Behavioral/Cognitive

# Neural Contributions to Reduced Fluid Intelligence across the Adult Lifespan

Daniel J. Mitchell,[1] Alexa L. S. Mousley,[1] Meredith A. Shafto,[2] Cam-CAN,[3*] and John Duncan[1]

[1]Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, CB2 7EF, United Kingdom, [2]Pomona College, Claremont, California 91711, and [3]Cambridge Centre for Ageing and Neuroscience, University of Cambridge and MRC Cognition and Brain Sciences Unit, CB2 7EF, United Kingdom

Fluid intelligence, the ability to solve novel, complex problems, declines steeply during healthy human aging. Using fMRI, fluid intelligence has been repeatedly associated with activation of a frontoparietal brain network, and impairment following focal damage to these regions suggests that fluid intelligence depends on their integrity. It is therefore possible that age-related functional differences in frontoparietal activity contribute to the reduction in fluid intelligence. This paper reports on analysis of the Cambridge Center for Ageing and Neuroscience data, a large, population-based cohort of healthy males and females across the adult lifespan. The data support a model in which age-related differences in fluid intelligence are partially mediated by the responsiveness of frontoparietal regions to novel problem-solving. We first replicate a prior finding of such mediation using an independent sample. We then precisely localize the mediating brain regions, and show that mediation is specifically associated with voxels most activated by cognitive demand, but not with voxels suppressed by cognitive demand. We quantify the robustness of this result to potential unmodeled confounders, and estimate the causal direction of the effects. Finally, exploratory analyses suggest that neural mediation of age-related differences in fluid intelligence is moderated by the variety of regular physical activities, more reliably than by their frequency or duration. An additional moderating role of the variety of nonphysical activities emerged when controlling for head motion. A better understanding of the mechanisms that link healthy aging with lower fluid intelligence may suggest strategies for mitigating such decline.

*Key words:* aging; exercise; fMRI; frontoparietal; intelligence; multiple-demand

---

### Significance Statement

Global populations are living longer, driving urgency to understand age-related cognitive declines. Fluid intelligence is of prime importance because it reflects performance across many domains, and declines especially steeply during healthy aging. Despite consensus that fluid intelligence is associated with particular frontoparietal brain regions, little research has investigated suggestions that under-responsiveness of these regions mediates age-related decline. We replicate a recent demonstration of such mediation, showing specific association with brain regions most activated by cognitive demand, and robustness to moderate confounding by unmodeled variables. By showing that this mediation model is moderated by the variety of regular physical activities, more reliably than by their frequency or duration, we identify a potential modifiable lifestyle factor that may help promote successful aging.

---

# Introduction

Aging affects many cognitive abilities, but the drop in fluid intelligence, the ability to solve novel, complex problems (Cattell, 1963; Carpenter et al., 1990; Duncan et al., 2017), is particularly steep (Horn and Cattell, 1967; Kievit et al., 2014; Samu et al., 2017). Moreover, fluid intelligence may be especially important for determining aging trajectories through contributions to ability across cognitive domains (Parkin and Java, 1999; Salthouse et al., 2003), performance in daily life (Diehl et al., 1995), and even health outcomes (Gottfredson and Deary, 2004). Understanding fluid intelligence decline is increasingly urgent as lifespans increase (Vaupel, 2010).

Extensive research associates fluid intelligence with a frontoparietal brain network (Jung and Haier, 2007; Duncan, 2010; Santarnecchi et al., 2017a). These regions, including the precentral sulcus, middle frontal gyrus, intraparietal sulcus, anterior insula, and anterior cingulate cortex, have been termed the multiple-demand network (MDN) because of their activation during many cognitively demanding tasks (Duncan and Owen, 2000; Naghavi and Nyberg, 2005; Duncan, 2010), including tests of fluid intelligence (Prabhakaran et al., 1997; Duncan et al., 2000). Individual differences in fluid intelligence correlate with MDN activity (Gray et al., 2003; Lee et al., 2006; Tschentscher et al., 2017; Assem et al., 2020b) and with its connectivity to other brain networks (Cole et al., 2012). The impact of focal lesions (Glascher et al., 2010; Woolgar et al., 2010; Barbey et al., 2014; Woolgar et al., 2018; Smith et al., 2022) and transient transcranial stimulation (Momi et al., 2020) suggests these regions' causal role in supporting fluid intelligence. We therefore test the possibility that functional differences in MDN activation mediate fluid intelligence decline during healthy aging (Phillips and Della Sala, 1998). Confirming this would help to understand the mechanism of age-related decline, and suggest potential targets for mitigation with interventions that might impact on the putative causal pathway from age to fluid intelligence via neural responsiveness.

Many studies have considered relations between brain activation and cognitive performance in older adults (Dennis and Cabeza, 2008; Eyler et al., 2011; Grady, 2012), although few have probed the specific three-way relationship between differences in age, brain activation, and fluid intelligence. In the context of broader questions on differential age effects across task domains, a recent study (Samu et al., 2017) reported results consistent with frontoparietal activity mediating age differences in performance during a fluid intelligence task. Activity of the default mode network (DMN), typically deactivated during attentionally demanding tasks (Buckner et al., 2008), was not found to mediate performance decline in the task, despite changing with age in tasks that showed behavioral decline.

