ARTICLE IN PRESS

Trends in Neuroscience and Education xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Trends in Neuroscience and Education

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

Individual differences in the neurobiology of fluid intelligence predict responsiveness to training: Evidence from a comprehensive cognitive, mindfulness meditation, and aerobic exercise intervention

Ana M. Daugherty[a,b,c,d,*], Bradley P. Sutton[d,e], Charles H. Hillman[f], Arthur F. Kramer[f,g,h], Neal J. Cohen[d,i,j,k], Aron K. Barbey[d,e,i,j,k,*]

[a] Department of Psychology, Wayne State University, Detroit, MI, USA
[b] Department of Psychiatry and Behavioral Neurosciences, Wayne State University, Detroit, MI, USA
[c] Institute of Gerontology, Wayne State University, Detroit, MI, USA
[d] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana Champaign, 405 North Mathews Avenue, Urbana, IL 61801, USA
[e] Department of Bioengineering, University of Illinois, Urbana, IL, USA
[f] Department of Psychology, Northeastern University, Boston, MA, USA
[g] Office of the Provost, Northeastern University, Boston, MA, USA
[h] Department of Mechanical & Industrial Engineering, Northeastern University, Boston, MA, USA
[i] Department of Psychology, University of Illinois, Urbana, IL, USA
[j] Neuroscience Program, University of Illinois, Urbana, IL, USA
[k] Center for Brain Plasticity, University of Illinois, Urbana, IL, USA

ARTICLE INFO

Keywords:
Fluid intelligence
Latent class analysis
Parahippocampal gyrus
Insula cortex
Middle frontal gyrus

ABSTRACT

Background: Fluid intelligence (G_f) is a critical cognitive ability that is predictive of real-world outcomes, and it has been a persistent aim to characterize its neural architecture.
Procedure: We advance our prior research by applying latent class analysis to evaluate individual differences in the neural and cognitive foundations of G_f over the course of a 16-week randomized, multi-modal intervention trial in neurologically healthy, younger adults ($N = 424$).
Results: Controlling for pre-intervention ability, three latent classes described individual performance at post-intervention and one group ($n = 71$) showed greater gains in visuospatial reasoning and high performance at post-intervention. The high performance group was predicted by larger anterior cingulate cortex, caudate and hippocampus volumes, and smaller middle frontal, insula and parahippocampal cortex volumes.
Conclusion: Regions that support cognitive control, working memory, and relational processes differentiated individuals who had higher G_f ability at pre-intervention and demonstrated a cumulative better response to the intervention.

1. Introduction

General intelligence predicts performance for a wealth of real-world outcomes across the lifespan—scholastic achievement [1,2], job performance [3,4], and career success [5]—and thus, it is an appealing target for interventions aimed to promote cognitive function and brain health. General intelligence reflects two primary facets of intellectual ability [6]: (i) *crystallized intelligence*, which supports problem solving in familiar contexts that rely upon prior knowledge and experience, and (ii) *fluid intelligence* (G_f), which enables problem solving in novel environments that require flexible, adaptive behavior [7,8]. G_f demonstrates dynamic changes throughout the lifespan, whereas crystalized intelligence continues to develop into middle age and remains stable until the ninth decade [9]. Interventions have aimed to bolster G_f with mixed results [10–13]. Recently, we reported a large-scale multimodal intervention that demonstrated modest gains in visuospatial reasoning within a group that engaged in equal parts cognitive training and exercise as compared to an active control [14]. We observed a large degree of individual differences not only at pre-intervention, but also in the magnitude of change in G_f test performance over the course of the intervention. Accumulating evidence further suggests that individual differences in brain structure and function account for variability in G_f

* Corresponding authors at: Decision Neuroscience Laboratory, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana Champaign, 405 North Mathews Avenue, Urbana, IL 61801, USA. Web: http://DecisionNeuroscienceLab.org.
E-mail addresses: ana.daugherty@wayne.edu (A.M. Daugherty), barbey@illinois.edu (A.K. Barbey).

https://doi.org/10.1016/j.tine.2019.100123

[15,16] and may therefore predict responsiveness to intervention. The present study sought to investigate this hypothesis, examining individual differences in the neurobiological foundations of fluid intelligence and their contributions to performance in a 16-week randomized, multi-modal intervention trial.

Cognitive tests of G_f measure individual differences in adaptive reasoning and problem solving skills, which reflect a broad set of abilities, including cognitive control, working memory, reasoning, and decision making [17]. Given the ontology of G_f, it is perhaps not surprising that an equally broad list of neural substrates have been identified. These brain regions can be conceptualized as falling into two categories: those that are putative substrates of fluid intelligence and those that belong to systems that enable it, including mechanisms for attention, learning, and memory.

For example, regions within the frontal and parietal cortices are known to play a central role in G_f [18–24]. The prefrontal cortex supports reasoning and decision making functions [16,18,19,25,26] and was among the first neural correlates identified [27]. However, prefrontal cortical function is not specific to G_f; for example, the middle frontal and orbitofrontal gyri are known to support working memory function [28,29]. Working memory functions are critical to G_f [18,19,30,31] and are a common target for interventions aimed to promote fluid intelligence. Working memory engages several brain regions, including the orbitofrontal cortex, caudate nucleus, cerebellum, insula cortex and prefrontal cortex [32] and it is plausible that factors that promote better function of these regions may bolster G_f.

A second system implicated in G_f is cognitive control [18,19]. Cognitive control is closely related to working memory function and response inhibition [33–35], which mediate its effects on G_f [36]. Indeed, cognitive control is a putative mechanism of training effects on G_f [37]. The anterior cingulate cortex is a canonical correlate of cognitive control and is commonly engaged in functional neuroimaging studies of G_f [21,38]. Cognitive control is vulnerable to stress and positivity bias, which has been attributed to its functional connection with the amygdala [39,40]. Although the amygdala has not been directly implicated in G_f, it mediates responses to stress and situational factors that are known to impair test performance [41–44]. Mindfulness techniques improve cognitive control [45] and mitigate effects of situational stress [42,43], and mindfulness interventions have been shown to improve test performance [41–44,46]. Thus, the anterior cingulate cortex and amygdala that support directed attention and mediate effects of stress may indirectly contribute to G_f.

Finally, a network of regions that are best described by their contribution to declarative or relational memory have been identified as correlates of G_f across the lifespan. Fluid reasoning is enabled by the capacity to flexibly adapt prior knowledge and experience to novel contexts. These core skills – adaptive learning and memory – are primary functions of medial temporal lobe regions [47], and although these regions are not commonly reported in studies specifically of G_f, there is accumulating evidence to support this notion. Larger hippocampal volume is associated with higher G_f in adults [48] and longitudinal shrinkage correlates with declines in G_f in adult aging [49,50]. The hippocampus, and the adjacent parahippocampal cortex, are critical for spatial cognition and visuospatial reasoning [51,52] that are typical targets of G_f assessments. The hippocampal internal circuitry and connectivity with the parahippocampal cortex are essential for pattern detection and comparisons in memory [53–55]. The ability to compare current information to past experience is necessary for adaptive reasoning and problem solving, and may be facilitated by the functional connection between the hippocampus and prefrontal cortex [47,56,57].

In summary, neural mechanisms for cognitive control, working memory, and relational processing have been identified to support G_f. However, few studies have considered the broad collection of brain regions simultaneously to evaluate their relative contributions to the neural architecture of G_f. Therefore, it is unclear what neural substrates may be most relevant to individual differences in G_f ability and, further, may be targets to promote cognitive function.

We address this issue in the current study and take a novel approach to evaluate individual differences in G_f over the course of a 16-week, multi-modal randomized controlled trial (RCT) via exploratory latent class analysis. The multi-modal intervention included aerobic exercise (Fit), combined exercise and cognitive training (Fit-MF), a third condition that further added mindfulness meditation (Fit-MF-Mind), and an active control that engaged in visual search tasks. In our prior report, we describe minimal differences between the active control condition and the three intervention groups, but noted profound individual differences in post-intervention performance on tests of G_f that were not explained by intervention assignment or pre-intervention ability [14]. Further, we noted in that report that even the individuals in the active control condition showed repeated-testing gains and high post-intervention performance on novel tests of G_f. Taken together, we speculated that all intervention activities, including the active control condition, may have engaged G_f function and that a follow-up study considering neural correlates of G_f as another source of individual differences was warranted.

In this manner, we aim to identify a possible subgroup of individuals who collectively demonstrate performance gains over the course of the intervention, as well as high performance on novel tests of G_f that were only administered at post-intervention. We further aim to identify neural correlates that best differentiate individuals who show greater gains following intervention, and possible differences between intervention groups. In this pursuit, we chose a set of regions *a priori* that, based upon the reviewed literature, are known correlates of cognitive control (anterior cingulate cortex, ACC; amygdala, Amy; middle frontal cortex, MFC), working memory (orbitofrontal cortex, OFC; caudate nucleus, Cd; cerebellum gray matter, Cb; insula cortex, Ins), and relational processing (hippocampus, Hc; posterior parahippocampal gyrus, PhG). We hypothesize that, when the sample is considered as a whole, a latent class of individuals who demonstrate gains and high G_f at post-intervention will be identified apart from a class of individuals who do not show gains and demonstrate low G_f at post-intervention. Second, we hypothesize that volumes of the prefrontal and parietal cortical regions will be the strongest predictors of individuals who show greater post-intervention gains in G_f, followed by regions that are correlated with cognitive control, working memory, and relational processing.

