Behavioral/Cognitive

Multivariate Lesion-Behavior Mapping of General Cognitive Ability and Its Psychometric Constituents

Mark Bowren Jr,[1] Ralph Adolphs,[2] Joel Bruss,[3] Kenneth Manzel,[3] Maurizio Corbetta,[4,5] Daniel Tranel,[1,3] and Aaron D. Boes[6]

[1]Department of Psychological and Brain Sciences, University of Iowa, Iowa City, Iowa 52242, [2]Division of Humanities and Social Sciences, California Technical Institute, Pasadena, California 91125, [3]Department of Neurology, Carver College of Medicine, Iowa City, Iowa 52242, [4]Department of Neuroscience, Venetian Institute of Molecular Medicine and Padova Neuroscience Center, University of Padua, Padova PD 32122, Italy, [5]Departments of Neurology, Radiology, and Neuroscience, Washington University School of Medicine, St. Louis, Missouri 63110, and [6]Departments of Neurology Psychiatry, and Pediatrics, Carver College of Medicine, Iowa City, Iowa 52242

General cognitive ability, or general intelligence (g), is central to cognitive science, yet the processes that constitute it remain unknown, in good part because most prior work has relied on correlational methods. Large-scale behavioral and neuroanatomical data from neurologic patients with focal brain lesions can be leveraged to advance our understanding of the key mechanisms of g, as this approach allows inference on the independence of cognitive processes along with elucidation of their respective neuroanatomical substrates. We analyzed behavioral and neuroanatomical data from 402 humans (212 males; 190 females) with chronic, focal brain lesions. Structural equation models (SEMs) demonstrated a psychometric isomorphism between g and working memory in our sample (which we refer to as g/Gwm), but not between g and other cognitive abilities. Multivariate lesion-behavior mapping analyses indicated that g and working memory localize most critically to a site of converging white matter tracts deep to the left temporo-parietal junction. Tractography analyses demonstrated that the regions in the lesion-behavior map of g/Gwm were primarily associated with the arcuate fasciculus. The anatomic findings were validated in an independent cohort of acute stroke patients ($n = 101$) using model-based predictions of cognitive deficits generated from the Iowa cohort lesion-behavior maps. The neuroanatomical localization of g/Gwm provided the strongest prediction of observed g in the new cohort ($r = 0.42$, $p < 0.001$), supporting the anatomic specificity of our findings. These results provide converging behavioral and anatomic evidence that working memory is a key mechanism contributing to domain-general cognition.

Key words: brain networks; general cognitive ability; general intelligence; lesion method; psychometrics; working memory

Significance Statement

General cognitive ability (g) is thought to play an important role in individual differences in adaptive behavior, yet its core processes remain unknown, in large part because of difficulties in making causal inferences from correlated data. Using data from patients with focal brain damage, we demonstrate that there is a strong psychometric correspondence between g and working memory – the ability to maintain and control mental information, and that the critical neuroanatomical substrates of g and working memory include the arcuate fasciculus. This work provides converging behavioral and neuroanatomical evidence that working memory is a key mechanism contributing to domain-general cognition.

Received June 4, 2020; revised Sep. 15, 2020; accepted Oct. 1, 2020.

Author contributions: M.B., R.A., M.C., D.T., and A.D.B. designed research; M.B., J.B., and K.M. performed research; M.B. and J.B. analyzed data; M.B. wrote the paper.

This study was supported by the National Institute of General Medical Sciences Grant T32GM108540, the National Institutes of Mental Health Grants 1 P50 MH094258 and 1 R21 MH120441-01, the Kiwanis Foundation, FC-Neuro University of Padua, National Institute of Neurological Disease and Stroke Grants 1 R01 NS114405-01 and NS095741, Flagship of the European Research Area Joint Transnational Call, Neuro-DiP: Progetto Dipartimenti di Eccellenza Italian Ministry of Research (MIUR), and CARIPARO Foundation Padova. This work was conducted on an MRI instrument funded by National Institutes of Health Grant 1S10OD025025-01.

The authors declare no competing financial interests.

Correspondence should be addressed to Mark Bowren Jr at mark-bowren@uiowa.edu or Daniel Tranel at daniel-tranel@uiowa.edu.

https://doi.org/10.1523/JNEUROSCI.1415-20.2020

Introduction

Individual differences in cognitive abilities tend to be positively correlated: for example, people with strong attentional and perceptual skills also tend to have strong memory and reasoning skills, on average (compare with Fig. 1*A*). Spearman was the first to conceptualize this "positive manifold" of performance correlations as reflective of a domain-general cognitive ability, or general intelligence (g; Spearman, 1904). Over the last century, g has become central to cognitive science (Wasserman, 2018). However, despite progress in our understanding of the neural correlates

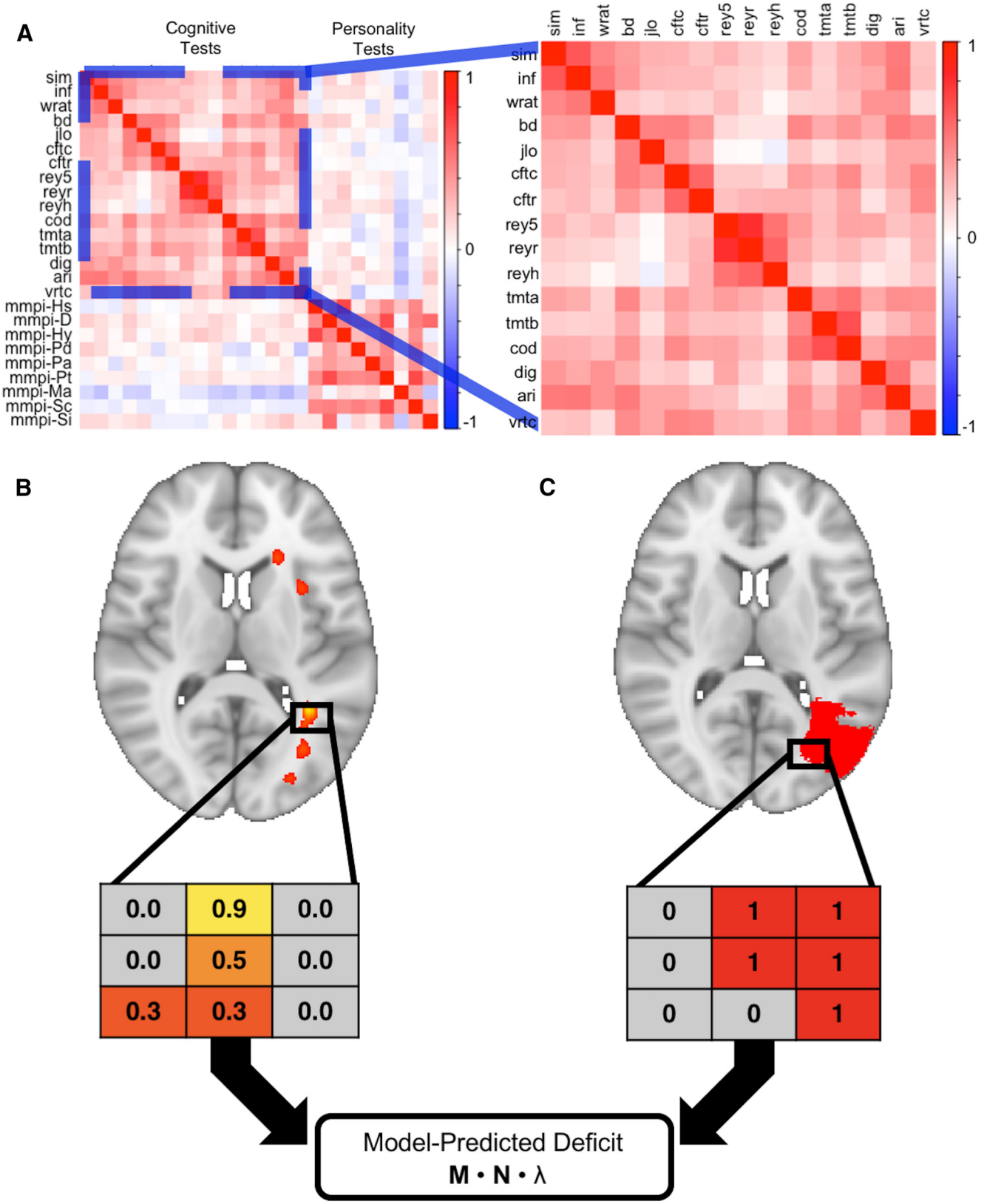

Figure 1. Schematic of overall approach. This figure summarizes how we used the Iowa cohort data to generate lesion-behavior maps (***A***, ***B***), which were then used to assign model-predicted factor scores to participants in the WU cohort based on lesion locations (***C***). Diverse cognitive abilities are positively correlated in this sample (expanded section), the basis for Spearman's g (***A***). Performances on individual tests were used to build SEMs, which were used in turn to estimate each latent variable for the subjects in the Iowa cohort. These factor scores were used to generate a lesion-behavior map (using SCCAN) for each factor, which is represented as a matrix of voxels weighted by their association with the factor in question (***B***; greater weights denoted by hotter colors). The lesion masks from the WU cohort were represented as a matrix of binary values [***C***; red (lesioned) or gray (not lesioned)]. A vector of model-predicted deficit scores was generated for the WU cohort by multiplying a matrix containing the voxel values from the lesion masks of the WU cohort subjects (M) by the matrix of voxel weights from a lesion-behavior map (N), and then by an eigenvalue derived from the lesion-behavior map in question (λ).

of g (Deary et al., 2010; Haier, 2017), a central question remains: what are the core processes that *constitute* g?

Working memory emerged as a leading explanation for domain-general cognition when Kyllonen and Christal (1990) demonstrated its near-perfect correlation with an estimate of g, sparking widespread interest into the relationship between these variables. Whereas some subsequent studies would find correlations of comparable or greater magnitude (Süß et al., 2002; Mackintosh and Bennett, 2003; Colom, 2004; Chen et al., 2019), others would find more modest associations between the two constructs (Ackerman et al., 2002; Conway et al., 2002; Kane et al., 2004). It is difficult to draw conclusions about the relationship between g and working memory based on this literature because of significant heterogeneity in how working memory has been defined and measured (Kane et al., 2005; Cowan, 2017) and the frequent use of fluid reasoning as a proxy for g in lieu of a large, diverse battery of cognitive tests (Johnson et al., 2004, 2008). Thus, although there is a general consensus that working memory is strongly associated with g, there is less agreement about whether working memory is a core process of g (Conway et al., 2003; Kane et al., 2005; Oberauer et al., 2005).