Here, we have five main aims. First, we replicate the finding of frontoparietal mediation of age differences in fluid intelligence (Samu et al., 2017), using an independent, nonoverlapping sample of participants from the population-based, healthy, adult lifespan cohort (Cambridge Center for Ageing and Neuroscience [Cam-CAN]) (Shafto et al., 2014) used by Samu et al. (2017). As the dependent variable, we replace concurrent task performance with a previously acquired, standardized fluid intelligence measure (Cattell and Cattell, 1973), avoiding external factors (e.g., arousal) that could comodulate simultaneous measures of brain activity and behavior. Second, after combining the nonoverlapping sample with that from Samu et al. (2017), we test whether mediation is specifically associated with voxels most responsive to cognitive demands, or also with voxels suppressed by cognitive demand. Third, we assess the robustness of the mediation result to possible unmodeled covariates. Fourth, since the mediation analysis cannot itself determine causality, we estimate causal directions under additional assumptions of an acyclic model with non-Gaussian errors. Finally, in exploratory analyses, we consider whether these relationships depend on potentially modifiable lifestyle factors. Specifically, increasing research proposes that physical exercise confers resilience to cognitive aging, although the nature and mechanism of this benefit remain unclear (Colcombe and Kramer, 2003; Smith et al., 2010; Liu-Ambrose et al., 2018). We therefore examine whether the mediation model is moderated by questionnaire measures that distinguish the variety, frequency, and duration of regular physical recreations.

# Materials and Methods

*Experimental design and statistical analyses.* The original experimental design for the Cam-CAN project (www.cam-can.com) is described in Shafto et al. (2014). Full details of the within- and between-subject variables, statistical tests, and software used in the current project are described in the following sections. Data can be requested after registration via the Cam-CAN dataset inventory (https://camcan-archive.mrc-cbu.cam.ac.uk/dataaccess/). Analysis code is available via the Open Science Framework (https://osf.io/xgw56/). The analyses were not preregistered.

*Participants.* Participants reported here are a subset of 252 participants from the population-based healthy adult cohort recruited for the Cam-CAN (for full details of the sample and exclusion criteria, see Shafto et al., 2014). The initial Cam-CAN data collection consisted of a background interview (Stage 1), detailed cognitive testing and core measures of brain structure and function (Stage 2), and targeted functional neuroimaging studies (Stage 3). The current study reports results from the behavioral test of fluid intelligence (Stage 2), the fMRI session using a similar fluid intelligence task (Stage 3), and self-reported recreational activities (from a questionnaire distributed in Stage 1). Of the participants recruited, 252 (133 female) completed the fMRI task, of whom all had completed the prior fluid intelligence test. Stage 3 testing occurred between 0.3 and 3.4 (mean 1.4) years following Stage 2. For analyses including age, we used the age midway between the two tests, for which ages ranged from 20.5 to 90.3 years (mean 55.1 years; approximately equal numbers per decile).

The first analysis, seeking to replicate the mediation observed by Samu et al. (2017), used a subset of 154 participants not included in the prior study. The remaining analyses used all 252 participants, except for the final analyses of moderated mediation, for which questionnaire data on physical recreation were missing from 13 participants and data on nonphysical activities were missing from 15 participants.

Participants gave written, informed consent, and the study was conducted in accordance with ethical approval obtained from the Cambridgeshire 2 (now East of England–Cambridge Central) Research Ethics Committee.

*Fluid intelligence measure.* Fluid intelligence was assessed using Scale 2, Form A of Cattell's Culture Fair Test (Cattell and Cattell, 1973), according to the standard protocol. This consists of four, nonverbal, multiple-choice, pencil-and-paper subtests of abstract reasoning (series completion, odd-one-out, matrix completion, and topological judgment) each introduced with examples and then completed under timed conditions, but with participants not informed of the precise time limits. When the total number of correct problems was converted to its standardized fluid intelligence score (IQ) using the conversion table in the manual (which is age-adjusted only below age 14) (Cattell and Cattell, 1973), its variance was found to decrease significantly with age (White-Wooldridge test, $\chi^2_{(2)} = 11.6$, $p = 0.003$). We therefore constructed a latent IQ variable from the first principle component across the subtests, similar to Kievit et al. (2014), which had homoscedastic residuals when predicted from age. This variable was standardized to have the same sample mean and SD as the normed scores based on the manual.

*Lifestyle activities measures.* Recent years have seen increasing consensus that physical exercise can be beneficial for neurocognitive health, including in older adults (Colcombe and Kramer, 2003; Smith et al., 2010; Liu-Ambrose et al., 2018). Most studies have examined a single measure of exercise, and so cannot distinguish which aspects of increased exercise might be most beneficial; however, some recent reports suggest that the duration of exercise may be less important than its intensity (Angevaren et al., 2007; Brown et al., 2012) or variety (Angevaren et al., 2007). In the Cam-CAN study, we had access to questionnaire measures that distinguished the variety, frequency, and duration of self-reported physical recreational activities, allowing us to investigate which of these different aspects of exercise might have the greatest impact on age-related decline in fluid intelligence.

