3) for DWI data and $p < .01$ (cluster size 800 mm^3) for VBM data. The second data set was used to extract the actual values from the ROIs yielded by the first analysis. In this way, these ROIs were not artificially inflated in terms of statistical fitting. From both the VBM and DTI data, the two ROIs that were strongest positively correlated with the FSIQ score on the WAIS-III, and the two ROIs that were strongest negatively correlated with FSIQ were used for further analysis. This procedure resulted in eight ROIs (four VBM and four DTI measures), of which the MNI coordinates of the center of gravity are in Table 1. All these ROIs have previously been associated with individual differences in general cognitive abilities (cf. Deary et al., 2010; Jung & Haier, 2007, and references therein); Brodmann's area (BA) 9 bordering on 46, BA area 9 bordering on 48, BA area 20, and BA area 18. The DTI loci were also in accordance with previous findings, such as for the inferior fronto-occipital fasciculus and various sections of the corpus callosum.

RESULTS

We report results for four latent variable models, which were fitted using Mplus (Muthén & Muthén, 1998–2007) by maximum likelihood estimation. In a covariance structure analysis, the model-implied

TABLE 1
MNI coordinates, voxel count, and morphological description of four VBM and four DTI measures

Measure	*Region*	*X coordinate*	*Y coordinate*	*Z coordinate*	*N voxels*
VBM 1	*9 and 46*	29	41	32	985
VBM 2	*9 and 48*	–41	22	44	565
VBM 3	*20*	38	–12	–39	117
VBM 4	*18*	21	–91	22	184
DTI 1	*Callosal body (directly below the basal ganglia)*	–9	15	–11	86
DTI 2	*Inferior fronto-occipital fasciculus*	–40	–22	–7	50
DTI 3	*Corpus callosum*	–14	–4	33	114
DTI 4	*Corpus callosum*	–18	–25	36	106

covariances are compared to the observed covariances. Fit indices represent the deviation or misfit of the observed covariance structure. We evaluate model fit by means of the chi-square test of model fit, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Bayesian information criterion (BIC). For discussions on the relative merits of these fit indices, see Schermelleh-Engel, Moosbrugger, and Müller (2003).

For the SEM analyses, a total of 12 observed variables were modeled: four domain scores of the WAIS-III, and measurements of four VBM, and four DTI ROIs. The WAIS sum scores were as follows: Verbal Comprehension ($M = 117.16$, $SD = 9.78$), Perceptual Organization ($M = 112.10$, $SD = 11.31$), Working Memory ($M = 111.32$, $SD = 13.11$), and Processing Speed ($M = 116.38$, $SD = 14.80$). We implement a simplified, nonhierarchical *g* model here for purposes of simplicity and model fitting. This model may be extended to include latent first-order factors without affecting the core ideas of the models.

To ensure that the indicators of *g* variables are psychometrically adequate, we first fit a confirmatory model with the four WAIS sum scores and one latent *g* factor. The chi-square test of model did not reject the one-factor model of *g* χ^2 (2, $n = 80$) = 2.15, $p > .30$. This fit was corroborated by the other fit indices, RMSEA (0.031) and CFI (0.996). We then fitted the four models discussed previously. Table 2 shows the resulting fit statistics. The standardized parameter estimates for all four neuro *g* models are presented in Table 3.

First, consider the simple, reflective model, in which all indicators are considered to reflect *g*. This model does not fit: It is rejected by the chi-square test, and other indicators corroborate this poor fit. That is, for this data set, neurological measurements *cannot be considered measurements of g*. Naturally, this may be a contingent fact about our data set. However, we think that the multifaceted nature of neurological properties that covary with *g* make this model an unlikely candidate for other data sets as well. For model 1 then, the report seems dire: Although this representation is in line with the terminology of "neuro *g*," it is not plausible a priori, and fits poorly when tested in a representative data set. In summary, these data indicate that

TABLE 2
Model fits for the four fitted SEM models. Selection of fit indices are based on Schermelleh-Engel et al. (2003)