A key barrier to progress in elucidating the critical processes of g has been the reliance, by and large, on correlational data collected from healthy individuals. This approach has inherent limitations in elucidating causal mechanistic relationships, making it difficult to draw a distinction between necessity and sufficiency, or between correlated and causal consequences. Studies using patients with focal brain damage have the potential to overcome some of these methodological limitations. The lesion method allows for inference on the independence of cognitive processes, as abilities that are strongly associated among healthy individuals can become dissociated following focal brain damage, such as expressive and receptive language. Thus, a cognitive ability that is necessary and sufficient for g cannot be psychometrically or neuroanatomically dissociable from g – even among patients with diverse focal brain lesions. Using this approach, one lesion study highlighted the importance of executive functioning for g (Barbey et al., 2012). However, it remains unclear which aspects of executive functioning are primarily responsible for this finding (Friedman et al., 2006). Moreover, previous lesion studies have not incorporated advanced analytic techniques, such as multivariate lesion-behavior mapping methods, which are better suited to localize functions that depend on distributed brain networks (Glascher et al., 2010; Woolgar et al., 2010; Barbey et al., 2012, 2013).

Here, we aim to shed new light on the processes necessary for g by applying powerful multivariate methods to neuropsychological and neuroimaging data from two large samples of individuals with focal brain damage (Table 1). For the primary analysis, we used structural equation modeling (SEM) to examine the relationships between g and each the following domain-specific cognitive abilities: crystalized intelligence (Gc), visuospatial ability (Gv), learning/memory (Gl), processing speed (Gs), and working memory (Gwm), including an evaluation of the variance in g explained by each domain. Next, we used multivariate lesion-behavior mapping to determine whether g depends on specific neuroanatomical structures, and whether this localization is shared with any specific cognitive ability. To address generalizability and neuroanatomical specificity, we used lesion-behavior maps from one cohort (University of Iowa) to predict g in a second cohort (Washington University) based on lesion location (Fig. 1).

Table 1. Demographic characteristics of the Iowa and WU cohorts

Cohort	Iowa ($n = 402$)	WU ($n = 101$)
Age (SD)	58.17 (13.81)	53.14 (11.23)
Gender	212 M/190F	57 M/44F
Education, years (SD)	13.46 (2.72)	13.32 (2.68)
Handedness	360R/15 M/26L	92R/9L
Lesion chronicity at scan, months (SD)	42.17 (57.58)	0.46 (0.16)
Lesion chronicity at neuropsychological testing, months (SD)	31.83 (0.25)	0.46 (0.17)
Lesion laterality	175L/141R/86B	46L/55R

Age and lesion chronicity were calculated at date of scan; age is in years; Iowa cohort age range: 20–88; WU cohort age range: 19–83; gender: M = male, F = female; handedness: R = right handed, L = left handed, M = mixed handedness; one participant was missing data on handedness in the Iowa cohort; units of lesion chronicity are the number of months between date of lesion onset and date of MRI scan; lesion chronicity at neuropsychological testing for the Iowa cohort was determined for the administration of the WAIS tests because of variability in the exact date of testing for the numerous tests; lesion laterality: R = right sided, L = left sided lesion, B = bilateral lesion.

Materials and Methods

Two cohorts of patients with focal brain lesions were included. The primary analyses were performed using data from participants in the Iowa Cognitive Neuroscience Patient Registry (Frank et al., 1997; Tranel, 2007). A "validation" cohort was derived from a previously published study of stroke patients ($n = 101$) recruited from Washington University in St. Louis (Corbetta et al., 2015). All participants have previously given written informed consent to participate in these studies, which were approved by the Institutional Review Boards of the participating institutions. To supplement our analyses, we also analyzed the correlation matrix of the primary subtests of the fourth edition of the Wechsler Adult Intelligence Scales (WAIS), which is based on data from 2200 healthy adults (Wechsler, 2008).

Iowa cohort

The Iowa Registry includes patients with focal brain lesions who completed a range of standardized neuropsychological tests and neuroimaging in the chronic epoch (more than three months) after lesion onset. Participants from the Registry were included in this study if they completed at least 75% of the selected tests (described in the following paragraph) and had a lesion mask registered to the MNI152 template. Exclusion criteria included: progressive neurologic disease, major psychiatric illness, developmental intellectual disability, developmental lesion onset (defined as age at lesion onset <18 years), and developmental-onset epilepsy. A total of 402 participants (212 males; 190 females) in the registry met these criteria, and lesion etiologies included: ischemic stroke ($n = 238$), hemorrhagic stroke ($n = 87$), benign tumor resection ($n = 64$), focal encephalitis ($n = 9$), focal contusions from head injury ($n = 3$), and penetrating head injury ($n = 1$). For patients with benign tumor resection, lesion location was determined using the boundaries of the tumor resection cavity from the postoperative imaging.

We focused on neuropsychological tests that have been most commonly administered to the participants in the Iowa Registry. We assigned each test to latent variables using the Cattell–Horn–Carroll (CHC) model of cognitive abilities (Schneider and McGrew, 2018), a widely used taxonomy of mental abilities that has been shown to robustly account for the covariance structure of the neuropsychological tests used in this cohort (Jewsbury et al., 2017). Of note, several of these tests are subtests of the WAIS, which is the most widely-used instrument for measuring intelligence and its facets, and which has an extensive psychometric literature supporting its reliability and validity (Wechsler, 2008). All subjects were administered the version of the WAIS that was the most up-to-date at the time of their assessment. We used administration A from the Benton Visual Retention Test, which requires maintenance and immediate reproduction of visuospatial information (working memory) without repeated exposures or delayed recall demands. When a test was administered to a patient more than once, we used the administration that was most contemporaneous with the date of the patient's MRI scan, acquired in the chronic epoch three months or more after lesion onset. Test scores were adjusted for age using

normative data from test manuals (Wilkinson and Robertson, 2006; Wechsler, 2008), meta-normative data (Mitrushina et al., 2005), or normative data from the Benton Neuropsychology Laboratory. All age-adjusted scores were then transformed into standard units (i.e., Z-scores) to place all scores on a similar metric. In those instances where test scores are omitted, these data were imputed using multiple imputation by chained equations as implemented in the MICE package (van Buuren and Groothuis-Oudshoorn, 2011) available in R (R Core Team, 2017). A total of 5% of the total dataset described above was imputed. We performed the analyses (described below) in both the full dataset (i.e., the dataset with imputed data) and the restricted dataset (i.e., data from only those patients with complete data; $n = 230$).

Lesion location

Lesion borders were defined using lesion masks generated using either the MAP-3 method, or a manual tracing on a native T1-weighted MRI scan. The MAP-3 method involves the manual tracing of lesion borders on a template brain using the lesion depicted in an MRI or CT scan as a guide, and has been previously described and validated (Damasio and Frank, 1992; Fiez et al., 2000). With improvements in MRI co-registration techniques, lesions added after 2006 were manually traced on the native T1-weighted scans with FSL (Smith et al., 2004) and then transformed into MNI space. Because lesions negatively affect the accuracy of the transformation to MNI space, transformations were performed using enantiomorphic normalization, which replaces the lesion volume with the voxel intensities from its non-damaged homolog to more closely align the transform with its template. Bilateral lesions were transformed by applying a cost function mask to the lesion volume (Brett et al., 2001), which reduces the influence of voxels within the lesion volume on the transformation process. Native MRIs of the brain and lesion mask were then co-registered to the MNI152 T1 1 mm atlas using linear and nonlinear registration available in ANTs (Avants et al., 2008). The anatomic accuracy of the lesion tracing was reviewed both in native and MNI space and edited as needed by a neurologist (A.D.B.) blinded to cognitive data.

Experimental design and statistical analysis

SEMs

SEMs were used to investigate the association between each specific cognitive ability and g using the Iowa data. We used a SEM framework over an exploratory approach where possible because of the greater hypothesis-testing flexibility afforded by the former, especially as it relates to bifactor modeling (Bollen and Noble, 2011; Muthukrishna and Henrich, 2019), and because previous work has evaluated how the Iowa neuropsychological tests fit into the architecture of the CHC model (Jewsbury et al., 2017). These analyses were performed using the lavaan library in R (Rosseel, 2012). All free parameters were estimated using maximum likelihood. Parameters were reported from the completely standardized solutions (i.e., all factor loadings were standardized between values of 0 and 1). Standard errors for all parameters were derived using the bootstrapping procedures (with 1000 draws) available in the SEM function in lavaan. For each model, local fit was inspected to identify parameters that were not statistically significantly different from zero. Overall model fit to the data were evaluated using the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). Acceptable model fit is indicated by CFI of at least 0.90, and RMSEA and SRMR ≤ 0.08 (Hu and Bentler, 1999). Scores for latent variables were generated using the lavPredict function in R, which estimates values of latent variables using a regression based on model parameters.

The following domain-specific cognitive abilities could be modeled from the observed data: crystallized intelligence (Gc), visuospatial ability (Gv), learning efficiency (Gl), processing speed (Gs), and working memory (Gwm). To account for method covariance, the unique variances of the Complex Figure Test Copy and Recall scores were allowed to covary in all models, as were the unique variances of parts A and B of the Trail Making Test, and the indices of the Rey Auditory Verbal Learning Test. A hierarchical model (Fig. 2*A*) was used to estimate g and to examine the variance in g that can be accounted for by each domain-specific ability. The square of the correlation between each ability and g served as an index of the variance in g explained by each specific cognitive ability; this is indicated by the factor loading of each ability on g (Kline, 2016). A confirmatory bifactor model (Fig. 2*B*) was used to evaluate the "distinctiveness" of each cognitive ability from g. In this model, all observed data were first set to load directly onto g, and then onto the