Measures of recent physical recreational activities were taken from a self-completion questionnaire, completed during Stage 1 (within 2 years of Stage 2), based on the recreation section of the EPIC-EPAQ2 questionnaire (Wareham et al., 2002), which was derived, in turn, from the Minnesota Leisure Time Activity questionnaire (Richardson et al., 1994). Questions included the approximate frequency (on a 7 point scale) and duration (in hours and minutes) of each of 35 recreational and DIY activities in which people had participated over the last year (e.g., cycling, mowing the lawn, dancing, golf). We considered just "regular" activities, which we defined as those occurring at least monthly. "Variety" of regular activities was measured as the number of different activities, "frequency" per regular activity was measured as the mean number of episodes in a year, and "duration per episode" was also measured as the mean across regular activities; "total duration" was calculated by multiplying the frequency and duration of each activity and summing over regular activities. In this way, we sought to address the question: if someone were to devote a fixed amount of time to extra physical activity, might it be better to perform their current activities for longer, to perform their current activities more often, or to engage in a greater range of activities?
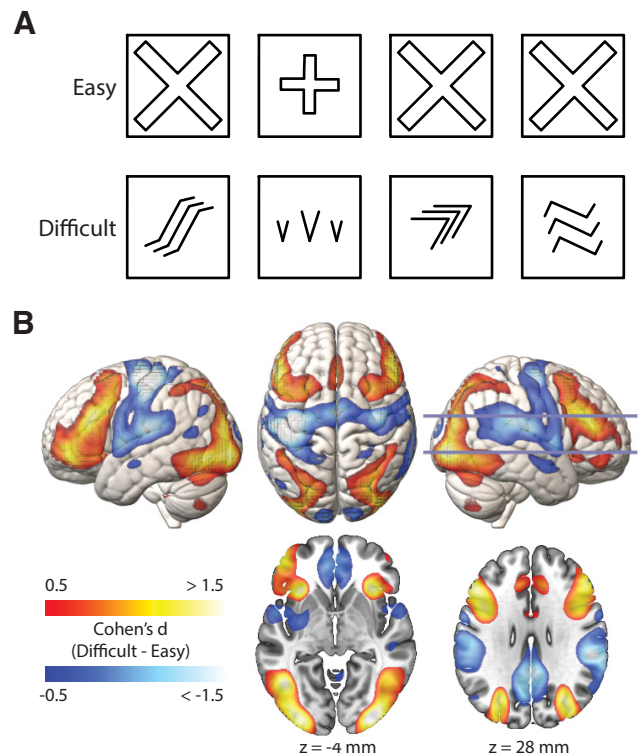
To assess whether the results for the variety of physical activities generalized to more intellectual activities, we derived a similar measure for the variety of recent "nonphysical" activities. For these activities, duration was never reported, and for most of them frequency was not reported, so variety was the only measure. The variety of recent nonphysical activities was quantified by summing the number of mental and social activities reported across two sources: (1) The "recent activities" portion of the self-completion questionnaire (based on elements of the Lifetime of Experiences Questionnaire) (Valenzuela and Sachdev, 2007) provided 38 items, including mental activities in a typical week (e.g., reading, art, crossword puzzles), events or entertainment in the last 2 months (e.g., cinema, pub, concert), usual means of acquiring information (e.g., TV, newspapers, Internet), and kinds of materials read on a regular basis (e.g., newspaper, novels, magazines); and (2) from a home interview at Stage 1, we included self-reports of 22 types of social interactions (e.g., phone friends, email friends, attend social clubs). As for the physical activity measures, for all questions where frequencies were reported, we only counted regular nonphysical activities (i.e., those that occurred at least monthly).

*MRI acquisition.* MRI was performed on a 3 Tesla Siemens TIM Trio System, using a 32 channel head coil.

A high-resolution 3D T1-weighted structural image was acquired using an MPRAGE sequence, with the following parameters: $TR = 2250$ ms; $TE = 2.99$ ms; $TI = 900$ ms; flip angle = 9 degrees; $FOV = 256 \times 240 \times 192$ mm; voxel size = 1 mm isotropic; GRAPPA acceleration factor = 2.

fMRI used a T2*-weighted gradient-echo EPI sequence with the following parameters: 32 axial slices (acquired in descending order); slice thickness of 3 mm, with an interslice gap of 25%; $TR = 2$ s; $TE = 30$ ms; flip angle = 78 degrees; $FOV = 192 \times 192 \times 120$ mm; voxel-size = $3 \times 3 \times 3.75$ mm.

*fMRI task.* During fMRI, participants performed a nonverbal reasoning task (see Fig. 1A) that has been previously shown to activate the frontoparietal MDN (Duncan et al., 2000; Woolgar et al., 2013) and is based on the odd-one-out subtest of Cattell's Culture Fair test (Cattell and Cattell, 1973). The task consisted of a series of problems in which participants were presented with a horizontal display of four panels and were instructed to select the panel that differed in some way from all of the



**Figure 1.** fMRI task, and response to difficult versus easy problem-solving. *A*, Examples of easy and difficult problems. In both examples, the second item is the correct answer. *B*, Regions significantly more active in the difficult than the easy condition (warm colors) and vice versa (cool colors), shown overlaid on surface renderings (top) and horizontal slices (bottom). N = 252. Slice positions are labeled in MNI coordinates and marked on the right hemisphere rendering. Activations are shown to a depth of 15 mm. Color scales indicate Cohen's *d*, thresholded at $|d| > 0.5$. (All effects greater than $|d| > 0.13$ are significant when controlling the FDR at $< 0.05$.)