2. Material and methods

2.1. Participants

The study sample included 424 adults (age $M = 23.35$, $SD = 4.84$; 46% female; 50% Caucasian), who scored within the normal range on Shipley vocabulary ($M = 110.63$, $SD = 9.21$). The sample was recruited as part of a multi-modal intervention study that is described in our prior report [14] and here, in brief. Participants were recruited from the Champaign-Urbana, IL metro region. To be eligible for the study, participants were age 18–44 years; had at least a high school education; spoke English fluently; had normal or corrected-to-normal vision and hearing; no current or recent medications affecting the central nervous system or present a risk during aerobic exercise; no history of psychological, neurological, or endocrine disease, no concussion within the past two years, and no learning disorders; did not smoke more than 10 cigarettes per day; did not have a body mass index greater than 35; and responded negatively to all items on the physical activity readiness questionnaire revised [58]. Participants were randomly assigned to four groups that engaged in different intervention activities: aerobic exercise (Fit); aerobic exercise and cognitive training (Fit-MF); aerobic exercise, cognitive training and mindfulness training (Fit-MF-Mind); and an active control condition (Control). Refer to our prior report for a complete description of intervention activities and analysis of group differences [14]; this report considers individual differences in G_f independent of

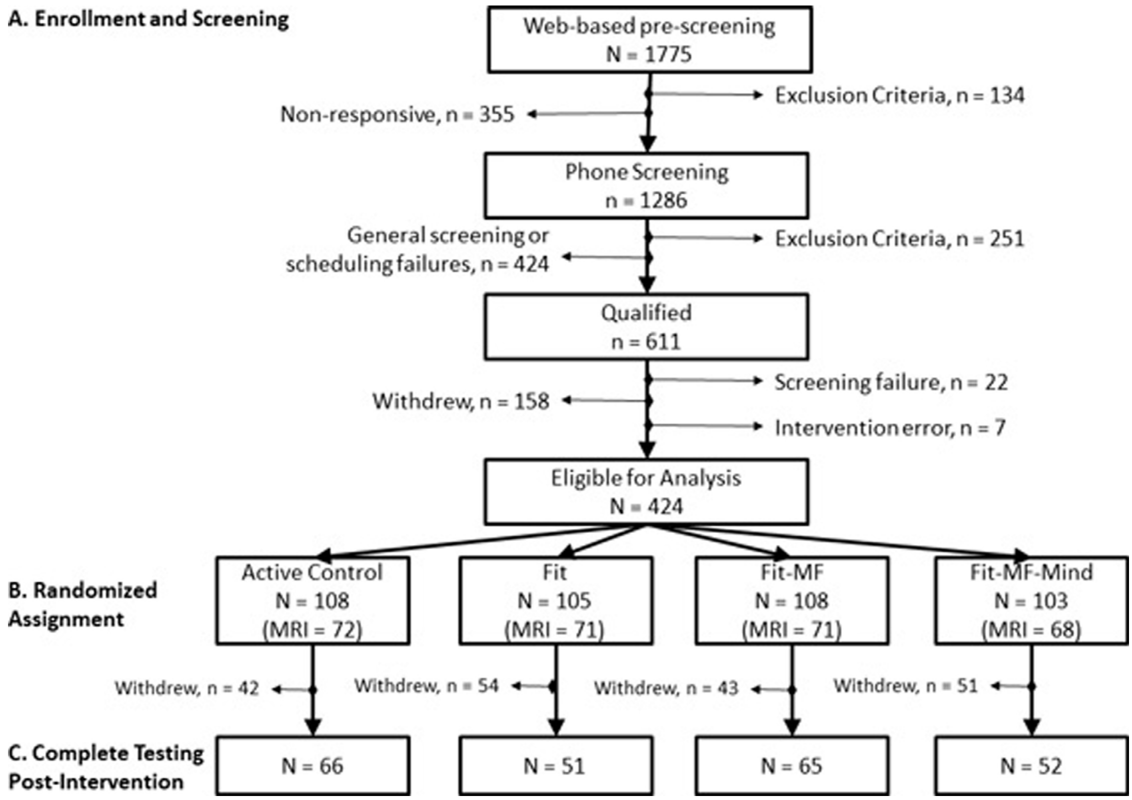

Fig. 1. Enrollment and attrition of the eligible sample in analyses. Participants were recruited and screened based on enrolment criteria. Based upon subject eligibility, responsiveness, and scheduling availability, 424 persons were enrolled and randomly assigned to one of four experimental conditions. A subset of individuals were chosen at random to undergo MRI at pre-intervention (n = 282). Intent-to-treat analyses included all eligible participants.

the intervention group assignments. All participants provided written informed consent in compliance with the university Institutional Review Board.

There were two sources of missing data. First, by study design, a subset of the eligible sample (N = 282; 67%) was selected at random to undergo MRI at pre-intervention. Second, of the eligible sample, 234 individuals returned for testing after the 16-week intervention (Fig. 1; n = 150 with complete data and MRI). In total, approximately 27% of data were missing at random (Little's χ^2 (30, N = 424) = 38.40, p = 0.14). The sample of individuals who had MRI were similar to those who did not in age (t (422) = −0.26, p = 0.80) and performance on Shipley vocabulary (t (422) = 1.72, p = 0.09), Figure Series (t (422) = 0.97, p = 0.97) and LSAT (t (422) = 1.88, p = 0.06) at pre-intervention, and were equivalently sampled across intervention conditions (χ^2 (3) = 0.13, p = 0.99). More women than men were missing an MRI scan (χ^2 (1) = 6.34, p = 0.01). The proportion of attrition was similar across intervention groups (χ^2 (3) = 2.21, p = 0.53) and was unrelated to demographic variables (all $p \geq 0.07$); individuals who returned performed similar to those who did not on Figure Series (t (422) = 0.09, p = 0.93) and LSAT at pre-intervention (t (422) = −0.50, p = 0.62). Based upon this analysis, missing data were handled via full information maximum likelihood (FIML) estimation, which leverages all available information in the covariance matrix to estimate latent effects without imputation [59,60]. The method is robust under the assumption of data missing at random [59,60], and inclusion of covariates of the missing data pattern (e.g., sex) in the model satisfies this assumption [61]. Use of all available data improves the validity of the model estimates and handling missing data via FIML is the current recommended practice for longitudinal studies [59]. Therefore, all analyses were completed with the total eligible sample of 424 participants.

2.2. *Assessment of fluid intelligence*

Fluid intelligence was assessed via standardized tests that were adapted to be administered on a computer with an online interface for data recording. Based on an independent sample, all tests had high internal consistency (Cronbach's $\alpha > 0.80$) and were administered with a time limit determined from response times in the 75th percentile, thus response speed was not considered as a confound to the assessment of G_f. Tests are reported in detail in our prior publication [14] and briefly, here.

2.2.1. *Repeated measures: LSAT and figure series*

Two tests of G_f were administered at pre- and post-intervention. The analytical reasoning subtest of the LSAT is a standardized achievement test used to determine admittance into law school and evaluates logical reasoning [62]. Figure Series is a canonical measure of G_f in which participants must choose the correct item missing in a series by deducing the rule governing the series [63]. For both tests, parallel forms were created and administered in a counterbalanced order at pre- and post-intervention. Total correct responses during the time limit on each test were included as indicators of G_f.

2.2.2. *Post-intervention G_f*

Four additional tests of G_f were administered only at post-intervention: Letter Series, Number Series, Matrix Reasoning, and Shipley Abstraction. Letter Series presents a string of letters to participants, who are required to choose the missing letter in the sequence [64]. In the Number Series Task, participants are presented a series of numbers and their task is to choose the missing number that completes the sequence [65]. The Matrix Reasoning Task also examined pattern completion, but within the context of abstract symbols presented in a matrix [63]. Finally, in the Shipley Abstraction Task, participants are presented a series of letters, numbers, or words, and their task is to choose the missing item that completes the sequence [66]. Tests were scored according to standard administration procedures.

2.2.3. MRI protocol and regional brain volumetry

MRI data were acquired with a Siemens Magnetom 3 T Trio scanner using a 32-channel head coil located in Beckman Institute Biomedical Imaging Center at the University of Illinois. A 3D high-resolution T1-weighted magnetization prepared gradient-echo (MPRAGE) sequence was collected with the following parameters: 0.9 mm^3 isotropic voxel; repetition time = 1900 ms; inversion time = 900 ms; echo time = 2.32 ms, with GRAPPA and an acceleration factor of 2.

Cortical reconstruction and volumetric segmentation was completed with the Freesurfer (v 6.0) image analysis suite, which is available for download online (http://surfer.nmr.mgh.harvard.edu/; last accessed 09/26/2018). Details of the procedures and software are described in prior publications [67–75]. Briefly, images were submitted to motion correction [74], removal of non-brain tissue using a hybrid watershed/surface deformation procedure [76], automated Talairach transformation, and segmentation of subcortical white matter and deep gray matter regions (including the hippocampus, amygdala and caudate nucleus) [69,71]. This was followed by tessellation of the gray-white matter boundary, automated topology correction [69,76], and surface deformation following intensity gradients to optimize boundary delineation between gray and white matter and cerebrospinal fluid [67,68,77], and regional parcellation of the cerebral cortex [72,78].

The automated procedure was evaluated by two trained raters for accurate gray-white matter separation, and was manually corrected with a procedure that was confirmed to be reliable: intra-class correlation coefficient (ICC(2); [79]) of at least 0.85. Subcortical gray matter segmentation was reviewed for complete parcellation and were not manually corrected. All regional brain volumes were corrected for individual intracranial volume via regression [80].