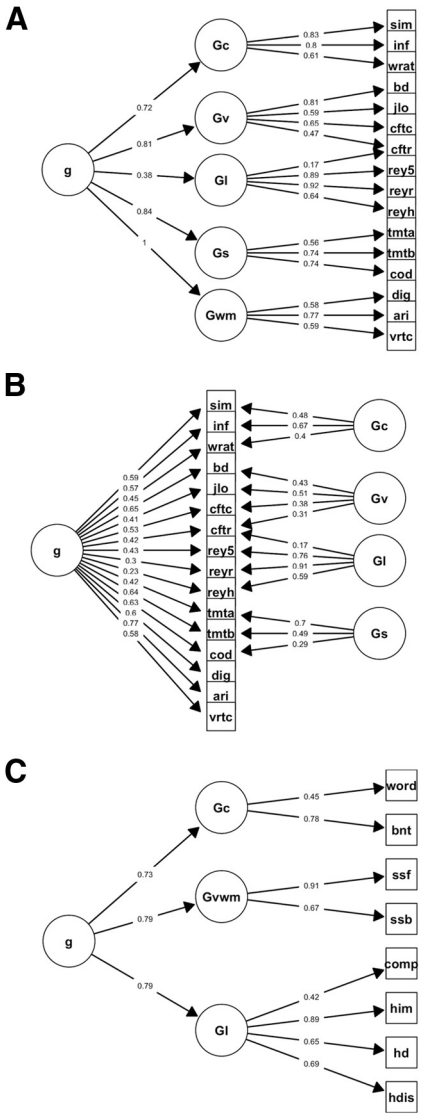

Figure 2. SEMs. A hierarchical model of cognitive abilities (***A***) was used to estimate g in the Iowa cohort and examine the correlation between each cognitive ability and g. Using a bifactor model of cognitive abilities (***B***), we investigated the extent to which crystalized intelligence (Gc), visuospatial ability (Gv), Gl, processing speed (Gs), and working memory (Gwm) were distinct from g. This was done by partitioning variance first to g, and then to the domain-specific factors. Correlated residuals were omitted to improve clarity. An exploratory hierarchical factor analysis was used to estimate g in the WU cohort (***C***; factor loadings were thresholded at 0.4 for clarity). Tests included: the similarities (sim), information (inf), block design (bd), digit-symbol coding (cod), digit span (dig), and arithmetic (ari) subtests from the Weschler Adult Intelligence Scales, the word reading test from the Wide Range Achievement Test (wrat), the Benton Judgment of Line Orientation Test (jlo), the copy (cftc) and recall (cftr) trials from the Rey-Osterrieth Complex Figure Test, trial 5 (rey5), delayed recall (reyr), and delayed recognition hits (reyh) from the Rey Auditory Verbal Learning Test, trials A (tmta) and B (tmtb) from the Trail Making Test, the number of correct items from the Benton Visual Retention Test (vrtc), the Word Comprehension test (word), the Complex Ideational Material test (comp), the Boston Naming Test (bnt) from the Boston Diagnostic Aphasia Examination, the forward (ssf) and backward (ssb) trials from the Spatial Span test of the Wechsler Memory Scales, and the immediate total recall (him), delayed recall (hd), and discrimination index (hdis) scores from the Hopkins Verbal Learning Test–Revised.

appropriate subfactor. As with all bifactor models, the latent variables were set to be orthogonal to one another under the assumption that the covariance among the factors is captured by the general factor. Using this model, ω hierarchical (ω_h) was calculated for each domain, which provided an index of the reliable variance associated with each domain-specific ability after removing the variance associated with g. Higher values indicate greater distinctiveness from g (i.e., greater reliable variance remaining after controlling for g).

Healthy cohort SEM

In addition to constructing models using the Iowa cohort with focal brain lesions, we also constructed analogous models using data from a large healthy population. This was done to evaluate whether conclusions from the models differed when derived from subjects with focal brain lesions versus healthy individuals. For this we used the WAIS-IV standardization sample correlation matrix, with a sample size of 2200 healthy adults described in detail elsewhere (Wechsler, 2008). These models were used to evaluate the correlations between specific domains (e.g., verbal comprehension) and g, and the distinctiveness of each domain from g, as performed with data from the subjects with focal brain lesions. Subfactors were labeled using the names of the associated WAIS-IV indices (e.g., the Verbal Comprehension Index).

Lesion-behavior mapping

Lesion-behavior mapping analyses were performed to identify neuroanatomical regions associated with impairment in domain-general and domain-specific cognitive abilities as estimated from the SEMs. Sparse canonical correlation analysis (SCCAN) was used for lesion-behavior mapping, as implemented in LESYMAP (Pustina et al., 2018), a package available in R (https://github.com/dorianps/LESYMAP). The SCCAN method uses an optimization procedure to derive voxel-weights that maximize the multivariate correlation between voxel values and behavioral scores. The predictive value and sparseness of the resulting maps were evaluated empirically using a fourfold, within-sample correlation between model-predicted and actual behavioral scores. LESYMAP deems a map "valid" if it is associated with a statistically significant predictive correlation. Thus, this approach tests the predictive value of the entire map at once and avoids the pitfalls associated with voxel-wise (i.e., mass univariate) methods, such as inflated rates of false-positive errors. This previously validated method of lesion-behavior mapping has been demonstrated to be more accurate than mass univariate methods and is better able to identify when multiple brain regions are associated with a behavioral variable (Pustina et al., 2018). Regions with minimal coverage (fewer than three lesions) were excluded, as performed previously (Hindman et al., 2018) to reduce processing time and the influence of outliers on the results while maintaining sufficient lesion coverage for the analyses. We also tested whether the spatial configuration of the voxel weights from the map of g was affected by changes in the sparseness parameter by manually manipulating this parameter to values above and below the final sparseness value as determined by LESYMAP.

A mass-univariate voxel-wise approach to lesion-behavior mapping was also employed to provide evidence for the robustness of our neuroanatomical results to different lesion-behavior mapping methodologies. This analysis was performed using the regresfast option in LESYMAP, which performs a voxel-wise regression of behavioral scores on voxel values. Multiple comparison correction was performed using permutation-based family-wise error correction as implemented by the FWERperm option for the lesymap command (https://github.com/dorianps/LESYMAP).

To evaluate the relative contribution of the regional clusters of the lesion-behavior map of g, the map was first separated into clusters using FSL's cluster function. For each cluster, the identified region was removed from each patient's lesion mask and the lesion-behavior map was re-calculated. The importance of each cluster was then evaluated using the difference in the within-sample predictive correlation value associated with these analyses, with a greater reduction in predictive accuracy indicating greater relative importance of the associated cluster.

Fiber tractography

We used the LEAD Connectome pipeline to explore the white matter tracts associated with regions identified in the lesion-behavior map of g (https://www.lead-dbs.org/about/lead-connectome/; Horn et al., 2014). As performed previously (Horn et al., 2017), this software uses normative diffusion-weighted MRI from 32 subjects from the Human Connectome Project dataset to perform deterministic fiber tractography and illustrate the course of white matter streamlines associated with a seed region-of-interest. Seed regions-of-interest were generated by clustering the map of g into spatial clusters using FSL's cluster tool. The resulting tracts were reviewed individually on applying a threshold to retain the top 5% of voxel values, and the unthresholded individual tracts derived from each cluster were combined to display white matter regions common among the clusters. The Human Connectome Project (Van Essen et al., 2013) white matter atlas was used to compare the results to normative white matter tracts.

Validation analysis

WU cohort

Patients in the WU cohort also completed multiple neuropsychological tests. Test scores were normed as described previously (Corbetta et al., 2015). We converted these data to Z-scores. Participants were selected for inclusion in this study if they completed at least 75% of the selected tests and had lesion masks derived from structural neuroimaging. Test scores and neuroimaging data were collected within the first three months following the stroke. A total of 101 stroke patients met these inclusion criteria and were included in this study. There were no missing data in this dataset.

Lesion segmentation

MRI and lesion mask acquisition for the WU cohort were described in the original study (Corbetta et al., 2015). For consistency with the Iowa cohort, the final lesion masks from the WU cohort were transformed to the MNI152 template using linear registration techniques in FSL.

Hierarchical factor analysis

Because most of the tests available in the WU dataset have not been previously evaluated in the context of the CHC model, we used an exploratory hierarchical factor analysis to estimate g in the WU cohort. Although the extraction of g was the primary goal of this analysis, for consistency with the Iowa cohort, we labeled the subfactors using a nomenclature similar to the CHC model by comparing the test groupings to descriptions of CHC domains (Schneider and McGrew, 2018). This analysis was performed using the psych package available in R (Revelle, 2015). We used maximum likelihood to estimate parameters, and an oblimin rotation to extract correlated factors for the first level of latent variables. The number of subfactors to retain at the first level was determined by comparing the fit of models with successively higher numbers of factors until an improvement in model fit was not found. The factor scores for these domains were then submitted to an exploratory factor analysis to extract a single hierarchical g factor. As with the Iowa cohort, an estimate of g was derived for each subject using regression methods. A bifactor model was not evaluated for this sample because of low sample size.

Predicting g from lesion location

The sparse canonical correlation models used to derive lesion-behavior maps from the Iowa Cohort were used to generate predictions of cognitive deficits based on lesion location in the WU Cohort (Fig. 1). Specifically, a three-dimensional matrix of the voxel-wise weights from a lesion-behavior map was multiplied by a three-dimensional binary matrix representing the lesion volumes from the WU cohort, and then by the eigenvalue from the lesion-behavior map to generate a vector of model-predicted scores. This was repeated for each lesion-behavior map derived from the Iowa cohort. Model-predicted scores were then correlated with observed g to provide an index of predictive accuracy. Additionally, semi-partial correlations were used to test whether these associations would hold after controlling for the effects of lesion volume. The statistical significance of the correlation between each predicted and

observed score was determined using permutation testing. Specifically, the order of the data was randomly generated, and the absolute value of the observed correlation strength was compared with that of randomly generated correlations derived from the same data across 10,000 permutations. The proportion of randomly generated correlations greater than the observed correlation served as the *p* value. Finally, a lesion-behavior map of g was also generated *de novo* for the WU cohort, and we calculated its spatial correlation with the lesion-behavior maps from the Iowa cohort by converting all of the lesion-behavior maps into vectors using ANTsR tools and then deriving the correlations among the vectors.

Data availability

The data that support the findings of this study are available from the corresponding author, on reasonable request.

Results

Iowa cohort SEM

The final SEMs converged normally, provided adequate fit to the data, and were without Heywood cases or negative estimated variances (Fig. 2*A*,*B*).

The initial hierarchical model failed to converge; inspection of the factor loadings suggested that this was because of an extremely high raw loading parameter that was associated with a perfect correlation (standardized factor loading = 1.00) between g and working memory. The model converged successfully when this raw estimate was constrained to a lower value, but where the standardized loading was still reported as one by the lavaan output (χ^2 = 384.55, CFI = 0.90, RMSEA = 0.09, SRMR = 0.07). To further investigate the reason for the failed convergence of the initial hierarchical model, we constructed (1) a hierarchical model in which the working memory tests loaded directly onto g rather than onto a Gwm factor, and (2) a correlated factors model in which the first-order latent variables (e.g., Gc, Gv, etc.) were allowed to covary without loading onto a superordinate g factor. Both of these models converged successfully, suggesting that there may have been a model misspecification occurring at the loading of Gwm onto g. Together, these analyses provided initial evidence that there is not a Gwm factor that is reliably separate from g in the Iowa cohort.