others. The horizontal extent of each display was approximately 12 degrees of visual angle. The task used a block design, with alternating blocks of easy and difficult problems. In the easy blocks, three panels were identical and the fourth was clearly different, rendering each decision trivial; in the difficult blocks, the four panels in each problem differed in many ways, requiring the identification of abstract patterns to select the odd-one-out. Participants completed four easy blocks and four difficult blocks, each preceded by a 3 s cue indicating whether the upcoming problems would be "Easy" or "Hard." Each problem remained on the screen until the participant responded, whereupon the next problem was presented after a 500 ms blank interval. Problems were presented in fixed order, drawn from a pool of 320 easy and 25 difficult problems. If a participant completed all problems of a given difficulty, problems were recycled from the beginning. (Across participants, the mean percentage of repeated problems was 0.2% for easy problems, and 12.5% for difficult problems.) Each block automatically ended after 30 s. Participants were encouraged to puzzle over each problem for as long as necessary, only responding when confident of the correct answer. Thus, the number of trials per block varied, while the time spent on each type of problem (easy and difficult) was held constant. The task was presented using E-Prime (Psychology Software Tools) and stimuli were back-projected onto a screen that was viewed through a mirror mounted on the head coil. Responses were made using a button box.

Before entering the scanner, participants were shown examples of the types of problems that they would encounter, and they practised selecting the odd-one-out until both they and the experimenter were happy that they understood the instructions.

*fMRI preprocessing.* Data were analyzed using "automatic analysis" software (Cusack et al., 2015) in MATLAB (The MathWorks), which called relevant routines from SPM12 (Wellcome Department of Imaging Neuroscience, London).

Each participant's structural volume was segmented into probabilistic maps of six tissue classes. Gray and white matter maps of all Cam-CAN Stage 2 participants whose structural volumes passed quality control (272, including 20 participants not reported here) were nonlinearly aligned using diffeomorphic registration (DARTEL) (Ashburner, 2007) to create a group template volume, which was then normalized to the MNI template via an affine transformation.

Functional volumes were rigidly realigned to correct for motion, and slice-time corrected. They were then coregistered to the structural volume, and normalized to template space using the combined transformations (native to group template, and group template to MNI template) derived from the structural volume. Functional volumes were spatially smoothed using a Gaussian kernel of 8 mm FWHM.

For each participant, at each voxel, a GLM was used to contrast the BOLD response to difficult versus easy problem-solving. One regressor for each condition was constructed by convolving the duration of each block with the canonical HRF. Additional covariates of no interest included the six movement parameters from the realignment step, and a constant regressor to model the session mean. The model and data were high-pass filtered with a cutoff of 1/128 Hz. Estimation of the model produced $\beta$ maps for each of the easy and difficult conditions. The difference map ($\Delta$BOLD, difficult minus easy) summarized the BOLD response to difficult problem-solving for each participant. The mean group effect was assessed using a $t$ test per voxel.

*ROIs and voxel-wise analyses.* For the analysis replicating the mediation observed by Samu et al. (2017), the fMRI response was summarized within each of 12 cortical networks reported by Ji et al. (2019), based on the multimodal parcellation from the Human Connectome Project (Glasser et al., 2016). Parcels were generated from https://neurovault.org/images/30759/, dilated to fill a gray-matter mask, and combined into one ROI per network. The core MDN regions, areas consistently and strongly activated by multiple cognitive demands, have been shown to comprise a subset of the larger frontoparietal resting state network (Assem et al., 2020a), so these regions were added as an extra "core MDN" ROI, expected to be most strongly responsive to the difficulty contrast. Similarly, a core DMN ROI was constructed using the same data and approach as in Assem et al. (2020a), but reversing the sign of the contrast (i.e., selecting parcels consistently most active in easier compared with harder conditions across a set of tasks). This identified a midline subset of the broader DMN network parcels. The response within each ROI was summarized by the mean across voxels.

Subsequent analyses used a voxel-wise approach to precisely identify those voxels where the mediation model was significant, and to assess the association between the strength of the mediation effect and the strength of the response to cognitive demand. Across most of the brain, fMRI data were available from all 252 participants; however, toward the edges of the brain data were missing from some participants depending on the position of the acquisition bounding box. We analyzed all voxels within an MNI template brain mask for which fMRI data were acquired from at least 100 participants. The sample size thus ranged from 100 to 252 across voxels (73% of voxels within the mask had data from all 252 participants; 90% of voxels had data from at least 90% of participants; 98% of voxels had data from >100 participants).

The set of voxels exhibiting both significant mediation and significant activation by cognitive demand (and with data from all 252 participants) then served as a functional ROI across which the BOLD response was averaged to fit a summary mediation model. This summary model was used in further analyses to examine robustness to unmodeled confounders, estimated causal direction, and moderated mediation.

For both ROI-based and voxel-based analyses, multiple comparisons were accounted for by controlling the false discovery rate (FDR) at < 0.05 (Benjamini and Yekutieli, 2001). Brain renderings are displayed using MRIcroGL software (https://www.nitrc.org/projects/mricrogl).