2.3. Statistical analysis

All analyses were completed in a structural equation modeling framework. Primary hypothesis testing was completed in a latent class analysis that we used here to identify subgroups of individuals who demonstrated above sample average G_f and greater responsiveness in the course of the intervention. The model construction for hypothesis testing was executed in two steps. In Step 1, two separate structural equation models were fit in order to estimate latent factor scores for performance on G_f tests and regional brain volumes. In Step 2, two separate latent class models were estimated to test hypotheses predicting performance subgroups at pre- and post-intervention.

2.3.1. Step 1: components of model construction

The model of G_f test performance was the same we reported in detail in our previous paper, and included appropriate model constraints for measurement invariance in repeated testing and between intervention groups [14] (Fig. 2, Step 1A). Figure Series and LSAT were administered at pre- and post-intervention, and these were used to identify latent constructs representing performance on each test at pre-intervention and latent change in performance. In addition, the four G_f measures that were only administered at post-intervention were used to identify a post-intervention fluid intelligence construct. All measures were converted to z-scores with the total sample average, and therefore negative values indicate below sample average performance and positive values indicate above sample average performance. Latent scores of pre-intervention Figure Series and LSAT, change in each test, and post-intervention G_f were extracted for further analysis in this report.

Also in Step 1, a second model of pre-intervention regional brain volumes was estimated (Fig. 2, Step 1B). Each region construct was identified by left and right volume measures, each with fixed loadings of 1 and freely-estimated measurement variance. This construction is conceptually similar to sum total volume but with the independent estimation of measurement error. The model included age, sex and Shipley vocabulary performance as covariates of regional brain volumes. In this model construction, missing values were handled via full information maximum likelihood estimation [59,60]. Latent factor scores of regional brain volumes were extracted for further analysis.

2.3.2. Step 2: hypothesis testing with latent class analysis

The hypotheses were tested in Step 2, which included estimating the number of latent classes that described performance subgroups at pre- (Fig. 2, Step 2A) and post-intervention (Step 2B), and predicting membership of the subgroups by regional brain volumes. The number of optimal classes was determined by relative change in model fit assessed with Akaike information criterion (AIC) and sample-size adjusted Bayesian information criterion (BIC), for which lower values indicated better fit [81]. As well as the Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMR) that compares change in the null hypothesis log-likelihood value between models of different numbers of classes, and statistical significance indicates that the additional class explains significantly more variability in the sample. This procedure was repeated in a step-wise manner until the optimal number of classes was determined.

Latent factor scores of regional brain volumes, age, and sex were included as covariates to predict latent class membership, as was intervention group. Brain region volumes were allowed to correlate, which provided an estimation of unique effects on latent class membership while accounting for repeated measurements. In model 2B of post-intervention performance, latent scores for pre-intervention Figure Series and LSAT were included as covariates [14] via regression, and therefore latent class membership was determined by individual variability in change in performance on repeated testing and on the novel tests at post-intervention, independent of pre-intervention G_f ability. Further, intervention group as a moderator of regional brain volume predicting performance subgroup was tested by including interaction terms, which if not significant were removed from the model. Significant prediction of latent class membership was evaluated at $p < 0.05$, and the unstandardized regression coefficients (b) are reported with the odds ratio and its 95% confidence intervals (OR 95% CI), which if not overlapping with 1.0 suggest differentiation between performance subgroups. In secondary analysis, planned comparisons between each intervention group and the active control condition were tested in predicting post-intervention performance subgroup membership, with Bonferroni correction of significance testing ($\alpha' = 0.02$).

3. Results

3.1. Identification of latent classes in performance at pre-intervention

In examining individual differences in pre-intervention performance on Figure Series and LSAT, two classes of individuals were identified. The model of two latent classes (AIC = 1591.13, BIC = 1597.27) fit significantly better than assuming a homogeneous sample (VLMR = 11.52, $p = 0.004$), and there was no evidence in support of three classes (AIC = 1566.12, BIC = 1574.89; VLMR = 122.51, $p = 0.61$). Class 1 identified 259 individuals (61%) and was defined by below sample average performance on Figure Series (latent mean = −0.30, $p < 0.001$) and LSAT (latent mean = −0.31, $p < 0.001$). Class 2 identified 165 individuals (39%), defined by above sample average performance on Figure Series (latent mean = 0.42, $p < 0.001$) and LSAT (latent mean = 0.45, $p < 0.001$). Based upon this observed pattern, we refer to Class 1 as "Lower Pre-Intervention G_f" and Class 2 as "Higher Pre-Intervention G_f", but we note that all participants were of normal intelligence.

3.1.1. Neural correlates of fluid intelligence performance groups at pre-intervention

Volumes of several brain regions were included as covariates to

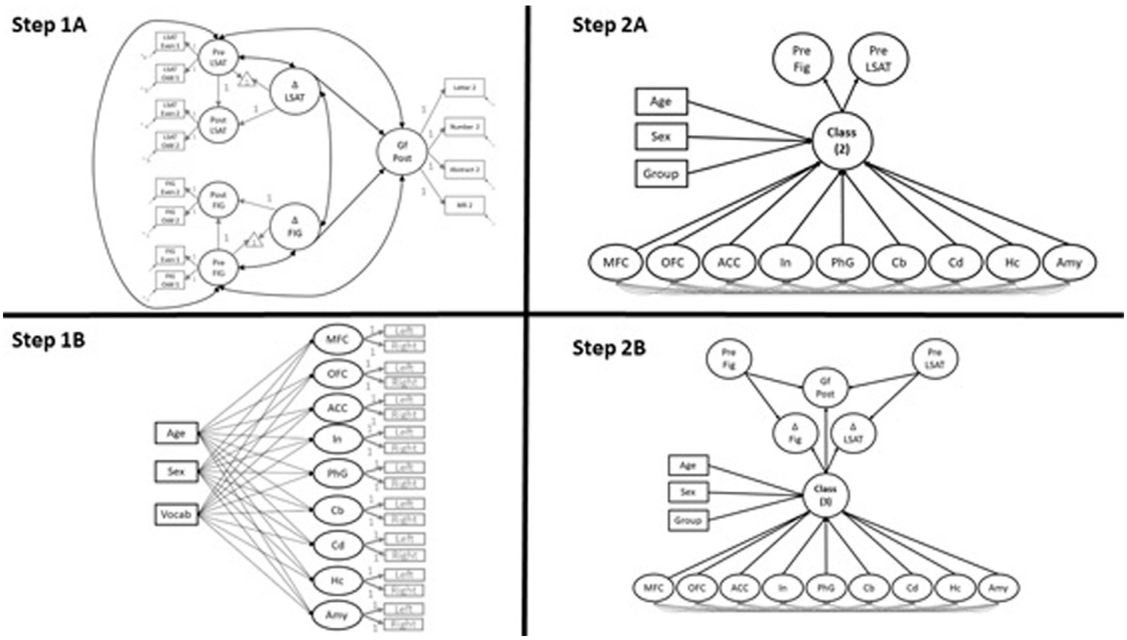

Fig. 2. An illustration of the structural equation modelling analysis. The analysis was completed in two steps. In step 1, two separate models (A and B) were estimated in order to extract latent factor scores to be used in the latent class analysis in step 2. The model depicted in Step 1A was published in our prior report [14]. The model in Step 1B was estimated to create latent estimates of regional brain volumes, and not illustrated, all brain region constructs were allowed to correlate. The latent factor scores that were estimated in Step 1A and 1B were extracted and used in the latent class analysis in Step 2. In the latent class model in Step 2, all covariates were allowed to correlate (illustrated with double-headed, curved arrow) and performance subgroups were predicted at pre-intervention (Step 2A) and at post-intervention (Step 2B).

predict latent class membership. As expected, the volumes of the brain regions were inter-correlated, ranging $r = 0.64$, $p < 0.001$ (ACC correlated with OFC) to $r = 0.04$, $p = 0.63$ (Ins correlated with PhG). Although not all regional volumes were significantly correlated with another, the degree of commonality in part reflects repeated measures taken of an individual, and we accounted for this by allowing the regional measures to correlate while predicting latent class membership.

A number of brain regions significantly differentiated the Higher Pre-Intervention G_f group from the lower G_f performance group (Table 1). Notably, the pattern of results was mixed and in rank order,

Table 1
Correlates of latent fluid intelligence performance classes at pre-intervention.

	Higher vs. Lower Pre-Intervention G_f			
Correlate	b	p-value	OR	OR 95% CI
Age	0.03	0.569	1.03	0.94/1.12
Sex (male)	**3.48**	**0.001**	**32.33**	**6.12/170.72**
Group	−0.01	0.949	0.99	0.69/1.40
ACC	**3.30**	**< 0.001**	**27.11**	**6.25/117.69**
MFC	**−0.23**	**< 0.001**	**0.80**	**0.73/0.88**
OFC	−0.26	0.059	0.77	0.62/0.97
Ins	**−1.33**	**< 0.001**	**0.27**	**0.15/0.48**
PhG	**−2.26**	**< 0.001**	**0.10**	**0.04/0.28**
Cb	**0.06**	**< 0.001**	**1.06**	**1.03/1.08**
Cd	**1.09**	**< 0.001**	**2.97**	**1.74/5.06**
Hc	**1.52**	**< 0.001**	**4.58**	**2.39/8.76**
Amy	**−0.84**	**0.045**	**0.43**	**0.22/0.86**

Note: Prediction of membership in the "Higher Pre-Intervention G_f" group (39% of the sample) versus the "Lower Pre-Intervention G_f" group (61% of the sample)—e.g., positive b-weight values and OR > 1.00 indicate larger regional brain volumes predict comparatively greater likelihood of classification into the higher performance group. Significant coefficients are bolded ($p < 0.05$). ACC—anterior cingulate cortex; MFC—middle frontal cortex; OFC—orbitofrontal cortex; Ins—insula cortex; PhG—posterior parahippocampal gyrus; Cb—cerebellum gray matter; Cd—caudate nucleus; Hc—hippocampus; Amy—amygdala.

larger ACC, Hc and Cd volumes were the strongest predictors, followed by smaller PhG volume. Repeating the analysis in the sample with complete data ($n = 150$) produced a similar pattern of results: 43% of the sample was classified as Higher Pre-Intervention G_f, which was significantly predicted by all brain regions (all $p \leq 0.03$), except for Amy ($p = 0.22$).