In the initial bifactor model (χ^2 = 312.79, CFI = 0.92, RMSEA = 0.08, SRMR = 0.06), the factor loadings of two tests of working memory (the digit span subtest from the WAIS and the Benton Visuospatial Retention Test) on the Gwm latent variable were not significantly different from zero, indicating that the covariance among the working memory tests was captured entirely by g. The hierarchical and bifactor models both implied a perfect correspondence between g and working memory, and thus indicated the use of a bifactor S-1 model, a variant of the bifactor model that should be used in situations in which anomalous (e.g., non-significant or negative) loadings are detected in the presence of a perfect or near-perfect equivalence between a subfactor and a general factor (Burns et al., 2020). In a bifactor S-1 model, the loadings of the indicators on the subfactor in question are omitted to achieve a final result that is free from anomalous loadings. The bifactor S-1 model for the Iowa cohort converged normally (χ^2 = 340.06, CFI = 0.91, RMSEA = 0.08, SRMR = 0.07). Of note, the bifactor-S1 model also provided a significantly better fit to the data than the hierarchical model (χ^2 difference = 44.49, $p < 0.001$).

In the hierarchical model, Gwm accounted for more variance in g (100%) than did Gc (51%), Gv (65%), Gl (15%), or Gs (71%; Fig. 1*A*). Model-based reliabilities for each domain-specific factor were estimated using ω hierarchical (ω_h). The values of ω_h from the bifactor-S1 model associated with Gc (0.39), Gv (0.27), Gl (0.63), and Gs (0.39) were greater than that associated with Gwm, which was 0 by definition, such that there was no reliable variance in the latent Gwm factor beyond g. Even in the initial bifactor model, which contained two non-significant loadings of Gwm tests on the latent Gwm factor, the remaining variance of the Gwm factor was virtually zero (ω_h = 0.02), and might be best characterized as the remaining reliable variance in the arithmetic subtest from the WAIS working memory index rather than the remaining variance in broad Gwm factor.

Healthy cohort SEM

To evaluate whether this pattern of findings generalized to a healthy sample, we performed analogous analyses on the correlation matrix of the WAIS-IV primary subtests (Wechsler, 2008). Overall, we obtained a similar pattern of results. A hierarchical model demonstrated that working memory shared more variance with g (88%) relative to verbal comprehension (68%), perceptual reasoning (76%), or processing speed (48%; χ^2 = 339.59, CFI = 0.97, RMSEA = 0.07, SRMR = 0.03). An initial bifactor model constructed from the WAIS correlation matrix did not successfully converge because of the variance-covariance matrix not being positive definite, indicating a problem with model specification. As was the case with the Iowa cohort, a bifactor S-1 model was applied to the WAIS data where the working memory factor was omitted. The bifactor S-1 model allowed this analysis to converge successfully (χ^2 = 223.64, CFI = 0.98, RMSEA = 0.06, SRMR = 0.02), suggesting that working memory (ω_h = 0.00) was not distinct from g. In contrast, this was not the case for the other cognitive abilities: verbal comprehension (ω_h = 0.27), perceptual reasoning (ω_h = 0.20), and processing speed (ω_h = 0.41). Even in the initial bifactor model for the WAIS data, which did not converge, the remaining variance of the Gwm factor was the lowest among all cognitive abilities (ω_h = 0.10). Finally, as was the case in the Iowa cohort, the bifactor-S1 model provided a significantly better fit to the data than did the hierarchical model (χ^2 difference = 115.95, $p < 0.001$).

Together, the psychometric results from the Iowa cohort and the WAIS-IV data all demonstrate an exceptionally strong to perfect correspondence between g and working memory. We refer to the psychometric isomorphism between g and Gwm in the Iowa sample's hierarchical model as g/Gwm.

Iowa cohort lesion-behavior mapping

We next investigated whether the neuroanatomical regions associated with g/Gwm were similar to those associated with the domain-specific cognitive abilities. The distribution of lesion coverage used for these analyses is shown in Figure 3*A*. We used the hierarchical model because hierarchical models have traditionally been used to represent g (of note, the hierarchical and bifactor model estimates of g were virtually identical; $r = 0.99$, $p < 0.001$). The lesion-behavior map of g/Gwm was associated with a statistically significant within-sample predictive correlation ($r = 0.27$, $p < 0.001$, optimal sparseness = 0.17; Fig. 3*B*). This analysis demonstrated that g/Gwm depends on the integrity of left temporo-parietal white matter, left parieto-occipital white matter, left frontal lobe white matter, the right subinsula/claustrum, the left posterior middle/inferior temporal gyrus, and the right posterior middle temporal gyrus. Regressing out the effects of sex and lesion chronicity did not change the lesion-behavior mapping results, and neither gender ($r = 0.06$, $p = 0.16$) nor lesion chronicity ($r = 0.08$, $p = 0.10$) was correlated significantly with g/Gwm.

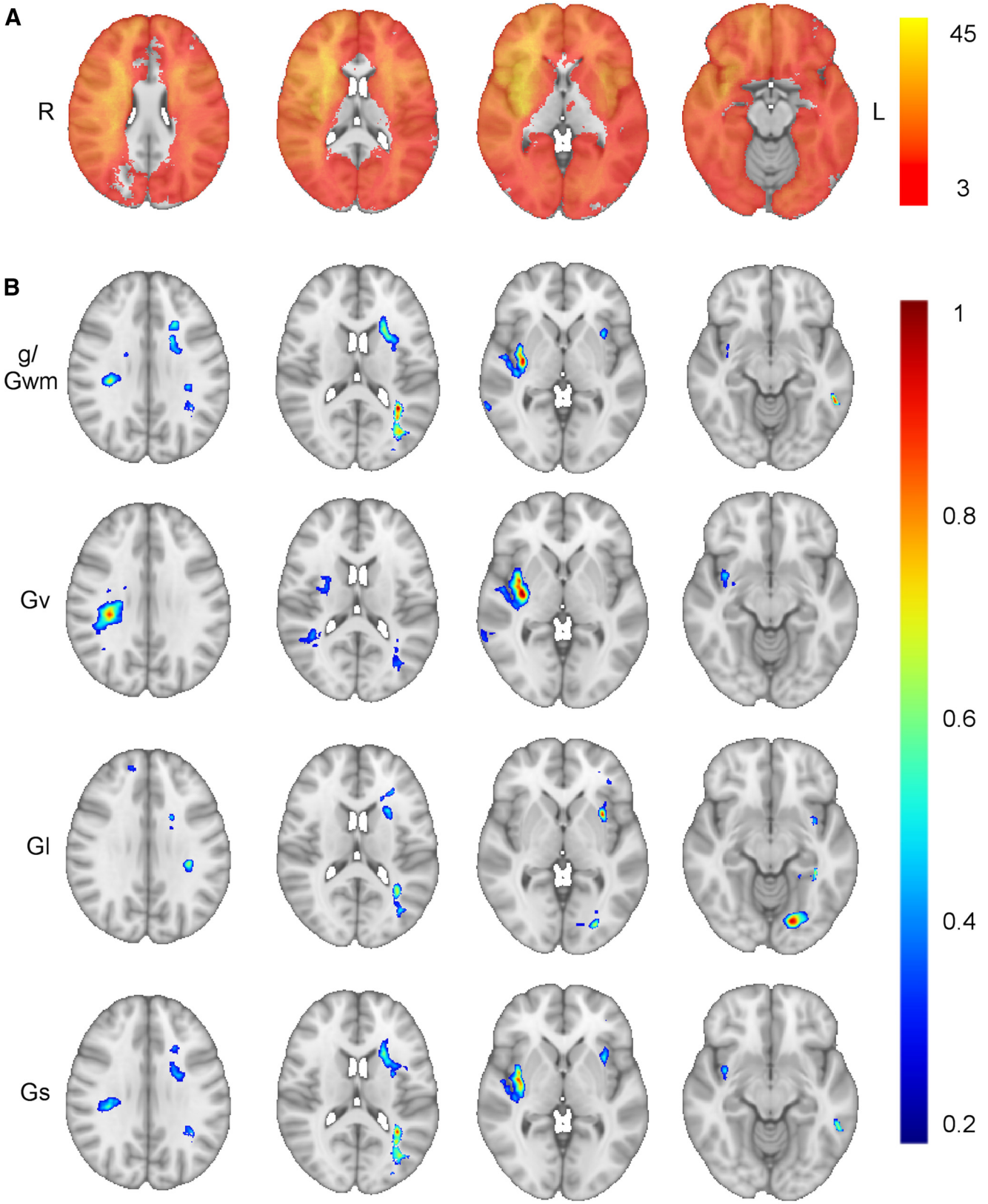

Figure 3. Lesion-behavior maps of general and domain-specific cognitive ability. ***A***, Distribution of the 402 lesions included in the lesion-behavior mapping, with the color scale representing the number of overlapping lesions at each voxel. Axial slices from the lesion-behavior maps derived from the hierarchical SEM are shown in panel ***B***, illustrating the brain regions associated with general cognitive ability/working memory (g/Gwm; $r = 0.27$, $p < 0.001$), visuospatial ability (Gv; $r = 0.30$, $p < 0.001$, optimal sparseness $= -0.20$), Gl ($r = 0.30$, $p < 0.001$, optimal sparseness $= 0.13$), and processing speed (Gs; $r = 0.28$, $p < 0.001$, optimal sparseness $= 0.19$). The map for crystalized intelligence was not statistically significant (Gc; $r = 0.10$, $p = 0.05$, optimal sparseness $= -0.15$). All lesion-behavior maps are shown at the same four axial slices of the MNI152 T1 1-mm template. Voxels weights were derived from the optimized SCCAN and were scaled to values between 0 and 1, with greater values indicating greater importance for the variable in question. LESYMAP's sparseness search routine generated negative sparseness values when there was a possibility for bidirectional voxel weights (for more information on sparseness values, see https://github.com/dorianps/LESYMAP/wiki/SCCAN-questions).