*Mediation analyses.* The mediation model tested whether the relationship between age and fluid intelligence could be (at least partially) accounted for by the relation between age and the brain response to difficult problem-solving. Mediation was assessed in MATLAB using a set of linear regressions (Baron and Kenny, 1986; MacKinnon et al., 2007). The first equation below expresses the total linear relation between age and IQ; the second equation expresses the unique linear relation between age and IQ when also modeling the effect of the BOLD response on IQ as follows:

$$C = i_1 + cA + e_1 \qquad (1)$$

$$C = i_2 + bB + c'A + e_2 \qquad (2)$$

$$B = i_3 + aA + e_3 \qquad (3)$$

where the variable $A$ is age, $B$ is the brain's BOLD response at a given voxel or ROI, and $C$ is Culture Fair IQ; the coefficient $a$ reflects the effect of age on the brain response, $b$ reflects the effect of the brain response on IQ while controlling for age, $c$ reflects the total effect of age on IQ, and $c'$ reflects the "conditional direct" effect of age on IQ while controlling for the brain response; $e_{1-3}$ are residuals and $i_{1-3}$ are intercept terms.

Where the data are consistent with mediation, $c'$ would have reduced magnitude compared with $c$; that is, the inclusion of the brain response in the model explains some of the variance in IQ that would otherwise have been explained by age. For a linear model, the difference between $c$ and $c'$ is equivalent to the product of $a$ and $b$ (MacKinnon et al., 2007), which describes the "indirect" effect of age on IQ as the effect of age on the brain response combined with the (age-adjusted) effect of the brain response on IQ.

Mediation is traditionally tested and easiest to interpret when there is no interaction between the mediator and independent variable in predicting the outcome variable (as assumed in Eq. 2, above). Therefore, before testing for mediation, we tested for ROIs or voxels where age and the fMRI response interact in predicting IQ, which would suggest that the brain response moderates (buffers or exacerbates) the direct effect of age on IQ. That is, we tested the interaction term ($d$) in the model as follows:

$$C = i_4 + bB + c'A + dAB + e_4 \qquad (4)$$

The significance of mediation can be assessed by separately testing the significance of $a$ and the significance of $b$, or by testing the significance of the product of $a$ and $b$ directly (Baron and Kenny, 1986; MacKinnon et al., 2002). While the latter approach is more powerful under the null hypothesis that both $a$ and $b$ are zero, it has inflated Type I error rates compared with a null hypothesis that either $a$ or $b$ might be zero (MacKinnon et al., 2002). We therefore used the more conservative conjunction of tests, which also allowed an efficient hierarchical approach in the context of testing multiple voxels across the brain. Specifically, we first identified voxels where the strength of the brain response (the potential mediator) showed a significant bivariate association with age (significant $a$ coefficient in Eq. 3), thresholding for significance at $p < 0.05$, two-tailed, while correcting for multiple comparisons (FDR). Of these voxels, we retained only those where there was also no evidence of moderation, defined as the interaction term ($d$) being both small (Cohen's $f^2 < 0.02$) and nonsignificant ($p > 0.05$, two-tailed, without correction for multiple comparisons). This generated a conservative set of voxels within which classical mediation could then be tested based on additional significance of the $b$ coefficient (Eq. 2). Since IQ has a negative relation with age, and we are interested in "consistent" mediation (where the direct and indirect effects have the same sign, such that modeling the mediator reduces the size of the total effect) (MacKinnon et al., 2000), we used a one-tailed test that $b$ had the opposite sign to $a$ (i.e., IQ increases with brain response, which decreases with age, or vice versa). Multiple comparisons were again accounted for by controlling the FDR <0.05, based on the conjunction of the tests for $a$ and $b$ (maximum $p$) (Heller et al., 2007).

*Correlation of mediation effect size with strength of BOLD response to cognitive demand, across voxels.* Testing whether voxels with the strongest mediation are also those most strongly activated by cognitive

demands is complicated by spatial autocorrelation across the brain. Nearby voxels tend to have similar responses, so standard significance tests that assume independence of samples are invalid. We therefore tested significance using a Monte Carlo approach based on Moran eigenvector spectral randomization (Wagner and Dray, 2015). The eigenvectors of a scaled proximity matrix derived from the Euclidian distance between voxels (Dray et al., 2006) comprehensively describe the spatial autocorrelation structure across all scales (Griffith and Peres-Neto, 2006). These eigenvectors were used to create a null model of the distribution of correlations that would be expected by chance, given the measured autocorrelation structure (Wagner and Dray, 2015). We used the MATLAB implementation in the Brainspace toolbox (Vos de Wael et al., 2020), using the "singleton" procedure and 10,000 random permutations.

*Diagnostic analyses of summary mediation model.* For maximum sensitivity, diagnostic analyses of the mediation model were run using the mean BOLD signal across all voxels with significant mediation, data from all 252 participants, and a preferential response to difficult problem-solving. To confirm correct specification of the functional form of Equations 1–3 (i.e., that additional nonlinear functions of the independent variables are not required to fit the data), we used the RESET test (Ramsey, 1969), as implemented in the Panel Data Toolbox (Alvarez et al., 2017). To test for heteroscedasticity of residuals from Equations 1–3 (i.e., whether residual variance varied as a function of the independent variables), we used the version of the White test proposed by Wooldridge (2012).