3.1.2. Equivalence of intervention groups prior to the intervention

In our prior report, we found that the randomly-assigned intervention groups were equivalent in pre-intervention G_f performance [14]. In the present analysis, intervention group assignment did not differentiate between the Lower and Higher Pre-Intervention G_f groups identified in the latent class analysis: $b = -0.01$, $p = 0.95$, OR = 0.99.

3.1.3. Identification of latent classes in performance at post-intervention

Higher pre-intervention performance on Figure Series ($b = -0.42$, $p < 0.001$) and LSAT ($b = -0.08$, $p < 0.001$) predicted lesser repeated-testing gains on each test, and both were associated with higher post-intervention G_f scores ($b = 0.53$ and 0.40, $p < 0.001$, respectively). Controlling for individual differences at pre-intervention, three classes described variability in performance at post-intervention. The model of three latent classes (AIC = 2396.52, BIC = 2415.80) fit significantly better than two classes (VLMR = 41.05, $p = 0.03$), and there was no evidence in support of four classes (AIC = 2293.98, BIC = 2318.52; VLMR = 120.24, $p = 0.39$).

Class 1 identified 39 individuals (9%) and was defined by significant declines in performance on repeated testing and below sample average performance on novel G_f measures at post-intervention (Table 2). Class 2 identified 314 individuals (74%), defined by significant gains in Figure Series, but significant decline in LSAT and below sample average performance on post-intervention G_f tests. Class 3 identified 71 individuals (17%) and was defined by significant gain in Figure Series that was greater than in Class 2 (see 95% CI in Table 2), stable performance on LSAT and above sample average performance on post-intervention G_f tests. By this collection of assessments, Class 3 described a group of individuals who showed the greatest gains following the

Table 2
Latent class identification by performance on fluid intelligence assessments over the course of the intervention.

Measure	Class 1 "Low Response" Mean	95% CI	p-value	R^2	Class 2 "Moderate Response" Mean	95% CI	p-value	R^2	Class 3 "High Response" Mean	95% CI	p-value	R^2
Δ Figure Series	−0.74	−0.89/−0.59	< 0.001	0.55	0.40	0.38/0.43	< 0.001	0.52	1.15	1.03/1.26	< 0.001	0.56
Δ LSAT	−0.17	−0.23/−0.11	< 0.001	0.07	−0.07	−0.09/−0.05	< 0.001	0.05	0.03	−0.01/0.07	0.52	0.06
Post-Intervention G_f	−0.46	−0.50/−0.41	< 0.001	0.97	−0.04	−0.06/−0.03	< 0.001	0.96	0.24	0.21/0.28	< 0.001	0.97

Note: Group means are reported with 95% confidence intervals, significance testing, and R^2 as the proportion of variance explained in each performance index per group within the model that included correlates. Figure Series and LSAT tests were repeated at pre- and post-intervention, and post-intervention G_f reflected performance on four novel tests that were not repeated. Performance on tests of G_f at pre- and post-intervention were normed to the sample average, and therefore negative values indicate performance below the sample average. Positive change scores indicate repeated-testing gains, and negative scores, decline over the course of the intervention. Pre-intervention performance was controlled as a covariate in the models. Based upon the observed pattern of individual differences in performance, we termed Class 1 as "Low Intervention Response" (9% of the sample), Class 2 as "Moderate Intervention Response" (74% of the sample), and Class 3 as "High Intervention Response" (17% of the sample).

intervention period and further analysis predicted classification of this group. We refer to Class 3 as the "High Intervention Response", Class 2 as the "Moderate Intervention Response" and Class 1 as the "Low Intervention Response" groups.

3.1.4. Neural correlates of fluid intelligence performance groups following an intervention

A number of regions significantly differentiated the High Intervention Response group from the Low Response group (Table 3). In rank order, larger ACC ($b = 1.47, p = 0.001$), Cd ($b = 0.78, p < 0.001$), and Hc ($b = 0.74$, $p < 0.001$), and smaller PhG ($b = -1.22$, $p = 0.002$), Ins ($b = -0.74$, $p < 0.001$), and MFC ($b = -0.12$, $p = 0.006$) volumes predicted greater likelihood of assignment to the High Intervention Response group relative to the lowest performance group. When differentiated from the remaining majority of the sample that demonstrated moderate response to the intervention, only three regions significantly predicted the High Intervention Response group: smaller PhG ($b = -0.50$, $p = 0.048$) was the strongest predictor, followed by larger Cd volume ($b = 0.30, p = 0.018$) and larger Cb volume ($b = 0.02$, $p = 0.038$). Taken together, smaller PhG volume was a consistently strong predictor of high intervention response apart from either the Low (OR = 0.29, OR 95% CI: 0.15/0.57) or Moderate Intervention Response groups (OR = 0.61, OR 95% CI: 0.40/0.92), followed by larger Cd volume (Table 2).

Repeating the analysis in the sample with complete data produced similar results: 38% of the sample was classified as High Intervention Response, followed by 55% Moderate Response and 7% Low Response. Volumes of ACC (Low OR = 2.46, $p = 0.48$; Moderate OR = 1.40, $p = 0.57$), Cd (Low OR = 1.83, $p = 0.03$; Moderate OR = 1.18, $p = 0.43$), PhG (Low OR = 0.87, $p = 0.86$; Moderate OR = 0.79, $p = 0.63$), Ins (Low OR = 0.66, $p = 0.20$; Moderate OR = 0.92, $p = 0.75$), and Amy (Low OR = 0.78, $p = 0.85$; Moderate OR = 0.48, $p = 0.05$) were the strongest predictors, although the significance testing was likely less sensitive.

3.1.5. Intervention groups did not differ in classification of performance

The main effect of intervention group did not significantly differentiate the High Intervention Response Group (vs. Low $b = -0.07$, $p = 0.74$; vs. Moderate $b = -0.02$, $p = 0.85$) and it did not moderate the relation between regional brain volume and performance classification (all interaction terms $p > 0.07$). In secondary analysis, planned comparisons against the active control condition further identified no evidence of differential effects by intervention type: Fit ($b = -0.33$ and -0.24, p's $\geq$ 0.44 vs. Low and Moderate, respectively), Fit-MF ($b = 0.64$ and 0.44, p's $\geq$ 0.08, respectively), and Fit-MF-Mind ($b = -0.07$ and -0.12, p's $\geq$ 0.67, respectively). Therefore, independent of prescribed intervention and pre-intervention ability, individuals with smaller PhG and larger Cd volumes were more likely to show greater improvement in G_f test performance.

4. Discussion

Here we take a novel approach of studying individual differences in the course of a RCT via latent class analysis to explore neural correlates of G_f. In a large sample of neurologically healthy, younger adults we observed three classes of individuals who showed low, moderate, and high responses following the intervention. Those in the High Intervention Response group demonstrated the largest repeated-testing

Table 3
Correlates of latent fluid intelligence performance classes at post-intervention.

Correlate	High Response vs. Low Response b	p-value	OR	OR 95% CI	High Response vs. Moderate Response b	p-value	OR	OR 95% CI
Age	−0.13	0.054	0.88	0.79/0.98	0.02	0.688	1.02	0.94/1.11
Sex	**1.48**	**0.030**	**4.40**	**1.43/13.50**	−0.38	0.463	0.68	0.29/1.61
Group	−0.07	0.754	0.94	0.66/1.32	−0.02	0.851	0.98	0.79/1.21
ACC	**1.47**	**0.001**	**4.33**	**2.09/9.00**	0.40	0.256	1.49	0.84/2.64
MFC	**−0.12**	**0.006**	**0.89**	**0.83/0.96**	−0.02	0.442	0.98	0.93/1.03
OFC	−0.12	0.114	0.88	0.78/1.01	0.01	0.935	1.01	0.91/1.11
Ins	**−0.74**	**< 0.001**	**0.48**	**0.35/0.65**	−0.25	0.063	0.78	0.63/0.97
PhG	**−1.22**	**0.002**	**0.29**	**0.15/0.57**	**−0.50**	**0.048**	**0.61**	**0.40/0.92**
Cb	0.02	0.116	1.02	1.00/1.03	**0.02**	**0.038**	**1.02**	**1.00/1.03**
Cd	**0.78**	**<0.001**	**2.18**	**1.61/2.94**	**0.30**	**0.018**	**1.36**	**1.10/1.67**
Hc	**0.74**	**< 0.001**	**2.10**	**1.52/2.92**	0.29	0.078	1.33	1.02/1.73
Amy	−0.65	0.064	0.52	0.29/0.93	−0.42	0.101	0.66	0.43/1.00

Note: Based upon the observed performance that defined each class, we describe those who showed a high intervention response in contrast to the low and moderate intervention response. All effects are calculated with regard to the high response group: e.g., positive b-weight values and OR > 1.00 indicate larger regional brain volumes predict comparatively greater likelihood of classification into the high response group. Significant coefficients are bolded ($p < 0.05$). ACC—anterior cingulate cortex; MFC—middle frontal cortex; OFC—orbitofrontal cortex; Ins—insula cortex; PhG—posterior parahippocampal gyrus; Cb—cerebellum gray matter; Cd—caudate nucleus; Hc—hippocampus; Amy—amygdala.

gains in visuospatial reasoning, stability (instead of decline) in analytic reasoning performance, and above sample average performance on novel tests of G_f administered at post-intervention. Test performance indicated average intelligence in this sample, and therefore individual differences that we have identified reflect natural variability that is representative of the normal population. Volumes of several brain regions predicted the subgroup with higher G_f at pre-intervention and distinguished individuals with high response to the intervention, including regions implicated in cognitive control, working memory, and relational processing.