Model	*Description*	*N*	*df*	*chi-square*	*p-value*	*RMSEA*	*CFI*	*BIC*
1	Reflective	80	54	119.98	0	0.124	0.629	7151.139
2	1 neuro *g*, 1 *g*	80	53	119.684	0	0.125	0.625	7155.224
3	1 *g*, 2 neuro *g*	80	45	72.979	0.0052	0.088	0.843	2738.439
4	MIMIC	80	26	32.194	0.1868	0.055	0.936	–39.174

Note: Df, Degrees of freedom; RMSEA, Root Mean Square Error of Approximation; CFI, Confirmatory Fit Index; BIC, Bayesian Information Criterion.

TABLE 3
Standardized parameter estimates for four fitted models. The final model contains two neural factors, loadings of which are represented side by side

Indicator	*CFA g model*	*1 (neuro) g*	*1 g, 1 neuro g*	*MIMIC*	*1 g, 2 neuro g*	
Verbal Comprehension	0.591	0.689	0.685	0.675	0.705	
Perceptual Reasoning	0.779	0.659	0.646	0.639	0.684	
Working Memory	0.486	0.545	0.544	0.534	0.511	
Processing Speed	0.476	0.442	0.442	0.479	0.387	
					Factor 1	*Factor 2*
VBM 1		0.3	0.315	0.295	–0.1	0.183
VBM 2		0.284	0.284	0.205	–0.336	–0.015
VBM 3		–0.471	–0.441	–0.051	1.07	0.002
VBM 4		–0.372	–0.35	–0.238	0.593	–0.036
DTI 1		0.514	0.483	0.161	–0.171	0.628
DTI 2		0.554	0.539	0.419	0.008	0.828
DTI 3		–0.218	–0.232	–0.248	0.065	–0.095
DTI 4		–0.224	–0.235	–0.224	0.013	–0.089

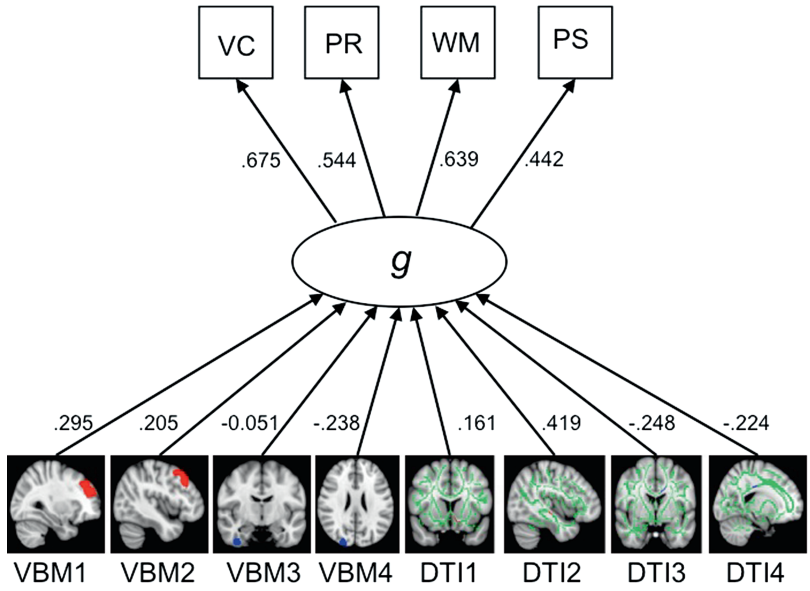

Figure 5. The best-fitting model: a MIMIC model. Eight neurological measurements, described in more detail in Table 1, jointly determine *g*. The model shows standardized parameter estimates for the eight formative indicators and four reflective indicators. Model fit is shown in Table 2.

neuro *g* should not be taken to refer to a unidimensional constellation of neural properties identical to *g*. Whether or not other constellations of neural indicators not explored in this study *will* fit such a unidimensional factor coincident with *g* is an empirical question that is still open, but, for the reasons we give above, we do not consider it very likely.