Being based on regression between observed variables, relationships in the mediation model could potentially be induced indirectly by unmodeled variables that drive covariation between the observations. For example, if some participants were more distracted in the scanner, this would likely lead to both lower performance and weaker task-induced BOLD signal. By using a measure of fluid intelligence acquired in a previous session, we avoid confounders that could affect simultaneous measures of performance and BOLD signal, and so ensure that the results generalize to a standard measure of fluid intelligence that is stable over time. Nonetheless, other potential confounds remain (e.g., general predisposition to distraction) and cannot be exhaustively excluded. Therefore, to assess the robustness of the mediation result to unmodeled confounders, we performed a sensitivity analysis using the Left Out Variables Error method (Mauro, 1990; MacKinnon and Pirlott, 2015). This analysis asks how much the observed mediation strength might be overestimated in the presence of hypothetical confounders that correlate to varying degrees with the modeled variables. We expect age to be a cause rather than an outcome, and so to be unaffected by unmodeled variables. In this analysis, we therefore focused on potential confounding of the relationship between the brain and fluid intelligence measures. The analysis was implemented in MATLAB based on the example in Valente et al. (2017).

The mediation analyses test whether the data are consistent with the hypothesized causal model (i.e., IQ and neural responsiveness are both affected by age, and IQ is also affected by neural responsiveness); however, they cannot distinguish between alternative causal models (Fiedler et al., 2011). We therefore estimated the causal relationship between each pair of variables using a "linear non-Gaussian acyclic model" (LiNGAM), which can recover the causal directions under the additional assumptions of an acyclic model with no more than one error term being perfectly Gaussian (Shimizu et al., 2006). Briefly, the method starts from the observation that the matrix description of the linear relations between a set of centred random variables $X$:

$$X = BX + E,$$

where $B$ is a coefficient matrix and $E$ are independent error terms, can be rearranged to the following:

$$X = (1 - B)^{-1}E.$$

If $E$ are assumed to be not perfectly Gaussian, then independent component analysis (ICA) (Comon, 1994; Hyvarinen, 1999) can be used to decompose $X$ into the independent error components multiplied by a mixing matrix, up to an undetermined scaling and permutation. The coefficient matrix $B$ can then be derived from the mixing matrix, with the correct permutation and scaling determined from the assumptions that the model is linear and acyclic, and so $B$ should be lower-triangular. The model was estimated with an algorithm based on this ICA-plus-permutation approach, using the MATLAB LiNGAM package (https://sites.google.com/view/sshimizu06/lingam). Since ICA-LiNGAM is not scale invariant, all variables were standardized to unit variance. The model was first estimated using the mean BOLD signal averaged across mediating voxels as defined above. Directionality of each path was defined as the estimated absolute connection strength in the hypothesized direction, minus the absolute connection strength in the reverse direction (of which only one is non-zero, given the acyclicity assumption). Thus, positive values would reflect causality in the hypothesized direction. The reliability of this estimate was assessed using bootstrap (15,000 resamples). Since the bootstrapped sampling distribution was found to be biased with respect to the observed statistic, and far from Gaussian, it was not conducive to constructing CIs. We therefore used the watershed algorithm to split the mediating voxels into 77 clusters (one per local minimum), ran LiNGAM on each cluster, discarded clusters where the sign of any undirected coefficient did not match that estimated from the all-voxel model (mean 44% across bootstrap resamples) or the LiNGAM algorithm warned that the coefficient matrix was not lower-triangular (mean 26% across bootstrap resamples), and calculated the mean directionality estimates across remaining clusters. This procedure was again repeated across 15,000 bootstrap resamples. In line with the central-limit theorem, using this mean estimate of directionality produced an approximately normal sampling distribution, which allowed a bias-corrected 95% CI to be constructed.
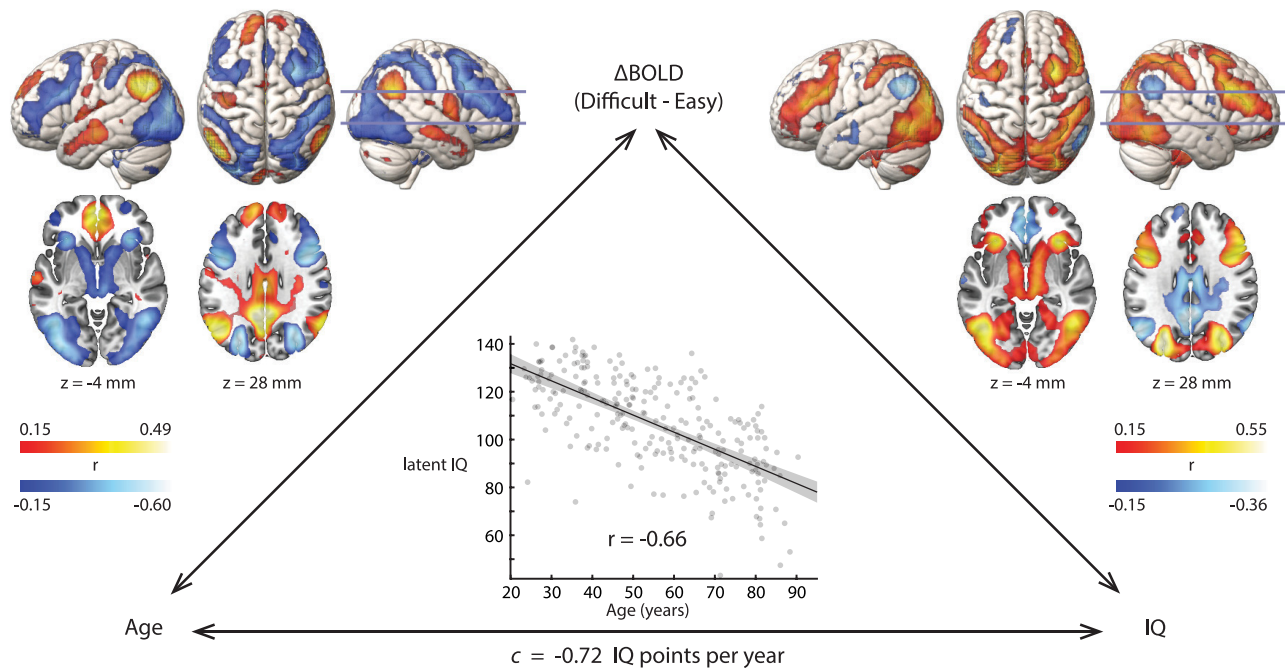
*Moderated mediation analyses.* We ran moderated mediation analyses to test whether the coefficients of the proposed mediation model depended on the level of a fourth variable (see Lifestyle activities measures). We used the approach of Edwards and Lambert (2007), allowing the moderator to affect any of the paths in the model, and testing for moderation of both single and compound paths. The analysis was implemented in MATLAB, based on the example in Edwards and Lambert (2007). Simple effects were calculated for levels of the moderator 1 SD above and below the mean. Differences between levels of the moderator were tested using standard parametric tests for the single paths ($a, b, c'$), and using bias-corrected percentile bootstrap (15,000 resamples) for the compound paths (indirect and total effects of age on fluid intelligence).