Individuals at post-intervention fell within three performance subgroups, and one group (17%) demonstrated the highest response following the intervention—largest gains in visuospatial reasoning, relative stability (opposed to decline) in analytical reasoning scores, and above sample average performance on tests of G_f that were only administered at post-intervention (controlling for pre-intervention ability). This group was distinct from the majority of the sample (74%) who showed modest gains only in visuospatial reasoning and otherwise poor performance, and a minority (9%) who experienced decline in repeated-testing and below sample average performance at post-intervention. The pattern of effects that include gains in performance on Figure Series but stability, and even decline, on LSAT is consistent with our prior report and we speculated that the psychometric properties of repeated-testing differ between the two tests [14]. Further, the latent class analysis accounted for a large proportion of variability in change in Figure Series performance and in novel post-intervention tests, but a small proportion of variance of change in LSAT performance. Following our previous report of individual differences in performance that were unexplained by intervention group [14], we hypothesized that there may be a subgroup of individuals across prescribed interventions who experienced differential improvement in G_f. Here, we found evidence in support of this hypothesis.

However, the type of multi-modal intervention activity—aerobic exercise, cognitive training, mindfulness meditation, or visual search and change detection (control)—did not differentiate between performance subgroups. The observation of greater individual variability than the magnitude of intervention group differences is in line with the mixed results that are common in the cognitive training and intervention literature [82]. This is further consistent with our previous report of minimal intervention-related effects, in which the combined fitness-cognitive training condition produced isolated repeated-testing gains in visuospatial reasoning greater than the active control condition [14]. We had identified that individuals even in the active control condition showed gains in performance following the intervention period, and large variability within the condition that additionally incorporated mindfulness meditation [14]. Each intervention activity was selected for the possibility of directly (i.e., cognitive training) or indirectly (i.e., aerobic exercise and mindfulness meditation) acting upon the cognitive and neural constituents of G_f. Moreover, the visual search task in the active control condition may have promoted attention control and processing speed [83] that later aided performance on timed tests of G_f. Indeed, here we found no single intervention activity that increased the likelihood of an individual belonging to the high performance subgroup at post-intervention. Instead, individual differences in regional brain volumes were the stronger predictor of performance at both pre- and post-intervention.

Volumes of regions within prefrontal and insular cortices differentiated between the lowest and highest intervention response subgroups. Specifically, smaller MFC and Ins predicted greater likelihood of showing G_f performance gains, opposed to deficits, following the intervention period. Similarly, smaller volumes of these regions also predicted higher pre-intervention G_f (39% of the sample) as compared to lower performance. These cortical regions follow a protracted developmental trajectory in which cortical thinning and shrinkage occur throughout adolescence and into the second decade before entering a period of stability [84], which is adaptive and supports normal cognitive development. As the study sample here was selected to be younger (average age 23.35 years), smaller volumes of these cortical regions predicting higher G_f function is in line with an expected developmental trajectory for adaptive cognitive function. In contrast, ACC follows a different non-linear developmental trajectory and develops earlier than the prefrontal cortex [84]. Therefore, it is plausible that larger ACC volumes in young adulthood may be adaptive for better cognitive control. In line with this, larger ACC volume predicted high G_f intervention response. The identification of these cortical regions as correlates of G_f performance subgroups is consistent with the theories of fluid intelligence engaging the fronto-parietal network [20] and many reports from MRI and patient lesion studies [18–24]. However, we do not find evidence of OFC volume significantly and uniquely differentiating performance groups at either pre- or post-intervention. In a longitudinal study of healthy aging, larger OFC volume predicted higher fluid intelligence in adults [49], and it is plausible that this region is a stronger correlate of function in the context of neurodegeneration than in the healthy, young adult sample of the present study.

Prefrontal and parietal cortical regions are canonical correlates of fluid cognitive ability and their connectivity with other brain regions galvanizes the theory that other cognitive abilities buttress G_f. For example, the functional connections between the prefrontal cortex and hippocampus may provide information stored in relational memory to be used in pattern detection, rule extraction and adaptive reasoning [47]. Although neural correlates of relational memory are not typically studied in the context of fluid intelligence, there is some initial evidence in support of this notion. Hc volume correlates with individual differences in fluid intelligence among musicians [48], and shrinkage predicts age-related declines in general intelligence across the adult lifespan [49,50]. Here, we identified that larger Hc predicted greater likelihood of classification to the high performing group at pre-intervention and in differentiating the highest intervention response group from the lowest. As participants were naïve to the tasks by design, and repeated tests employed parallel forms, the contribution of Hc cannot be ascribed to recollection of the specific test *per se*. Instead, relational memory function may have bolstered performance on tests of adaptive reasoning and fluid ability, and individuals with larger Hc volumes may have been better equipped to take advantage of this.

Neural correlates of fluid ability differentiated between performance groups, yet the strongest predictors were regions implicated in working memory and relational processes. Smaller PhG volume was among the strongest predictors of pre-intervention performance group, as well as in differentiating those who showed the greatest response following intervention. The comparison between the extreme lowest and highest performance groups is informative. Yet, most intriguing is that PhG, Cd and Cb were the only brain regions examined that also differentiated the highest response group from the remaining majority of the sample, who demonstrated lesser but significant gains in visuospatial reasoning, decline in analytical reasoning, and lower post-intervention performance on novel tests. Therefore, contrary to our hypothesis, neural correlates of relational processes and working memory, which are associated with fluid ability, were the strongest predictors of G_f over the course of the intervention.

The close relationship between working memory and performance on tests of G_f has been widely and thoroughly reported [18,19,30,31]. Cognitive control directly modifies working memory functions and mediates the effect of working memory capacity on G_f test performance [33–36]. The prefrontal-parietal network regions that were identified as significant predictors of performance subgroup membership—middle frontal, insula, and anterior cingulate cortices—are not specific to fluid intelligence and reasoning functions, and are implicated in a broader working memory network that includes the cerebellum and caudate nucleus [85]. Further, considering the test content in the present study, many of the assessments relied upon visuospatial cognitive ability and this constituted the majority of variance explained in the latent performance groups. Visuospatial cognitive ability correlates with parahippocampal cortical structure and function that integrates information and facilitates contextual judgment [51,86]. Relational processes, in conjunction with working memory, are integral to G_f ability [87,88]. Of the regions examined, PhG may have presented as a strong performance predictor for its function, and for its integrative projections between the hippocampus, prefrontal and parietal cortical regions in a network for information and experiences, integrated over time and space [56,57,89]. Therefore, the evidence we report here is congruent with a model of fluid intelligence that is dependent upon executive functions [19] and buttressed by relational processes [47,87,88]. Two of the three multi-modal intervention conditions included adaptive cognitive training that targeted working memory functions [14]. The reported evidence suggests that individuals who possessed a more developed and robust neural architecture may have been better equipped to benefit from the multi-modal cognitive, exercise and mindfulness intervention that we employed here.

However, prescribed intervention alone did not differentiate between performance subgroups or modify the strength of neural correlates. No differential effects by group may indicate that all combinations of activities produced comparable effects, or that no activity affected G_f. The intervention activities encompassed cognitive training, physical exercise and mindfulness activities, all of which correlate with development and maintenance of cognitive ability across the lifespan to varying degree [90–94]. Yet, acute interventions aimed to promote cognitive ability, including G_f, appear to not be "one size fits all" and individuals vary widely in response to the intervention. Here we demonstrate that multi-modal intervention does not provide differential gains but the forces that shaped the developing brain may be of greater importance. This observation begs for future research to adopt methods to specifically evaluate the magnitude and source of individual variability, even in the context of RCT [95].

In application to formal education, this observation bolsters the continued trend to create diverse lesson plans and flexible assignments that accommodate an individual student's interests and abilities [96]. Education and learning experiences shape neural cognitive developmental trajectories across the lifespan; and as shown here, pre-intervention ability and regional brain volumes modified an individual's performance following the acute intervention. Neural cognitive development in early life may not only determine adulthood cognitive ability but also modify efficacy of interventions aimed to improve it.

The evidence we report should be interpreted with consideration of study strengths and limitations. First, we use a novel approach to study individual differences in the course of a randomized, multi-modal intervention. However, we evaluated covariates of latent G_f class membership selected *a priori*, and due to the size of the sample, we cannot include volumetry of all brain regions. Therefore, the analysis provides insight into multiple neural cognitive systems that predict individual variability in fluid intelligence functions following a multi-modal intervention. Yet, the analysis is not representative of the entire brain. Second, we selected volumes of gray matter regions as a proxy indicator of brain structure. Additional white matter regions, as well as regional functional activation and metabolic markers, are expected correlates of fluid intelligence function that we did not examine here. Future studies may endeavor to integrate multi-modal neuroimaging data to provide a thorough account of the neural architecture of fluid intelligence and sources of individual differences in ability.