Model 2, shown in Figure 2, represents neuro *g* as a separate latent variable that correlates with *g*. As can be seen in Table 2, this model was also rejected for our data set, reflecting the lack of unidimensionality in the neurological measurements. That is, at least in this data set, a constellation of neurological measurements cannot be considered a unidimensional property of individuals, akin to a kind of "brain fitness factor." Despite the fact that these measures correlate independently with *g*, they do not intercorrelate positively.

Model 3 (shown in Figure 3) estimates multiple (in this case, two) latent neural factors that correlate with *g*. To fit this model, we implemented a new method available in Mplus, namely a combined exploratory and confirmatory factor analysis. As Table 2 shows (model 3), this model is also rejected, although the fit is considerably less poor than the fit of models 1 and 2. This model, although rejected for this data set, seems more intuitively plausible and offers more in the way of interpretation, and in our view may well be an option worth considering for further research.

Finally, we consider the formative, or MIMIC model, as represented in Figure 4. This model is not rejected by the chi-square test of model fit, and other indicators also represent adequate to good fit (see Table 2, model 4). For this data set, we can consider neurological measurements to jointly predict a unidimensional *g,* although they do not themselves form a unidimensional scale. This situation is consistent with the idea that the neurological properties together determine individual differences in *g*. Figure 5 shows this best-fitting model, including standardized parameter estimates of each of the neural indicators.

Based on this data set, the simplest candidates for neuro *g* are rejected. The more complex models, which either attempt to capture several latent dimensions of neurological covariation (model 3) or model an asymmetric relationship between *g* and the brain (model 4), perform better.

GENERAL DISCUSSION

Neuro *g* is unlikely to be a physical characteristic of the human brain as measured by our neurological indicators, and it is also unlikely to be a unidimensional physical variable correlated with psychometric *g*. More complex models are needed to explain the joint covariance structure of neurological and behavioral measures. These models feature either an asymmetric relationship between *g* and the brain, or capture several dimensions of covariation. Thus, it seems at least unlikely that there is a simple "neuro *g*". Of course, these conclusions are partly contingent on the data we acquired here: It may be that other data sets yield

different conclusions regarding the models. Moreover, there is always the possibility that unobserved third factors are the source of covariation between neurological indicators and intelligence (e.g., white matter, cardiovascular factors, and cognitive ability; Marks et al., 2007), although this problem is not unique for SEM. More generally, although SEM has many statistical benefits over traditional techniques, there are issues that require careful attention, including model equivalence (Lee & Hershberger, 1990), model selection (Myung & Pitt, 1997), and judging model fit (Schermelleh-Engel et al., 2003). In addition, SEM generally requires larger sample sizes than other approaches, although required samples are not prohibitive (e.g., Marsh & Hau, 1999). Nonetheless, SEM offers various essential benefits over more traditional methods. Most importantly, they allow for the flexible comparison of various models in such a way that they can be compared across studies. In this paper, we were able to compare various competing hypotheses, reject certain alternatives, and tentatively conclude that the MIMIC perspective currently offers the best explanation of the data analyzed here. Moreover, SEM is a flexible tool: It could be extended by focusing on different neural indicators (that may show higher factor loadings or fit to stricter models), or by examining hierarchical models of *g* (e.g., Carroll, 1993).

This finding is in line with recent work in other fields, such as that of emotion research. Lindquist, Wager, Kober, Bliss-Moreau, & Barrett (in press) show that, for the relation between emotions and the brain, a model similar to the MIMIC model is better supported empirically than the essentialist view (where the activity of one specialized subsystem is considered the core feature of a particular emotion, as in model 1). In addition, the statistical and conceptual properties of the MIMIC model are most compatible with contemporary perspectives on the genetic influences on *g* (cf. Penke, Denissen, & Miller, 2007). In terms of the MIMIC model, genetic effects may therefore feature as predictors alongside the neurological variables.