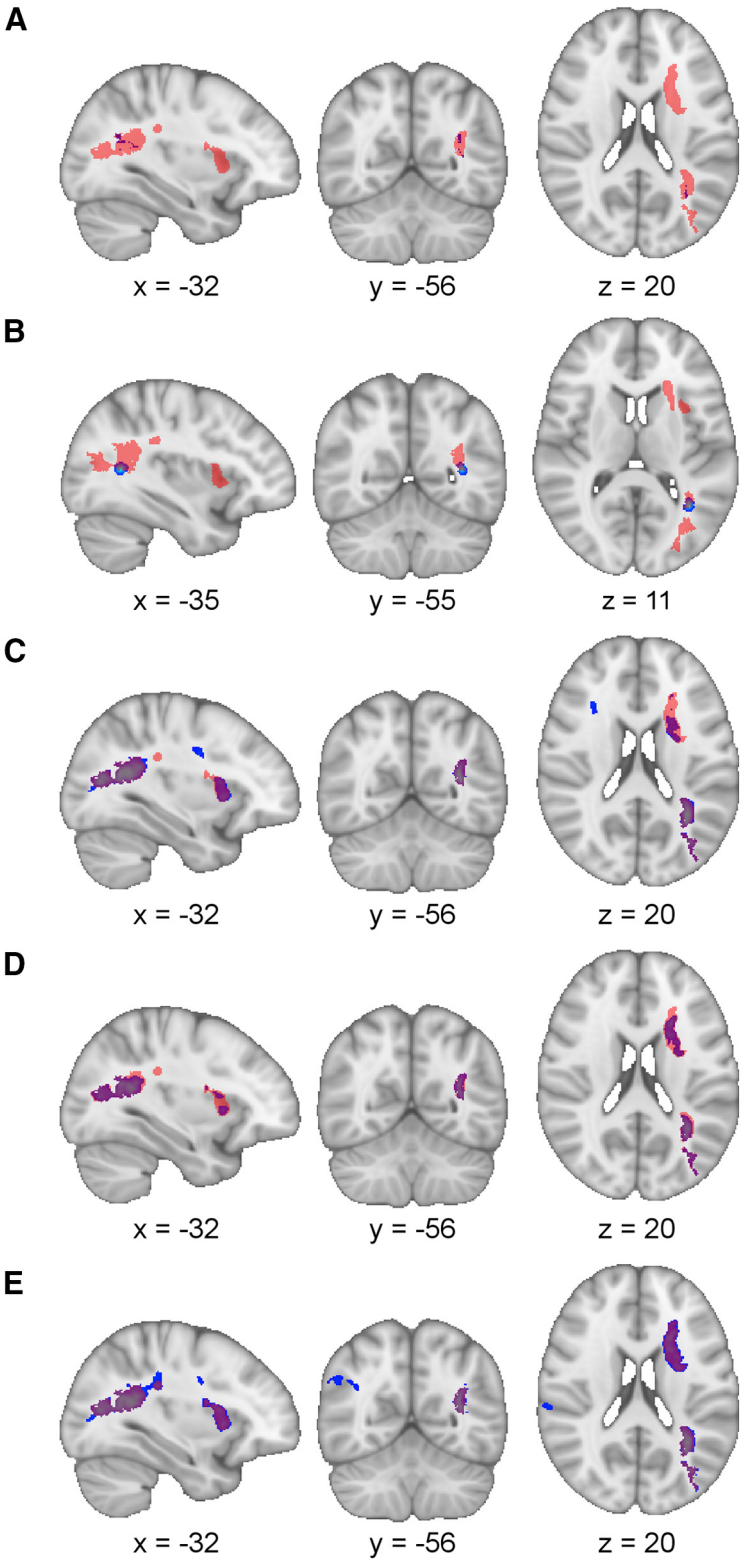

Figure 4. Additional lesion-behavior mapping analyses. We performed additional lesion-behavior mapping analyses (blue) to explore the aspects of the neuroanatomical localization of g (red) that were most robust to the analytic approach. The region of left posterior white matter deep to the temporo-parietal junction was still implicated when (***A***) using a voxel-wise ("mass univariate") approach to lesion-behavior mapping, (***B***) restricting analyses to only subjects with complete data ($n = 230$), (***C***) controlling for the effects of age and education, and decreasing (***D***) and increasing (***E***) the sparseness parameter of the final lesion-behavior map.

An evaluation of the importance of the individual clusters from this map showed that the cluster located in the left temporo-parietal white matter was the most important to the final solution ($\Delta r = -0.05$). This cluster also contained the voxel with the strongest association with g/Gwm in the entire map (MNI coordinate: −35, −50, 12). By comparison, the clusters located in the left temporal lobe white matter ($\Delta r = -0.02$), the left occipital lobe white matter ($\Delta r = 0.02$), the right posterior middle temporal gyrus ($\Delta r = -0.03$), the right parietal white matter ($\Delta r = 0.00$), the right subinsula/claustrum ($\Delta r = -0.03$), and the left frontal white matter and anterior insula ($\Delta r = -0.03$) were associated with lower changes in the within-sample predictive correlation. The left temporo-parietal white matter region was also implicated when: using a mass-univariate, voxel-wise approach to lesion-behavior mapping (Fig. 4*A*), restricting our analyses to the subsample of patients with complete data (Fig. 4*B*), controlling for age and education (Fig. 4*C*), and altering the sparseness parameter of the model (Fig. 4*D*,*E*). Examples of select individual subjects with relatively circumscribed damage to this region are shown in Figure 5.

Next, we performed two follow-up lesion-behavior mapping analyses to further evaluate the importance of this left temporo-parietal white matter region for domain-general cognition. First, if this region was important for cognition across multiple domains, a similar neuroanatomical localization for g would be expected when estimating g from a SEM that does not include tests of working memory, as variance related to g should still be included in the estimates of each specific domain of cognition. Second, if we remove the influence of g from the domain-specific abilities (using the bifactor model), the regions most associated with domain-general cognition should no longer be present in the domain-specific lesion-behavior maps. In both cases, our results support the importance of the left temporo-parietal white matter region.

First, we created a lesion-behavior map of g estimated without tests of working memory. This showed a strong peak in the same left temporo-parietal white matter region, and was highly similar to the map of g/Gwm in which working memory tests were included (spatial correlation $r = 0.91$). Moreover, the left temporo-parietal white matter region was also present in lesion-

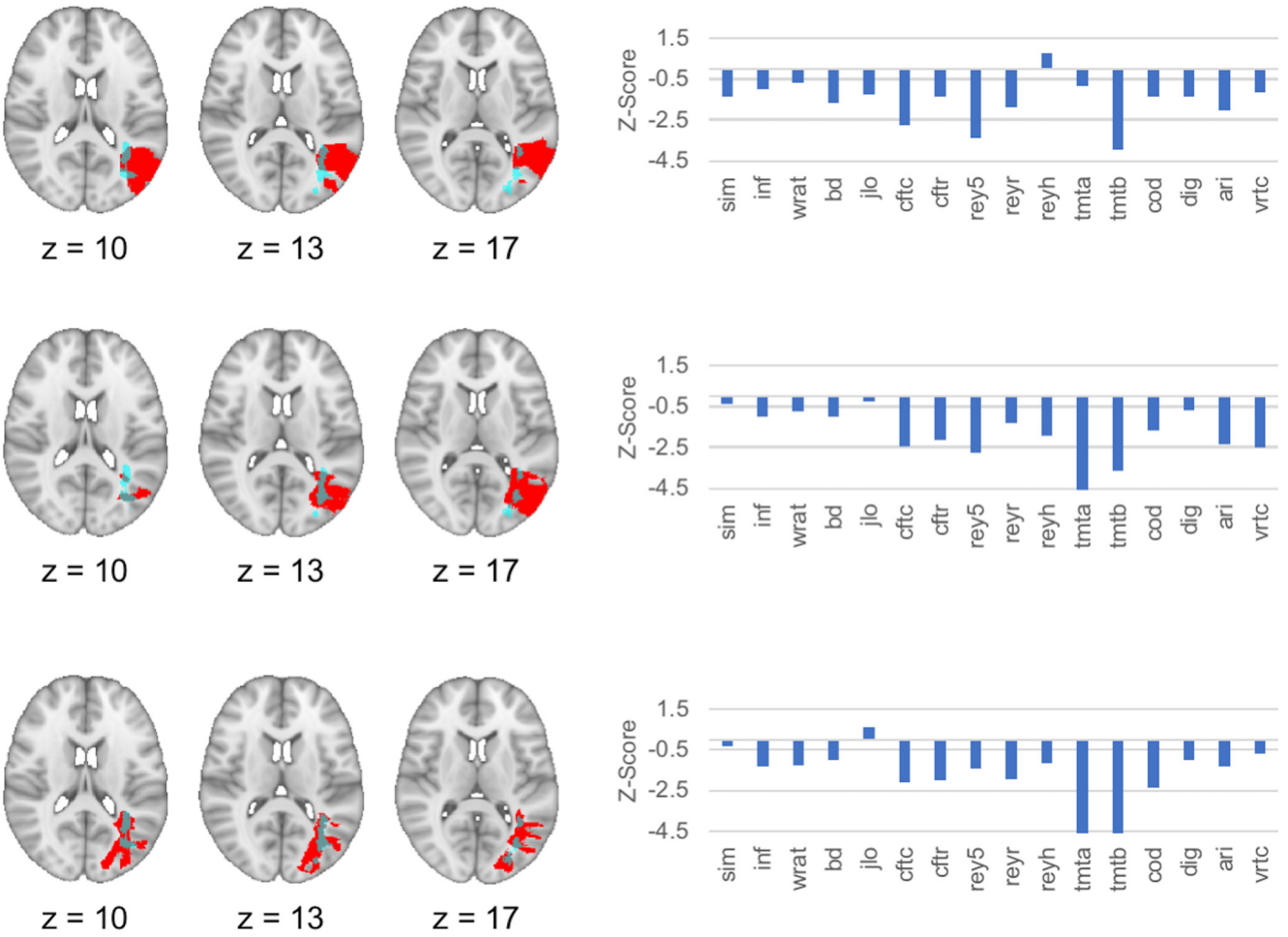

Figure 5. Representative individual cases with lesions involving the peak cluster associated with general cognitive ability. Lesion mask (red) and neuropsychological data are presented from three representative subjects with relatively focal lesions that involved the location of the peak cluster of the lesion-behavior map of general cognitive ability/working memory (g/Gwm). Neuropsychological test performances were broadly low and/or impaired across multiple domains for each subject, as represented by negative Z-scores relative to other subjects with focal brain lesions in the registry. The peak cluster from the lesion-behavior mapping analysis of g/Gwm is depicted in light blue. Abbreviations for individual tests are found in the legend of Figure 2.

behavior maps of the domain-specific cognitive abilities, but was absent after the influence of g was removed from these abilities in the bifactor model, which can be seen by comparing the maps in Figures 3*B*, 6, and through an examination of the spatial correlations between each hierarchical model's lesion-behavior map and its bifactor analog: Gv (0.60), Gl (0.48), and Gs (0.11). The spatial correlations between the map of g/Gwm and the maps of the other domain-specific cognitive abilities estimated from the bifactor model were low (spatial correlation r range: 0.04–0.16).