Third, only two tests of G_f were repeated in the course of study, and performance on Figure Series and LSAT did not identify a common factor and showed different patterns of change with repeated testing [14]. This constrained the specification of the latent modeling and the estimates of repeated testing gains are specific to each test, opposed to latent ability reflective of performance on multiple tests. This was better addressed with the collection of tests that were administered only at post-intervention that identified a common factor; nonetheless, this model limitation should be considered when interpreting the reported results.

Fourth, approximately a third of data were missing either by design or due to study attrition. Data were found to satisfy the statistical assumption of missing at random. However, the motivations and causes for study attrition may not be random and correlate with cognitive ability, attention and motivation [97], which are relevant to performance on tests of G_f. The current report bases primary hypothesis testing in the complete sample with intent-to-treat analyses and handled missing data via FIML estimation, which is the current recommended approach to longitudinal studies with missing data, including intervention designs [59,60,98,99]. Yet, we can only assess possible correlates of attrition based upon the available data and our limited assessment should be considered when interpreting the evidence. Analyses were repeated in the sample with complete data and a similar pattern of effects was identified, although some significance testing was likely less sensitive in the smaller sample. Further, inclusion of the entire available sample improves the validity of the estimates. In this manner, possible nonrandom causes for attrition that may correlate with performance are represented (albeit imperfectly) in the estimates, opposed to narrowly defining the model by those individuals who chose to complete the prescribed intervention.

Fifth, the sample includes healthy, younger adults who lived in a Midwest United States city, and most of whom were enrolled in a four-year, public university. Performance on the tests of G_f and vocabulary fell within the range of normal intelligence, and therefore we describe individual differences representative of the typical population. However, the generalizability of this sample is specific to its demographic characteristics and it does not represent the diversity of the broader population.

Finally, the analysis of latent classes only characterized the pattern of individual differences in G_f performance and its neural correlates. The analysis cannot evaluate hypotheses of causal relationships between brain structures and functions, and should be considered as a description of the observed data.

5. Conclusions

We present a novel approach to study individual differences in the course of a randomized, multi-modal intervention study via latent class analysis. Three latent classes described performance at post-intervention, while controlling for pre-intervention ability: one subgroup (17%) demonstrated the greatest gains in visuospatial reasoning, stability (and not decline) in analytical reasoning and above sample-average performance on tests that were only administered at post-intervention. This performance subgroup was differentiated from the lowest performance subgroup (9% of the sample), and the remaining sample majority, who notably experienced lesser, but significant, gains in visuospatial ability and otherwise performed poorly. Volumes of several brain regions that are implicated in cognitive control, working memory, and relational processes differentiated between the highest and lowest intervention response groups. However, regions important to relational processing and working memory, especially the posterior parahippocampal gyrus and caudate nucleus, emerged as the strongest predictors of the high performance subgroup. We find evidence supporting a theoretical account of fluid intelligence ability that is dependent upon executive function and buttressed by relational processes. To acquire a better understanding of the neural architecture of fluid intelligence, future large-scale studies that adopt multi-modal neuroimaging, multiple complementary cognitive assessments, and statistical methods for the assessment of individual differences are necessary.

Declaration of Competing Interest

A.M.D., B.P.S., C.H.H., A.F.K., N.J.C., and A.K.B. have no financial or personal conflicts of interests to disclose.

Acknowledgments

We wish to extend special thanks to Dr. Patricia Jones, Dr. Chris Zwilling, Mr. Nikolai Sherepa, Mr. Evan Anderson, Ms. Courtney Allen, and the INSIGHT research team for their assistance in study administration and data collection. The research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via Contract 2014-13121700004 to University of Illinois at Urbana-Champaign (PI: Barbey). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes not- withstanding any copyright annotation thereon. Ana M. Daugherty was supported by a Beckman Institute Postdoctoral Fellowship at the University of Illinois at Urbana-Champaign, with funding provided by the Arnold and Mabel Beckman Foundation.

References

[1] L.S. Gottfredson, Why g matters: the complexity of everyday life, Intelligence 24 (1) (1997) 79–132.
[2] N.R. Kuncel, S.A. Hezlett, Standardized tests predict graduate students' success, Science 315 (2007) 1080–1081, https://doi.org/10.1126/science.1136618.
[3] J.E. Hunter, Cognitive ability, cognitive aptitudes, job knowledge, and job performance, J. Vocat. Behav. 29 (1986) 340–362.
[4] J.F. Salgado, N. Anderson, S. Moscoso, C. Bertua, F. de Fruyt, J.P. Rolland, A meta-analytic study of general mental ability validity for different occupations in the European community, J. Appl. Psychol. 88 (2003) 1068–1081, https://doi.org/10.1037/0021-9010.88.6.1068.
[5] P. Hagmann-von Arx, J.T. Gygi, R. Weidmann, A. Grob, Testing relations of crystallized and fluid intelligence and the incremental predictive validity of conscientiousness and its facets on career success in a small sample of German and Swiss workers, Front. Psychol. (2016) 7, https://doi.org/10.3389/fpsyg.2016.00500.
[6] A.K. Barbey, Network neuroscience theory of human intelligence, Trends Cognit. Sci. 22 (2018) 8–20, https://doi.org/10.1016/j.tics.2017.10.001.
[7] B.J. Carroll, F. Cassidy, D. Naftolowitz, N.E. Tatham, W.H. Wilson, A. Iranmanesh, P.Y. Liu, J.D. Veldhuis, Pathophysiology of hypercortisolism in depression, Acta Psychiatr. Scand. 115 (2007) 90–103, https://doi.org/10.1111/j.1600-0447.2007.00967.x.
[8] R.B. Cattell, Theory of fluid and crystallized intelligence: a critical experiment, J. Educ. Psychol. 54 (1963) 1–22, https://doi.org/10.1037/h0046743.
[9] A.S. Kaufman, J.L. Horn, Age changes on tests of fluid and crystallized ability for women and men on the kaufman adolescent and adult intelligence test (KAIT) at ages 17–94 years, Arch. Clin. Neuropsychol. 11 (1996) 97–121, https://doi.org/10.1016/0887-6177(95)00003-8.
[10] W.-.T. Chooi, L.A. Thompson, Working memory training does not improve intelligence in healthy young adults, Intelligence 40 (2012) 531–542, https://doi.org/10.1016/j.intell.2012.07.004.
[11] S.M. Jaeggi, B. Studer-Luethi, M. Buschkuehl, Y. Su, J. Jonides, W.J. Perrig, Intelligence the relationship between n-back performance and matrix reasoning — implications for training and transfer, Intelligence 28 (2010) 625–635, https://doi.org/10.1016/j.intell.2010.09.001.
[12] T.S. Redick, Z. Shipstead, T.L. Harrison, K.L. Hicks, D.E. Fried, D.Z. Hambrick, M.J. Kane, R.W. Engle, No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study, J. Exp. Psychol. Gen. 142 (2013) 359–379, https://doi.org/10.1037/a0029082.
[13] T.W. Thompson, M.L. Waskom, K.-L.A. Garel, C. Cardenas-Iniguez, G.O. Reynolds, R. Winter, P. Chang, K. Pollard, N. Lala, G.A. Alvarez, J.D.E. Gabrieli, Failure of working memory training to enhance cognition or intelligence, PLoS One (2013) 8, https://doi.org/10.1371/journal.pone.0063614.
[14] A.M. Daugherty, C. Zwilling, E.J. Paul, N. Sherepa, C. Allen, A.F. Kramer, C.H. Hillman, N.J. Cohen, A.K. Barbey, Multi-modal fitness and cognitive training to enhance fluid intelligence, Intelligence 66 (2018) 32–43, https://doi.org/10.1016/j.intell.2017.11.001.
[15] R. Hammer, E.J. Paul, C.H. Hillman, A.F. Kramer, N.J. Cohen, A.K. Barbey, Individual differences in analogical reasoning revealed by multivariate task-based functional brain imaging, Neuroimage (2018), https://doi.org/10.1016/j.neuroimage.2018.09.011.
[16] T. Talukdar, F.J. Román, J.T. Operskalski, C.E. Zwilling, A.K. Barbey, Individual differences in decision making competence revealed by multivariate fMRI, Hum. Brain Mapp. 39 (2018) 2664–2672, https://doi.org/10.1002/hbm.24032.
[17] C. Spearman, THE abilities of man, Science 68 (1928) 38, https://doi.org/10.1126/science.68.1750.38-a.
[18] A.K. Barbey, R. Colom, E.J. Paul, J. Grafman, Architecture of fluid intelligence and working memory revealed by lesion mapping, Brain Struct. Funct. 219 (2014) 485–494, https://doi.org/10.1007/s00429-013-0512-z.
[19] A.K. Barbey, R. Colom, J. Solomon, F. Krueger, C. Forbes, J. Grafman, An integrative architecture for general intelligence and executive function revealed by lesion mapping, Brain J. Neurol. 135 (2012) 1154–1164, https://doi.org/10.1093/brain/aws021.
[20] R.E. Jung, R.J. Haier, The parieto-frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence, Behav. Brain Sci. 30 (2007) 135–154, https://doi.org/10.1017/S0140525x07001185.
[21] F. Preusse, E. Van Der Meer, G. Deshpande, F. Krueger, I. Wartenburger, Fluid intelligence allows flexible recruitment of the parieto-frontal network in analogical reasoning, Front. Hum. Neurosci. (2011) 5, https://doi.org/10.3389/fnhum.2011.00022.
[22] F. Preusse, E. van der Meer, D. Ullwer, M. Brucks, F. Krueger, I. Wartenburger, Long-term characteristics of analogical processing in high-school students with high fluid intelligence: an fMRI study, ZDM 42 (2010) 635–647, https://doi.org/10.1007/s11858-010-0259-4.
[23] I. Wartenburger, H.R. Heekeren, F. Preusse, J. Kramer, E. van der Meer, Cerebral correlates of analogical processing and their modulation by training, Neuroimage 48 (2009) 291–302, https://doi.org/10.1016/j.neuroimage.2009.06.025.
[24] C.M. Wharton, J. Grafman, S.S. Flitman, E.K. Hansen, J. Brauner, A. Marks, M. Honda, Toward neuroanatomical models of analogy: a positron emission tomography study of analogical mapping, Cognit. Psychol. 40 (2000) 173–197, https://doi.org/10.1006/cogp.1999.0726.
[25] M.W. Cole, T. Ito, T.S. Braver, Lateral prefrontal cortex contributes to fluid intelligence through multinetwork connectivity, Brain Connect. 5 (2015) 497–504, https://doi.org/10.1089/brain.2015.0357.
[26] D. Tranel, E. Vianna, K. Manzel, H. Damasio, T. Grabowski, Neuroanatomical correlates of the benton facial recognition test and judgment of line orientation test, J. Clin. Exp. Neuropsychol. 31 (2009) 219–233, https://doi.org/10.1080/13803390802317542.
[27] W.C. Halstead, Brain and Intelligence; a Quantitative Study of the Frontal Lobes, University of Chicago Press, Chicago, IL, US, 1947.
[28] R. Colom, M. Burgaleta, F.J. Román, S. Karama, J. Alvarez-Linera, F.J. Abad, K. Martínez, M.Á. Quiroga, R.J. Haier, Neuroanatomic overlap between intelligence and cognitive factors: morphometry methods provide support for the key role of the frontal lobes, Neuroimage 72 (2013) 143–152, https://doi.org/10.1016/j.neuroimage.2013.01.032.
[29] A.K. Barbey, M. Koenigs, J. Grafman, Orbitofrontal contributions to human working memory, Cereb. Cortex 21 (2011) (1991) 789–795, https://doi.org/10.1093/cercor/bhq153.
[30] R.W. Engle, S.W. Tuholski, J.E. Laughlin, A.R. Conway, Working memory, short-term memory, and general fluid intelligence: a latent-variable approach, J. Exp. Psychol. Gen. 128 (1999) 309–331.
[31] M.J. Kane, D.Z. Hambrick, S.W. Tuholski, O. Wilhelm, T.W. Payne, R.W. Engle, The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. J. Exp. Psychol. Gen. 133 (2004) 189–217.
[32] C. Rottschy, R. Langner, I. Dogan, K. Reetz, A.R. Laird, J.B. Schulz, P.T. Fox, S.B. Eickhoff, Modelling neural correlates of working memory: a coordinate-based meta-analysis, Neuroimage 60 (2012) 830–846, https://doi.org/10.1016/j.neuroimage.2011.11.050.
[33] R.W. Engle, M.J. Kane, Executive attention, working memory capacity, and a two-factor theory of cognitive control, Psychol. Learn. Motiv. (2003) 145–199, https://doi.org/10.1016/S0079-7421(03)44005-X Elsevier.
[34] J.C. McVay, M.J. Kane, Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention, J. Exp. Psychol. Gen. 141 (2012) 302–320, https://doi.org/10.1037/a0025250.
[35] N. Unsworth, G.J. Spillers, Working memory capacity: attention control, secondary memory, or both? A direct test of the dual-component model, J. Mem. Lang. 62 (2010) 392–406, https://doi.org/10.1016/j.jml.2010.02.001.
[36] N. Unsworth, K. Fukuda, E. Awh, E.K. Vogel, Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval, Cognit. Psychol. 71 (2014) 1–26, https://doi.org/10.1016/j.cogpsych.2014.01.003.
[37] P.M. Greenwood, R. Parasuraman, The mechanisms of far transfer from cognitive training: review and hypothesis, Neuropsychology 30 (2016) 742–755, https://doi.org/10.1037/neu0000235.
[38] J.R. Gray, C.F. Chabris, T.S. Braver, Neural mechanisms of general fluid intelligence, Nat. Neurosci. 6 (2003) 316–322, https://doi.org/10.1038/nn1014.
[39] K. Nashiro, M. Sakaki, M. Mather, Age differences in brain activity during emotion processing: reflections of age-related decline or increased emotion regulation? Gerontology 58 (2012) 156–163, https://doi.org/10.1159/000328465.
[40] S.J. Vine, L.J. Moore, M.R. Wilson, An integrative framework of stress, attention, and visuomotor performance, Front. Psychol. (2016) 7, https://doi.org/10.3389/fpsyg.2016.01671.
[41] J.B. Banks, M.S. Welhaf, A. Srour, The protective effects of brief mindfulness meditation training, Conscious. Cognit. 33 (2015) 277–285, https://doi.org/10.1016/j.concog.2015.01.016.
[42] L.A. Brown, E.M. Forman, J.D. Herbert, K.L. Hoffman, E.K. Yuen, E.M. Goetter, A randomized controlled trial of acceptance-based behavior therapy and cognitive therapy for test anxiety: a pilot study, Behav. Modif. 35 (2011) 31–53, https://doi.org/10.1177/0145445510390930.
[43] M.D. Mrazek, M.S. Franklin, D.T. Phillips, B. Baird, J.W. Schooler, Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering, Psychol. Sci. 24 (2013) 776–781, https://doi.org/10.1177/0956797612459659.
[44] C. Noone, B. Bunting, M.J. Hogan, Does mindfulness enhance critical thinking?