*Control analyses adjusting for head motion.* Head motion is expected to increase with age and to decrease with fluid intelligence (Siegel et al., 2017), and is therefore a potential confound as well as a source of noise. While this is of particular concern for measures of functional connectivity (Ciric et al., 2018; Parkes et al., 2018), it can also degrade task fMRI data (Siegel et al., 2014). Therefore, on the suggestion of a reviewer, all analyses were repeated after regressing summary measures of individual differences in head motion. Two measures of head motion were calculated, using SPM Utility Plus (Pernet, 2021): the mean frame-wise displacement (FD) across the fMRI scan (Power et al., 2012) and the proportion of high-motion frames (FD > 0.9 mm, with this threshold being suitable for task-based fMRI in cohorts with moderate motion) (Siegel et al., 2014). All analyses were repeated after regressing these measures from the key variables of age, IQ, and the BOLD response at each ROI/voxel, as well as lifestyle measures in analyses of moderated mediation. This carries a risk of removing actual effects of interest, to the extent that they happen to covary with head motion (Power et al., 2014); however, as a supplementary analysis it can identify results where head motion may be (or is unlikely to be) a confound, and could unmask results that might otherwise be obscured by the variance that head motion shares with other variables.

## Results

### Behavior
Fluid intelligence (IQ), measured using the Culture Fair test (Cattell and Cattell, 1973), had a mean of 106.6 and SD of 19.5 ($N = 252$; range 47-158 using the conversion in the manual, and

**Figure 2.** Bivariate relationships between age, fMRI response, and fluid intelligence ($N = 252$). Voxel-wise correlation coefficients are thresholded based on significance at FDR $< 0.05$. Slice positions are labeled in MNI coordinates and marked on the right hemisphere rendering. Surface renderings show activations to a depth of 15 mm. The gray band in the scatter plot represents the 95% CI of the fitted line.

43-142 when estimated as a latent variable; see Materials and Methods). Since the test was originally constructed to have a population mean of 100 and SD of 16, our higher mean may reflect the Flynn effect (Colom and Garcia-Lopez, 2003), while the higher variance may reflect our wide age range. It is unclear whether the lowest values reflect genuinely low fluid intelligence in older participants, or a qualitative difference in their ability to understand or perform the tasks. Since we had no principled reason to exclude the lowest performers, all participants were included in the analysis. However, on the suggestion of a reviewer, all analyses were repeated after excluding participants whose latent IQ score was >3 SDs below the mean (one participant in the ROI analysis plus a second in the subsequent analyses; manual-normed IQ scores 47 and 57; latent IQ scores 43 and 48; ages 71 and 86). The main conclusions were unchanged.

For the fMRI task, reaction time (RT) was calculated as the median time spent on each accurately answered problem at first attempt, per condition and participant. Participants were, as expected, substantially faster on the easy (mean RT = 1.02 s) than the difficult (mean RT = 4.83 s) problems ($t_{(251)} = 37.9$; $p = 7.4 \times 10^{-106}$). Participants were also substantially more accurate on the easy (mean accuracy = 97.6%) compared with the difficult (mean accuracy = 56.2%) problems ($t_{(251)} = 34.2$; $p = 2.4 \times 10^{-96}$).

**fMRI difficulty contrast**
The group mean BOLD response associated with difficult versus easy problem-solving is shown in Figure 1B. As expected, the difficult condition produced greater activation of the MDN bilaterally, extending into occipital cortex which may reflect enhanced attention to the visual stimuli. The difficult condition was also associated with reduced response of the DMN, auditory, and sensorimotor cortex. Reduced responses in sensorimotor cortex, especially in the left hemisphere, are expected because of the less frequent (right-hand) button presses; reduced activation in auditory cortex may reflect attentional suppression of the scanner noise with increased focus on the visual modality.