Finally, in light of the observation that estimates of g and Gwm were perfectly correlated in the hierarchical model from this sample, we conducted an additional follow-up analysis to explore whether g and working memory indeed share neuroanatomical correlates. We evaluated whether Gwm estimated from a correlated factors model (i.e., a model without a superordinate g factor) produced a similar lesion-behavior map to that of g/Gwm. This analysis produced a highly similar result (spatial correlation $r = 0.99$), with a peak in the same left temporo-parietal white matter region. Together, these analyses demonstrate evidence that working memory and g can be localized to this same white matter region outside of the data from the full Iowa hierarchical model.

Iowa cohort tractography

The map for g was separated into spatial clusters and each cluster served as a seed region-of-interest for a tractography analysis, allowing us to infer which white matter tracts are associated with these regions in healthy adults. The clusters, including the cluster of temporo-parietal white matter that was most strongly associated with g, were most prominently and consistently associated with the arcuate fasciculus (Fig. 7), a large white matter tract that interconnects the frontal, temporal, and parietal cortices (Catani and Thiebaut de Schotten, 2008).

Using the Iowa-derived lesion-behavior maps to predict g in the WU cohort

Next, we evaluated whether the lesion-behavior maps from the Iowa cohort could be used to predict observed deficits in g in the independent WU cohort based exclusively on lesion location. The exploratory factor analysis demonstrated that a three-factor solution fit the cognitive data the best, which converged normally and was without Heywood cases ($\chi^2 = 3.95$, CFI > 0.99, RMSEA < 0.01, SRMR < 0.01; Fig. 2*C*). The three subfactors corresponded most closely to Gc, visuospatial working memory (Gvwm), and Gl. We evaluated the validity of estimating g from this more limited set of domain-specific variables by also

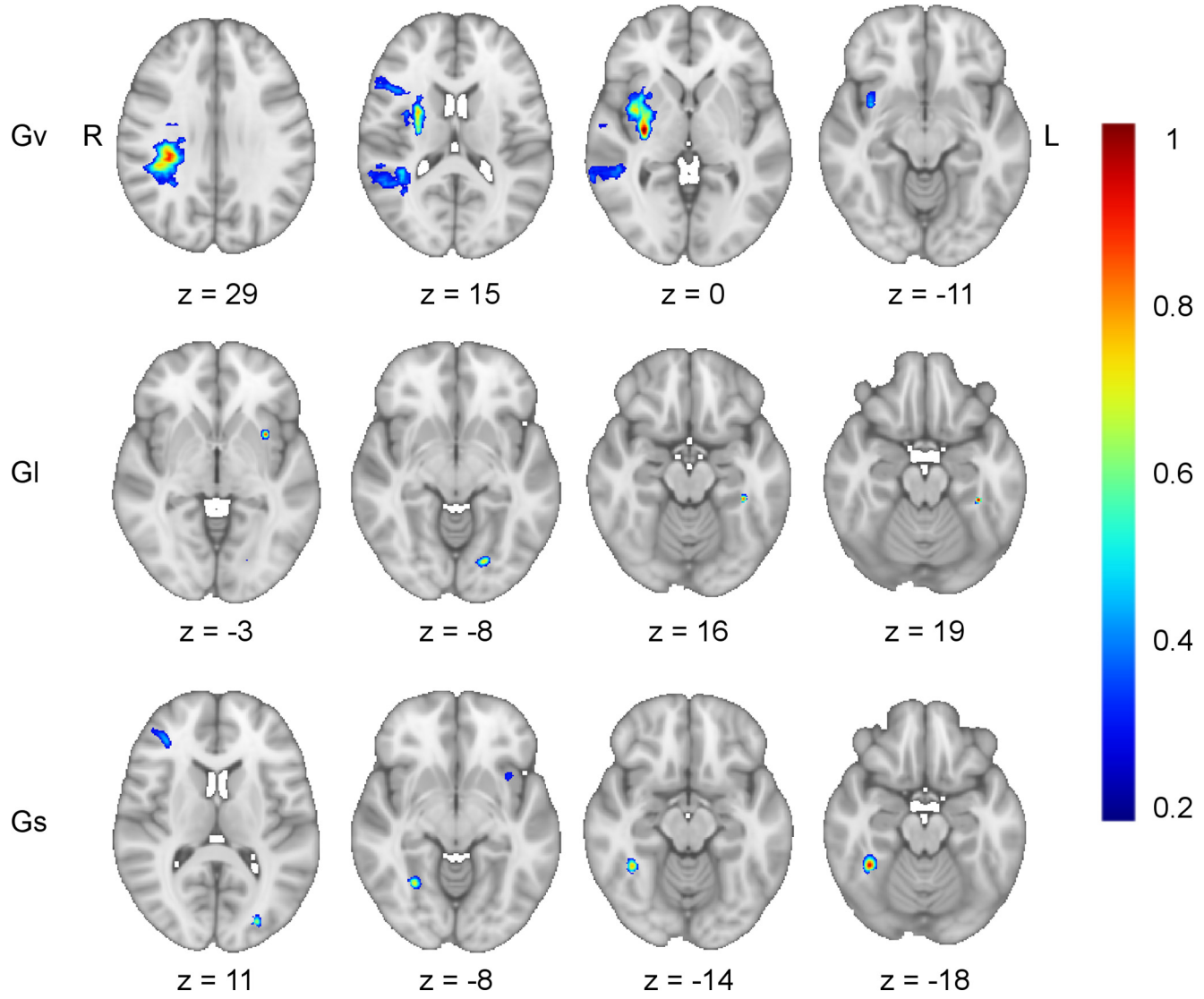

Figure 6. Bifactor model lesion-behavior mapping. Axial slices from the lesion-behavior maps estimated from the bifactor S-1 model illustrate the brain regions uniquely associated with each domain-specific cognitive ability after removing the variance attributable to general cognitive ability (g; see Fig. 2*B*). This included: visuospatial ability (Gv; $r = 0.34$, $p < 0.001$, optimal sparseness = −0.20), learning/memory (Gl; $r = 0.32$, $p < 0.001$, optimal sparseness = −0.01), and processing speed (Gs; $r = 0.19$, $p < 0.001$, optimal sparseness = 0.08). The map for crystalized intelligence (Gc; $r = 0.19$, $p < 0.001$, optimal sparseness = 0.77) did not return any regions associated with impairment in this domain. Working memory (Gwm) demonstrated no remaining reliable variance after accounting for g and so was not depicted. For each map, the left hemisphere temporo-parietal white matter region implicated as important for g/Gwm (compare with Fig. 3) was no longer present in any of these maps. Voxel weights were derived from the optimized SCCAN and were scaled to values between 0 and 1.

estimating g in the Iowa data, but restricting the analysis to these three domain-specific variables. The correlation between g estimated with all the data versus with the restricted domain-specific variables was 0.90, demonstrating that g can be estimated accurately with fewer domain-specific variables, as were available in the WU dataset (Johnson et al., 2004, 2008).

A matrix of lesion masks from the WU cohort (Fig. 1*C*) was used in conjunction with the three-dimensional weighted voxels derived from the hierarchical lesion-behavior maps from Iowa (Fig. 1*B*) in an effort to predict g in the WU cohort. To evaluate the anatomic specificity of these predictions, we compared the accuracy of the predictions from the g/Gwm map to the accuracy of the predictions from the other domain-specific lesion-behavior maps. A caveat to this cross-cohort prediction analysis is that the Iowa cohort had more extensive coverage in the brain whereas the coverage in the WU cohort was predominantly, although not exclusively, in the distribution of the middle cerebral artery, which resulted in limited coverage of the left temporo-parietal white matter region identified in the Iowa lesion-behavior g map. Many of the lesions in the WU cohort did not intersect with the lesion-behavior maps of the Iowa cohort. To account for this, we also evaluated the accuracy of the model-based predictions for only those WU subjects with lesions that overlapped with each of the Iowa lesion-behavior maps.

Across these analyses, the lesion-behavior map of g/Gwm from the Iowa sample provided consistent significant predictions of observed g in the WU sample, even after controlling for lesion volume. Table 2 lists the correlations (*r*) and semi-partial correlations (*sr*), the latter of which control for lesion volume, between the model-predicted domain scores and observed g in the WU sample. Importantly, the unique aspects of the other domain-specific maps (i.e., predictions made from the bifactor maps) were not able to significantly predict g in the WU cohort, supporting the anatomic specificity of these results.

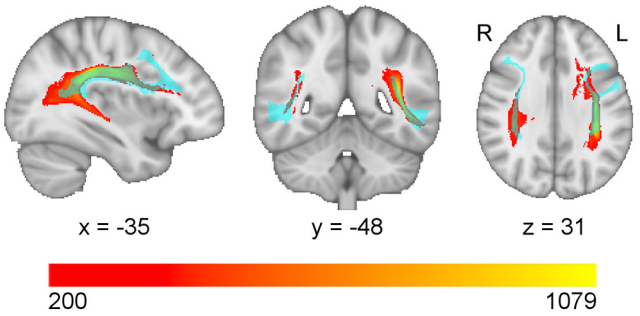

Figure 7. Seed-based fiber tractography. The white matter regions of the lesion behavior map of general cognitive ability/working memory (g/Gwm) were segmented into clusters and then each cluster was submitted as a seed region-of-interest in a deterministic fiber tractography analysis. This analysis produced voxel-wise maps indicating the number of white matter streamlines passing through each voxel that originated from the seed region-of-interest. The tractography maps were summed into a single map (depicted in a red-yellow palette, where more yellow colors indicate a higher number of streamlines), thresholded at 200 streamlines, and compared with the Human Connectome Project White Matter Atlas. There was a prominent involvement of the arcuate fasciculus (light blue).