evidence for the mediating effects of executive functioning in the relationship between mindfulness and critical thinking, Front. Psychol. 6 (2015) 2043, https://doi.org/10.3389/fpsyg.2015.02043.

[45] Y.-.Y. Tang, B.K. Hölzel, M.I. Posner, The neuroscience of mindfulness meditation, Nat. Rev. Neurosci. 16 (2015) 213–225, https://doi.org/10.1038/nrn3916.

[46] P.B. Sharp, B.P. Sutton, E.J. Paul, N. Sherepa, C.H. Hillman, N.J. Cohen, A.F. Kramer, R.S. Prakash, W. Heller, E.H. Telzer, A.K. Barbey, Mindfulness training induces structural connectome changes in insula networks, Sci. Rep. 8 (2018) 7929, https://doi.org/10.1038/s41598-018-26268-w.

[47] J.X. Wang, N.J. Cohen, J.L. Voss, Covert rapid action-memory simulation (CRAMS): a hypothesis of hippocampal–prefrontal interactions for adaptive behavior, Neurobiol. Learn. Mem. 117 (2015) 22–33, https://doi.org/10.1016/j.nlm.2014.04.003.

[48] M.S. Oechslin, C. Descloux, A. Croquelois, J. Chanal, D.V.D. Ville, F. Lazeyras, C.E. James, Hippocampal volume predicts fluid intelligence in musically trained people, Hippocampus 23 (2013) 552–558, https://doi.org/10.1002/hipo.22120.

[49] N. Raz, U. Lindenberger, P. Ghisletta, K.M. Rodrigue, K.M. Kennedy, J.D. Acker, Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors, Cereb. Cortex 18 (2008) (1991) 718–726, https://doi.org/10.1093/cercor/bhm108.

[50] A. Reuben, A.M. Brickman, J. Muraskin, J. Steffener, Y. Stern, Hippocampal atrophy relates to fluid intelligence decline in the elderly, J. Int. Neuropsychol. Soc. JINS 17 (2011) 56–61, https://doi.org/10.1017/S135561771000127X.

[51] E.M. Aminoff, K. Kveraga, M. Bar, The role of the parahippocampal cortex in cognition, Trends Cognit. Sci. 17 (2013) 379–390, https://doi.org/10.1016/j.tics.2013.06.009.

[52] H.E.H. Pol, H.G. Schnack, D. Posthuma, R.C.W. Mandl, W.F. Baaré, C. van Oel, N.E. van Haren, D.L. Collins, A.C. Evans, K. Amunts, U. Bürgel, K. Zilles, E. de Geus, D.I. Boomsma, R.S. Kahn, Genetic contributions to human brain morphology and intelligence, J. Neurosci. 26 (2006) 10235–10242, https://doi.org/10.1523/JNEUROSCI.1312-06.2006.

[53] M. Azab, S.M. Stark, C.E.L. Stark, Contributions of human hippocampal subfields to spatial and temporal pattern separation, Hippocampus 24 (2014) 293–302, https://doi.org/10.1002/hipo.22223.

[54] E.T. Rolls, Pattern separation, completion, and categorisation in the hippocampus and neocortex, Neurobiol. Learn. Mem. 129 (2016) 4–28, https://doi.org/10.1016/j.nlm.2015.07.008.

[55] M.A. Yassa, C.E.L. Stark, Pattern separation in the hippocampus, Trends Neurosci. 34 (2011) 515–525, https://doi.org/10.1016/j.tins.2011.06.006.

[56] N.J. Cohen, H. Eichenbaum, Memory, Amnesia, and the Hippocampal System, The MIT Press, Cambridge, MA, US, 1993.

[57] H. Eichenbaum, N.J. Cohen, From Conditioning to Conscious Recollection: Memory Systems of the Brain, Oxford University Press, New York, NY, US, 2001.

[58] S. Thomas, J. Reading, R.J. Shephard, Revision of the physical activity readiness questionnaire (PAR-Q), Can. J. Sport Sci. J. Can. Sci. Sport 17 (1992) 338–345.

[59] R. Larsen, Missing data imputation versus full information maximum likelihood with second-level dependencies, Struct. Equ. Model. Multidiscip. J. 18 (2011) 649–662, https://doi.org/10.1080/10705511.2011.607721.