Our approach emphasizes the importance of conceptual and statistical clarity for neuroscientific research in intelligence. Neuroscientific findings should, whenever possible, go beyond simple measures of association. Psychometric modeling techniques, as discussed in this paper, allow us to see beyond simple correlations. This enables the investigation of conceptual hypotheses on the relation between intelligence and the brain that were hitherto the province of mere speculation. By explicitly representing, modeling, and testing competing hypotheses, we may be able to finally get a grip on this complex problem. We can only hope that, as was the case in the history of general intelligence and the development of factor analytic methods, competing methods and models will lead the way to new conceptual and empirical developments. We think that this will prove to be the most interesting, insightful, and productive road to a better understanding of the neurological basis of intelligence.

Original manuscript received 22 June 2011
Revised manuscript accepted 27 September 2011
First published online 18 January 2012

REFERENCES

Andersson, J. L. R., Jenkinson, M., & Smith, S. M. (2007). Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2. Retrieved from: www.fmrib.ox.ac.uk/analysis/techrep/tr07ja2/tr07ja2.pdf (accessed March 12, 2009).

Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge, UK: Cambridge University Press.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York, NY: Cambridge University Press.

Chiang, M. C., Barysheva, M., Shattuck, D. W., Lee, A. D., Madsen, S. K., Avedissian, C., ... Thompson, P. M. (2009). Genetics of brain fiber architecture and intellectual performance. *Journal of Neuroscience*, *29*, 2212–2224.

Choi, Y. Y., Shamosh, N. A., Cho, S. H., DeYoung, C. G., Lee, M. J., Lee, J. M., ... Lee, K. H. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *Journal of Neuroscience*, *28*, 10323–10329.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*, 201–211.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155–174.

Haier, R. J., Colom, R., Schroeder, D. H., Condon, C. A., Tang, C., Eaves, E., et al. (2009). Gray matter and intelligence factors: Is there a neuro-*g*? *Intelligence*, *37*, 136–144.

Henson, R. (2005). What can functional neuroimaging tell the experimental psychologist? *Quarterly Journal of Experimental Psychology*, *58*, 193–233.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger/Greenwood.

Jöreskog, K., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*, 631–639.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*, 135–187.

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Scholte, H. S., Wicherts, J. M., & Borsboom, D. (2011). Mind the gap: A

psychometric approach to the reduction problem. *Psychological Inquiry*, *22*, 67–87.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research*, *25*, 313–334.

Li, Y., Liu, Y., Li, J., Qin, W., Li, K., Yu, C., et al. (2009). Brain anatomical network and intelligence. *PLoS Computational Biology*, *5*, e1000395.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (in press). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*.

Luders, E., Narr, K. L., Thompson, P. M., & Toga, A. W. (2009). Neuroanatomical correlates of intelligence. *Intelligence*, *37*, 156–163.

Mackintosh, N. J. (1998). *IQ and human intelligence*. New York, NY: Oxford University Press.

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.

Marks, B. L., Madden, D. J., Bucur, B., Provenzale, J. M., White, L. E., Cabeza, R.,, et al. (2007). Role of aerobic fitness and aging on cerebral white matter integrity. *Annals of the New York Academy of Sciences*, *1097*, 171–174.

Marsh, H. W., & Hau, K.-T. (1999). Confirmatory factor analysis: Strategies for small sample sizes. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 251–306). Thousand Oaks, CA: Sage.

McIntosh, A. R., & Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, *2*, 2–22.

Muthén, L. K., Muthén, B. O. (1998–2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin and Review*, *4*, 79–95.

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, *21*, 549–587.

Penke, L., Muñoz Maniega, S., Murray, C., Gow, A. J., Valdés Hernández, M. C., Clayden, J. D., … Deary, I. J. (2010). A general factor of brain white matter integrity predicts information processing speed in healthy older people. *Journal of Neuroscience*, *30*, 7569–7574.

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., & Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, *18*, 712–721.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of fit measures. *Methods of Psychological Research Online*, *8*, 23–74.

Smith, S. E., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*, 83–98.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., & Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*, 1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, 208–219.

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, *29*, 7619–7624.

Wechsler, D. (2005). *WAIS-III Nederlandstalige bewerking*. Amsterdam, The Netherlands: Harcourt Assessment.

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*, 45–57.