Table 2. Association between model-predicted factors and observed g in the WU cohort

Model-predicted factor score	Correlation (*r*)	Hierarchical regression (*sr*)	Correlation (*r*) in subsample	Hierarchical regression (*sr*) in subsample
g/Gwm	0.42***	0.26**	0.50***	0.23*
Gv	0.24*	0.06	0.20*	−0.03
Gv-bifactor	0.15	−0.01	0.12	−0.11
Gl	0.31**	0.21*	0.51***	0.23*
Gl-bifactor	−0.06	0.01	−0.05	0.04
Gs	0.41***	0.25*	0.47***	0.19
Gs-bifactor	0.15	−0.02	0.14	−0.07

Subsamples in columns 4 and 5 were determined by removing subjects whose lesion masks did not intersect with the lesion-behavior map in question. The sample sizes for these more restricted subsample analyses were: g/Gwm ($n = 67$), Gv ($n = 50$), Gv-bifactor ($n = 62$), Gl ($n = 34$), Gl-bifactor ($n = 49$), Gs ($n = 74$), Gs-bifactor ($n = 66$). Statistical significance was determined using permutation tests; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Separately, a lesion-behavior map of g derived *de novo* from the WU cohort data was generated and produced results similar to those observed from the Iowa cohort: fronto-parietal and temporo-parietal white matter regions (including the arcuate fasciculus) were implicated bilaterally with stronger associations found in the left hemisphere. Additionally, the map of g from the WU cohort was the most similar to the map of g/Gwm (spatial correlation $r = 0.17$) from the Iowa cohort relative to the other domain-specific maps derived from the bifactor model: Gv ($r = 0.12$), Gl ($r = 0.04$), and Gs ($r = 0.07$). Together, the findings in the WU sample provide evidence corroborating the anatomic findings from the Iowa sample.

Discussion

In this study, we asked the question: what is the relation between general cognitive ability (g) and its subordinate factors? This question cannot be fully addressed in healthy individuals, as g is defined by the shared variance across these factors. However, it can be further elucidated in subjects with focal brain lesions since it is possible to discover dissociations between g and its subordinate factors that are not seen in healthy individuals. SEMs indicated that there is a strong correspondence between g and working memory; because these two variables could not be reliably separated in the primary sample, we refer to them collectively here as g/Gwm. Lesion-behavior mapping demonstrated that g and working memory depend on the integrity of a posterior region of left hemisphere white matter located in the arcuate fasciculus, a white matter bundle connecting the frontal, parietal, and temporal cortices that has also variably been referred to as the superior longitudinal fasciculus (Catani and Thiebaut de Schotten, 2008; Dick and Tremblay, 2012). The anatomic localization of g/Gwm was validated in an independent sample of patients with focal brain damage (the WU cohort), where a lesion-behavior map of g/Gwm from the Iowa cohort significantly predicted g in the WU cohort based on lesion location, demonstrating that our findings generalize robustly from chronic lesions to acute lesions. This finding is notable given that the recovery of cognitive functions is variable in the early months following stroke and this variability would likely make it more difficult to predict cognition across cohorts. We elaborate on our behavioral and neuroanatomical findings below and their relationship to the existing literature.

Converging perspectives from cognitive and systems neuroscience suggest that the core processes of working memory include the maintenance and control of representational information encoded in neuronal ensembles (D'Esposito and Postle, 2015; Eriksson et al., 2015; Miller et al., 2018). The strong association that we observed between g and working memory is comparable to what has been reported by previous research on this topic (Kyllonen and Christal, 1990; Süß et al., 2002; Mackintosh and Bennett, 2003; Colom, 2004; Chen et al., 2019). However, it is difficult to make direct comparisons between previous work and the present study. Here, we estimated g using the covariance among a large and diverse battery of cognitive tests (Johnson et al., 2004, 2008). In contrast, many previous studies used fluid reasoning as a proxy for g. There is also significant heterogeneity in how previous studies have measured working memory, reflecting the expansive proliferation of different working memory theories over the last century (Engle, 2002; Unsworth and Engle, 2007; Baddeley, 2012; Cowan, 2017; Oberauer, 2019). Here, we modeled working memory using diverse, widely used clinical tests, which can capture the prominent covariance among working memory tests that exists regardless of their theoretical assumptions (Wilhelm et al., 2013; Chen et al., 2019) or their canonical status as "clinical" or "laboratory" tests (Shelton et al., 2009). Many previous studies have relied on tests from a specific measurement paradigm.

Our findings suggest that working memory is constitutive of g and that we can build on our understanding of the mechanisms of domain-general cognition by reframing individual differences in g as being largely driven by individual differences in working memory. Although working memory is a complex cognitive ability associated with the interactions among widely distributed brain regions (Christophel et al., 2017), our data, and data from others (Pineda-Pardo et al., 2016), highlight the importance of efficient, direct connectivity between the prefrontal and posterior cortices in particular. This is consistent with the emphasis of the fronto-parietal integration theory on the arcuate fasciculus and fronto-parietal connectivity (Jung and Haier, 2007). This may be the case for at least two reasons: First, direct projections from the prefrontal cortex to posterior brain regions allow goals and task demands represented in the former to efficiently modulate mental representations encoded in the latter (Miller et al., 2018). Second, because the prefrontal and posterior cortices both encode stimulus information during working memory tasks (Ester et al., 2015), the direct connections between them might

facilitate the protection of representations from disruption by ongoing sensory processing (Jacob and Nieder, 2014). The region of the brain with the highest density of these fronto-posterior projections is located in the left temporo-parietal white matter region highlighted by the present study (Owen et al., 2015), which may be why this region is important for working memory.

Our results suggest that g is a "left hemisphere dominant" ability. This could reflect the relative importance of sequential and/or logical information processing, for which the left hemisphere is specialized (Gazzaniga et al., 2014), for domain-general cognition. Alternatively, it may be the case that our neuropsychological tests rely heavily on verbal processing, including the comprehension of test instructions. It remains unclear whether the present results demonstrate a true left hemisphere dominance for g, or whether this finding is derivative to the way tasks are typically instructed to persons.

Our results highlight the importance of white matter to a greater extent than gray matter (Li et al., 2009; Tang et al., 2010). Our lesion-behavior mapping analyses were sufficiently powered to detect findings over much of the cerebral cortex, yet identified deep white matter regions as the most strongly associated with g. Although white matter sites are difficult to relate to specific functional brain networks using normative data, white matter lesions have been observed to disrupt functional networks, a mechanism which might have contributed to the current findings (Griffis et al., 2019). Prior work has shown that the functional connectivity between frontal and parietal brain regions is associated with diverse cognitive tasks (Duncan, 2010), cognitive control (Yeo et al., 2011), and g (Vakhtin et al., 2014; Santarnecchi et al., 2017; Dubois et al., 2018), possibly reflecting the involvement of executive functions, including working memory, in multiple forms of cognition (Camilleri et al., 2018). Consistent with the results we report here, previous lesion studies of g have also implicated left fronto-parietal white matter; however, previous studies have also implicated larger clusters of voxels, potentially because of the use of mass-univariate approaches (Glascher et al., 2010; Woolgar et al., 2010; Barbey et al., 2012, 2013). The current findings sharpen previous work by showing a particular emphasis on the arcuate fasciculus.

This study highlights the need for process specificity in studying brain-behavior relationships. We show that the domain-general influence of working memory contributes to individual differences in other mental functions. Without correcting for the influence of working memory, research that aims to focus on a specific cognitive process may be unintentionally studying the confluence of working memory and the intended domain of study, rather than the unique aspects of the targeted process. We believe this concept has been borne out in several language studies highlighting the importance of the same left temporo-parietal white matter region that we identified here (Dronkers et al., 2004; Geva et al., 2012; Baldo et al., 2013; Butler et al., 2014; Henseler et al., 2014; Harvey and Schnur, 2015; Yourganov et al., 2016; Griffis et al., 2017; Diachek et al., 2020). Although language tasks in our sample also implicate this left temporo-parietal region, this region is absent upon regressing out the influence of g. The integration of quantitative cognitive models into cognitive neuroscience may become increasingly important in elucidating domain-specific brain-behavior relationships.

There are important caveats to our conclusions. Our latent variable for working memory was estimated by collapsing across verbal and visuospatial modalities. Although a meta-analysis of fMRI studies has shown that this conjunction into "supramodal" working memory is associated with fronto-parietal activity (Rottschy et al., 2012), verbal and Gvwm can also constitute their own subdomains of working memory. There was only one test of Gvwm widely available in the Iowa sample, and so we were not able to estimate separate latent variables for each domain. Relatedly, working memory is a multifaceted construct and further research would be required to evaluate whether a specific aspect of working memory drives its association with g. Additional research will be needed to further evaluate the contribution of working memory to g relative to other aspects of executive functioning (Friedman et al., 2006), and to explore potential sex (Haier et al., 2005) and lesion-mechanism differences in the relationships among g, working memory, and neuroanatomy. Moreover, future research could seek to clarify the causal direction of the association between g and lesions to the posterior white matter region identified (Modig Wennerstad et al., 2010). Finally, we wish to emphasize the possibility that future studies with larger samples or different tests, or case-studies with very focal lesions, might find some degree of dissociation between g and working memory.

In summary, the present study provides evidence that working memory is constitutive of general cognitive ability (g). We present the largest lesion-behavior mapping analysis of g to date that combines SEM and multivariate lesion-behavior mapping, and where further support is garnered by predicting g from lesion location in an independent cohort. These results inform the study of domain-general cognition.

References

Ackerman PL, Beier ME, Boyle MO (2002) Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. J Exp Psychol Gen 131:567–589.

Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12:26–41.

Baddeley A (2012) Working memory: theories, models, and controversies. Annu Rev Psychol 63:1–29.

Baldo JV, Arévalo A, Patterson JP, Dronkers NF (2013) Grey and white matter correlates of picture naming: evidence from a voxel-based lesion analysis of the Boston Naming Test. Cortex 49:658–667.

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

Barbey AK, Colom R, Grafman J (2013) Dorsolateral prefrontal contributions to human intelligence. Neuropsychologia 51:1361–1369.

Bollen KA, Noble MD (2011) Structural equation models and the quantification of behavior. Proc Natl Acad Sci USA 108:15639–15646.

Brett M, Leff AP, Rorden C, Ashburner J (2001) Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage 14:486–500.

Burns GL, Geiser C, Servera M, Becker SP, Beauchaine TP (2020) Application of the bifactor S - 1 model to multisource ratings of ADHD/ODD symptoms: an appropriate bifactor model for symptom ratings. J Abnorm Child Psychol 48:881–894.

Butler RA, Lambon Ralph MA, Woollams AM (2014) Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain 137:3248–3266.

Camilleri JA, Müller VI, Fox P, Laird AR, Hoffstaedter F, Kalenscher T, Eickhoff SB (2018) Definition and characterization of an extended multiple-demand network. Neuroimage 165:138–147.

Catani M, Thiebaut de Schotten M (2008) A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44:1105–1132.

Chen Y, Spagna A, Wu T, Kim TH, Wu Q, Chen C, Wu Y, Fan J (2019) Testing a cognitive control model of human intelligence. Sci Rep 9:2898.

Christophel TB, Klink PC, Spitzer B, Roelfsema PR, Haynes JD (2017) The distributed nature of working memory. Trends Cogn Sci 21:111–124.