[60] B. Muthen, D. Kaplan, M. Hollis, On structural equation modeling with data that are not missing completely at random, Psychometrika (1987) 431–462.

[61] C.K. Enders, Applied Missing Data Analysis, Guilford Press, New York, 2010.

[62] L. Roussos, L. Norton, LSAT Item-Type Validity Study: Law School Admission Council Technical Report 98-01, (1998).

[63] R.B. Cattell, A.K.S. Cattell, Measuring Intelligence With the Culture Fair Tests, Institute for Personality and Ability Testing, Institute for Personality and Ability Testing, Champaign, Ill., 1960.

[64] K.W. Schaie, T.G. Thurstone, L.L. Thurstone, Schaie-Thurstone Adult Mental Abilities Test, Consulting Psychologists Press, Palo Alto, Calif., 1985.

[65] K. McGrew, F. Schrank, R. Woodcock, Woodcock-Johnson III Normative update: Technical manual, Riverside Publishing, 2007.

[66] W. Shipley, C. Gruber, T. Martin, A. Klein, Shipley-2, Western Psychological Services, 2009.

[67] A.M. Dale, B. Fischl, M.I. Sereno, Cortical surface-based analysis: I. segmentation and surface reconstruction, Neuroimage 9 (1999) 179–194, https://doi.org/10.1006/nimg.1998.0395.

[68] A.M. Dale, M.I. Sereno, Improved localizadon of cortical activity by combining EEG and meg with MRI cortical surface reconstruction: a linear approach, J. Cognit. Neurosci. 5 (1993) 162–176, https://doi.org/10.1162/jocn.1993.5.2.162.

[69] B. Fischl, A. Liu, A.M. Dale, Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex, IEEE Trans. Med. Imaging 20 (2001) 70–80, https://doi.org/10.1109/42.906426.

[70] B. Fischl, M. Sereno, A.M. Dale, Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system, Neuroimage 9 (1999) 195–207, https://doi.org/10.1006/nimg.1998.0396.

[71] B. Fischl, D.H. Salat, A.J.W. van der Kouwe, N. Makris, F. Ségonne, B.T. Quinn, A.M. Dale, Sequence-independent segmentation of magnetic resonance images, Neuroimage (23 Suppl 1) (2004) S69–S84, https://doi.org/10.1016/j.neuroimage.2004.07.016.

[72] B. Fischl, A. van der Kouwe, C. Destrieux, E. Halgren, F. Ségonne, D.H. Salat, E. Busa, L.J. Seidman, J. Goldstein, D. Kennedy, V. Caviness, N. Makris, B. Rosen, A.M. Dale, Automatically parcellating the human cerebral cortex, Cereb. Cortex 14 (2004) (1991) 11–22.

[73] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. Macfall, B. Fischl, A. Dale, Reliability in multisite structural mri studies: effects of gradient non-linearity correction on phantom and human data, Neuroimage 30 (2006) 436–443, https://doi.org/10.1016/j.neuroimage.2005.09.046.

[74] M. Reuter, H.D. Rosas, B. Fischl, Highly accurate inverse consistent registration: a robust approach, Neuroimage 53 (2010) 1181–1196, https://doi.org/10.1016/j.neuroimage.2010.07.020.

[75] M. Reuter, N.J. Schmansky, H.D. Rosas, B. Fischl, Within-subject template estimation for unbiased longitudinal image analysis, Neuroimage 61 (2012) 1402–1418, https://doi.org/10.1016/j.neuroimage.2012.02.084.

[76] F. Ségonne, A.M. Dale, E. Busa, M. Glessner, D. Salat, H.K. Hahn, B. Fischl, A hybrid approach to the skull stripping problem in MRI, Neuroimage 22 (2004) 1060–1075, https://doi.org/10.1016/j.neuroimage.2004.03.032.

[77] B. Fischl, A.M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 11050–11055, https://doi.org/10.1073/pnas.200033797.

[78] R.S. Desikan, F. Ségonne, B. Fischl, B.T. Quinn, B.C. Dickerson, D. Blacker, R.L. Buckner, A.M. Dale, R.P. Maguire, B.T. Hyman, M.S. Albert, R.J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, Neuroimage 31 (2006) 968–980, https://doi.org/10.1016/j.neuroimage.2006.01.021.

[79] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol. Bull. 86 (1979) 420–428.

[80] C.R. Jack, C. Twomey, A.R. Zinsmeister, F.W. Sharbrough, R.C. Petersen, G.D. Cascino, Anterior temporal lobes and hippocampal formationL normaive volumetric measurements from mr images in young adults, Radiology 172 (1989) 549–554.

[81] Q. Chen, W. Luo, G.J. Palardy, R. Glaman, A. McEnturff, The efficacy of common fit indices for enumerating classes in growth mixture models when nested data structure is ignored: a Monte Carlo study, Sage Open 7 (2017) 215824401770045, , https://doi.org/10.1177/2158244017700459.

[82] J. Au, E. Sheehan, N. Tsai, G.J. Duncan, M. Buschkuehl, S.M. Jaeggi, Improving fluid intelligence with training on working memory: a meta-analysis, Psychon. Bull. Rev. 22 (2015) 366–377, https://doi.org/10.3758/s13423-014-0699-x.

[83] K. Clark, L.G. Appelbaum, B. van den Berg, S.R. Mitroff, M.G. Woldorff, Improvement in visual search with practice: mapping learning-related changes in neurocognitive stages of processing, J. Neurosci. 35 (2015) 5351–5359, https://doi.org/10.1523/JNEUROSCI.1152-14.2015.

[84] P. Shaw, N.J. Kabani, J.P. Lerch, K. Eckstrand, R. Lenroot, N. Gogtay, D. Greenstein, L. Clasen, A. Evans, J.L. Rapoport, J.N. Giedd, S.P. Wise, Neurodevelopmental trajectories of the human cerebral cortex, J. Neurosci. Off. J. Soc. Neurosci. 28 (2008) 3586–3594, https://doi.org/10.1523/JNEUROSCI.5309-07.2008.

[85] J. Eriksson, E.K. Vogel, A. Lansner, F. Bergström, L. Nyberg, Neurocognitive architecture of working memory, Neuron. 88 (2015) 33–46, https://doi.org/10.1016/j.neuron.2015.09.020.

[86] O. Baumann, J.B. Mattingley, Functional organization of the parahippocampal cortex: dissociable roles for context representations and the perception of visual scenes, J. Neurosci. 36 (2016) 2536–2542, https://doi.org/10.1523/JNEUROSCI.3368-15.2016.

[87] K. Oberauer, Design for a working memory, Psychol. Learn. Motiv. 51 (2009) 45–100, https://doi.org/10.1016/S0079-7421(09)51002-X Elsevier Academic Press.

[88] K. Oberauer, H.-.M. Süß, O. Wilhelm, N. Sander, Individual differences in working memory capacity and reasoning ability, Variation in Working Memory, Oxford University Press, 2007, pp. 49–75.

[89] A. Konkel, N.J. Cohen, Relational memory and the hippocampus: representations and methods, Front. Neurosci. 3 (2009) 166–174, https://doi.org/10.3389/neuro.01.023.2009.

[90] E.S. Sharp, C.A. Reynolds, N.L. Pedersen, M. Gatz, Cognitive engagement and cognitive aging: is openness protective? Psychol. Aging 25 (2010) 60–73, https://doi.org/10.1037/a0018748.

[91] E.A.L. Stine-Morrow, B.R. Payne, B.W. Roberts, A.F. Kramer, D.G. Morrow, L. Payne, P.L. Hill, J.J. Jackson, X. Gao, S.R. Noh, M.C. Janke, J.M. Parisi, Training versus engagement as paths to cognitive enrichment with aging, Psychol. Aging 29 (2014) 891–906, https://doi.org/10.1037/a0038244.

[92] L. Corno, E.B. Mandinach, The role of cognitive engagement in classroom learning and motivation, Educ. Psychol. 18 (1983) 88–108, https://doi.org/10.1080/00461528309529266.

[93] D.B. Bellinger, M.S. DeCaro, P.A.S. Ralston, Mindfulness, anxiety, and high-stakes mathematics performance in the laboratory and classroom, Conscious. Cognit. 37 (2015) 123–132, https://doi.org/10.1016/j.concog.2015.09.001.

[94] C.H. Hillman, K.I. Erickson, A.F. Kramer, Be smart, exercise your heart: exercise effects on brain and cognition, Nat. Rev. Neurosci. 9 (2008) 58–65, https://doi.org/10.1038/nrn2298.

[95] T. Smoleń, J. Jastrzebski, E. Estrada, A. Chuderski, Most evidence for the compensation account of cognitive training is unreliable, Mem. Cognit. 46 (2018) 1315–1330, https://doi.org/10.3758/s13421-018-0839-z.

[96] D. Dockterman, Insights from 200+ years of personalized learning, NPJ Sci. Learn. 3 (2018) 15, https://doi.org/10.1038/s41539-018-0033-x.

[97] D.C. Arbiv, N. Meiran, Performance on the antisaccade task predicts dropout from cognitive training, Intelligence 49 (2015) 25–31, https://doi.org/10.1016/j.intell.2014.11.009.

[98] T.D. Little, Longitudinal Structural Equation Modeling, Guilford Press, New York, NY, US, 2013.

[99] T. Raykov, Analysis of longitudinal studies with missing data using covariance structure modeling with full-information maximum likelihood, Struct. Equ. Model. Multidiscip. J. 12 (2005) 493–505, https://doi.org/10.1207/s15328007sem1203_8.