Colom R (2004) Working memory is (almost) perfectly predicted by g. Intelligence 32:277–296.

Conway ARA, Cowan N, Bunting MF, Therriault DJ, Minkoff SRB (2002) A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. Intelligence 30:163–183.

Conway AR, Kane MJ, Engle RW (2003) Working memory capacity and its relation to general intelligence. Trends Cogn Sci 7:547–552.

Corbetta M, Ramsey L, Callejas A, Baldassarre A, Hacker CD, Siegel JS, Astafiev SV, Rengachary J, Zinn K, Lang CE, Connor LT, Fucetola R, Strube M, Carter AR, Shulman GL (2015) Common behavioral clusters and subcortical anatomy in stroke. Neuron 85:927–941.

Cowan N (2017) The many faces of working memory and short-term storage. Psychon Bull Rev 24:1158–1170.

D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. Annu Rev Psychol 66:115–142.

Damasio H, Frank R (1992) Three-dimensional in vivo mapping of brain lesions in humans. Arch Neurol 49:137–143.

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

Diachek E, Blank I, Siegelman M, Affourtit J, Fedorenko E (2020) The domain-general multiple demand (MD) network does not support core aspects of language comprehension: a large-scale fMRI investigation. J Neurosci 40:4536–4550.

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

Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. Cognition 92:145–177.

Dubois J, Galdi P, Paul LK, Adolphs R (2018) A distributed brain network predicts general intelligence from resting-state human neuroimaging data. Philos Trans R Soc Lond B Biol Sci 373:20170284.

Duncan J (2010) The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn Sci 14:172–179.

Engle RW (2002) Working memory capacity as executive attention. Curr Dir Psychol Sci 11:19–23.

Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L (2015) Neurocognitive architecture of working memory. Neuron 88:33–46.

Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87:893–905.

Fiez JA, Damasio H, Grabowski TJ (2000) Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. Hum Brain Mapp 9:192–211.

Frank RJ, Damasio H, Grabowski TJ (1997) Brainvox: an interactive, multimodal visualization and analysis system for neuroanatomical imaging. Neuroimage 5:13–30.

Friedman NP, Miyake A, Corley RP, Young SE, Defries JC, Hewitt JK (2006) Not all executive functions are related to intelligence. Psychol Sci 17:172–179.

Gazzaniga MS, Ivry RB, Mangun GR (2014) Hemispheric specialization. In: Cognitive neuroscience: the biology of the mind, Ed 4. New York: W. W. Norton and Company.

Geva S, Baron JC, Jones PS, Price CJ, Warburton EA (2012) A comparison of VLSM and VBM in a cohort of patients with post-stroke aphasia. Neuroimage Clin 1:37–47.

Glascher J, Rudrauf D, Colom R, Paul LK, Tranel D, Damasio H, Adolphs R (2010) Distributed neural system for general intelligence revealed by lesion mapping. Proc Natl Acad Sci USA 107:4705–4709.

Griffis JC, Nenert R, Allendorfer JB, Szaflarski JP (2017) Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. Neuroimage Clin 14:552–565.

Griffis JC, Metcalf NV, Corbetta M, Shulman GL (2019) Structural disconnections explain brain network dysfunction after stroke. Cell Rep 28:2527–2540.e9.

Haier RJ (2017) The neuroscience of intelligence. New York: Cambridge University Press.

Haier RJ, Jung RE, Yeo RA, Head K, Alkire MT (2005) The neuroanatomy of general intelligence: sex matters. Neuroimage 25:320–327.

Harvey DY, Schnur TT (2015) Distinct loci of lexical and semantic access deficits in aphasia: evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. Cortex 67:37–58.

Henseler I, Regenbrecht F, Obrig H (2014) Lesion correlates of patholinguistic profiles in chronic aphasia: comparisons of syndrome-, modality- and symptom-level assessment. Brain 137:918–930.

Hindman J, Bowren MD, Bruss J, Wright B, Geerling JC, Boes AD (2018) Thalamic strokes that severely impair arousal extend into the brainstem. Ann Neurol 84:926–930.

Horn A, Ostwald D, Reisert M, Blankenburg F (2014) The structural-functional connectome and the default mode network of the human brain. Neuroimage 102:142–151.

Horn A, Neumann WJ, Degen K, Schneider GH, Kühn AA (2017) Toward an electrophysiological "sweet spot" for deep brain stimulation in the subthalamic nucleus. Hum Brain Mapp 38:3377–3390.

Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 6:1–55.

Jacob SN, Nieder A (2014) Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. Neuron 83:226–237.

Jewsbury PA, Bowden SC, Duff K (2017) The Cattell–Horn–Carroll model of cognition for clinical assessment. J Psychoeduc Assess 35:547–567.

Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II (2004) Just one g: consistent results from three test batteries. Intelligence 32:95–107.

Johnson W, Nijenhuis J, Bouchard TJ (2008) Still just 1 g: consistent results from five test batteries Intelligence 36:81–95.

Jung RE, Haier RJ (2007) The parieto-frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav Brain Sci 30:135–154, discussion 154–187.

Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RW (2004) The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. J Exp Psychol Gen 133:189–217.

Kane MJ, Hambrick DZ, Conway ARA (2005) Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). Psychol Bull 131:66–71.

Kline RB (2016) Principles and practice of structural equation modeling, Ed 4. New York: The Guilford Press.

Kyllonen PC, Christal RE (1990) Reasoning ability is (little more than) working-memory capacity? Intelligence 14:389–433.

Li Y, Liu Y, Li J, Qin W, Li K, Yu C, Jiang T (2009) Brain anatomical network and intelligence. PLoS Comput Biol 5:e1000395.

Mackintosh N, Bennett ES (2003) The fractionation of working memory maps onto different components of intelligence. Intelligence 31:519–531.

Miller EK, Lundqvist M, Bastos AM (2018) Working memory 2.0. Neuron 100:463–475.

Mitrushina M, Boone KB, Razani J, D'Elia LF (2005) Handbook of normative data for neuropsychological assessment. New York: Oxford University Press.

Muthukrishna M, Henrich J (2019) A problem in theory. Nat Hum Behav 3:221–229.

Modig Wennerstad K, Silventoinen K, Tynelius P, Bergman L, Rasmussen F (2010) Association between intelligence and type-specific stroke: a population-based cohort study of early fatal and non-fatal stroke in one million Swedish men. J Epidemiol Community Health 64:908–912.

Oberauer K (2019) Working memory capacity limits memory for bindings. J Cogn 2:40.

Oberauer K, Schulze R, Wilhelm O, Süß HM (2005) Working memory and intelligence–their correlation and their relation: comment on Ackerman, Beier, and Boyle (2005). Psychol Bull 131:61–65, author reply 61–65.

Owen JP, Chang YS, Mukherjee P (2015) Edge density imaging: mapping the anatomic embedding of the structural connectome within the white matter of the human brain. Neuroimage 109:402–417.

Pineda-Pardo JA, Martínez K, Román FJ, Colom R (2016) Structural efficiency within a parieto-frontal network and cognitive differences. Intelligence 54:105–116.

Pustina D, Avants B, Faseyitan OK, Medaglia JD, Coslett HB (2018) Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. Neuropsychologia 115:154–166.

R Core Team (2017) R: a language and environment for statistical computing, Ed 3.6.1. Vienna: Austria.

Revelle W (2015) psych: procedures for personality and psychological research. In: R package version 1.9.12 edition. Evanston: Northwestern University.

Rosseel Y (2012) lavaan: an R package for structural equation modeling. J Stat Soft 48:1–36.

Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB (2012) Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage 60:830–846.

Santarnecchi E, Emmendorfer A, Tadayon S, Rossi S, Rossi A, Pascual-Leone A (2017) Network connectivity correlates of variability in fluid intelligence performance. Intelligence 65:35–47.

Schneider WJ, McGrew KS (2018) The Cattell-Horn-Carroll theory of cognitive abilities. In: Contemporary intellectual assessment: theories tests, and issues, Ed 4 (Flanagan DP, McDonough EM, eds). New York: The Guilford Press.

Shelton JT, Elliott EM, Hill BD, Calamia MR, Gouvier WD (2009) A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. Intelligence 37:283.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23[Suppl 1]:S208–S219.

Spearman C (1904) General intelligence objectively determined and measured. Am J Psychol 15:201–293.

Süß HM, Oberauer K, Wittmann WW, Wilhelm O, Schulze R (2002) Working-memory capacity explains reasoning ability—and a little bit more. Intelligence 30:261–288.

Tang CY, Eaves EL, Ng JC, Carpenter DM, Mai X, Schroeder DH, Condon CA, Colom R, Haier RJ (2010) Brain networks for working memory and factors of intelligence assessed in males and females with fMRI and DTI. Intelligence 38:293–303.

Tranel D (2007) Theories of clinical neuropsychology and brain-behavior relationships: Luria and beyond. In: Textbook of clinical neuropsychology (Morgan JE, Ricker JH, eds), pp 27–37. New York: Taylor and Francis.

Unsworth N, Engle RW (2007) On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. Psychol Bull 133:1038–1066.

Vakhtin AA, Ryman SG, Flores RA, Jung RE (2014) Functional brain networks contributing to the parieto-frontal integration theory of intelligence. Neuroimage 103:349–354.

van Buuren S, Groothuis-Oudshoorn K (2011) mice: multivariate imputation by chained equations in R. J Stat Soft 45:1–67.

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, WU-Minn HCP Consortium (2013) The WU-Minn Human Connectome Project: an overview. Neuroimage 80:62–79.

Wasserman JD (2018) A history of intelligence assessment: the unfinished tapestry. In: Contemporary intellectual assessment: theories, tests, and issues, Ed 4 (Flanagan DP, Harrison PL, eds), pp 3–55. New York: Guilford Press.

Wechsler D (2008) Wechsler adult intelligence scale: WAIS-IV technical and interpretive manual. London: Pearson.

Wilhelm O, Hildebrandt A, Oberauer K (2013) What is working memory capacity, and how can we measure it? Front Psychol 4:433.

Wilkinson GS, Robertson GJ (2006) WRAT4 wide range achievement test professional manual, Ed 4. Lutz: Psychological Assessment Resources.

Woolgar A, Parr A, Cusack R, Thompson R, Nimmo-Smith I, Torralva T, Roca M, Antoun N, Manes F, Duncan J (2010) Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. Proc Natl Acad Sci USA 107:14899–14902.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165.

Yourganov G, Fridriksson J, Rorden C, Gleichgerrcht E, Bonilha L (2016) Multivariate connectome-based symptom mapping in post-stroke patients: networks supporting language and speech. J Neurosci 36:6668–6679.