**Bivariate relationships between age, fMRI, and fluid intelligence**
We start by describing the three bivariate relationships in the full sample. First, the association between age and fluid intelligence (Fig. 2, bottom) showed strong and approximately linear cross-sectional decline across the age range, as reported previously using related measures and overlapping samples of Cam-CAN participants (Kievit et al., 2014; Samu et al., 2017). The Pearson correlation coefficient ($r = -0.66$) corresponded to an average loss of 7.2 IQ points per decade of age.

Second, we assessed how the fMRI response to difficult problem-solving depended on the age of the participants (Fig. 2, left). To a first approximation, the results corresponded to a weakening of the typical response, as age increased: regions that were activated by the difficult condition on average (Fig. 1B, warm colors) showed less activation (Fig. 2, left, cool colors), while DMN regions that were typically suppressed in the difficult condition showed less suppression. We note two partial exceptions to this general pattern: auditory and sensorimotor cortex, which were strongly suppressed in the difficult condition, showed relatively less difference across age, or even increased suppression (anterior IPS); similarly, parts of the basal ganglia showed increasingly negative response to difficulty with age, despite the difficulty response being nonsignificant (caudate) or weakly negative (putamen) on average.

Third, we assessed how the fMRI response to difficult problem-solving correlated with individual differences in fluid intelligence (Fig. 2, right). The pattern was remarkably similar to the correlation with age, but inverted in sign. That is, regions whose activity was associated with higher fluid intelligence tended to be those with the greatest age-related decline in activity, and vice versa. One exception was again auditory and sensorimotor regions: although these were suppressed in the difficult condition, and suppression somewhat reduced with age, we saw little association with IQ, consistent with the expectation that these regions are responding to specific sensory and motor aspects of

**Table 1. BOLD response to cognitive difficulty, and its moderation and mediation of age effects on fluid intelligence, for cortical network ROIs (N = 154)**

| Network | ΔBOLD response (difficult − easy) | | | Moderation (age:ΔBOLD interaction) | | | Mediation (age → ΔBOLD → IQ) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Cohen's d | p | FDR-adjusted p | Cohen's d | p | FDR-adjusted p | ab | p (a and b) | FDR-adjusted p (a and b) |
| Core multiple-demand[a] | 1.282 | <0.001 | <0.001[b] | 0.163 | 0.045 | 0.636 | −0.076 | <0.001 | 0.002[b] |
| Frontoparietal | 1.026 | <0.001 | <0.001[b] | 0.097 | 0.228 | 0.800 | −0.049 | 0.004 | 0.015[b] |
| Dorsal-attention | 0.849 | <0.001 | <0.001[b] | 0.043 | 0.593 | 0.980 | −0.087 | <0.001 | 0.002[b] |
| Secondary visual | 0.822 | <0.001 | <0.001[b] | −0.019 | 0.815 | 0.980 | −0.067 | <0.001 | 0.002[b] |
| Primary visual | 0.425 | <0.001 | <0.001[b] | −0.103 | 0.205 | 0.800 | −0.024 | 0.200 | 0.399 |
| Ventral multimodal | 0.219 | 0.007 | 0.011[b] | −0.033 | 0.687 | 0.980 | −0.023 | 0.032 | 0.089 |
| Language | 0.179 | 0.027 | 0.035[b] | −0.018 | 0.823 | 0.980 | −0.004 | 0.778 | 0.980 |
| Orbito-affective | 0.052 | 0.520 | 0.520 | −0.006 | 0.939 | 0.980 | −0.017 | 0.090 | 0.211 |
| Posterior multimodal | −0.126 | 0.120 | 0.129 | 0.002 | 0.980 | 0.980 | 0.010 | 0.980 | 0.980 |
| Cingulo-opercular | −0.135 | 0.095 | 0.111 | −0.006 | 0.944 | 0.980 | <0.001 | 0.972 | 0.980 |
| Default | −0.180 | 0.027 | 0.035[b] | 0.010 | 0.901 | 0.980 | 0.023 | 0.964 | 0.980 |
| Core default-mode[a] | −0.572 | <0.001 | <0.001[b] | −0.016 | 0.846 | 0.980 | 0.026 | 0.856 | 0.980 |
| Auditory | −0.722 | <0.001 | <0.001[b] | −0.079 | 0.331 | 0.928 | 0.017 | 0.938 | 0.980 |
| Somatomotor | −1.018 | <0.001 | <0.001[b] | −0.126 | 0.120 | 0.800 | 0.007 | 0.910 | 0.980 |

[a]Networks defined based on conjunction of response to easy/difficult conditions using data from three Human Connectome Project tasks as presented in Assem et al. (2020a); remaining networks are defined from functional connectivity as presented in Ji et al. (2019).
[b]Significant results after FDR correction (adjusted p < 0.05).

the particular task design, rather than having a more general role in fluid intelligence.

**Mediation of age-related differences in IQ by neural response to cognitive demand: network-based replication**
Given the pattern of bivariate relationships, it is possible that, for some voxels, age independently impacts fluid intelligence and the BOLD response. To identify regions where a reduced response to difficult problem-solving potentially mediates the effect of age on fluid intelligence, we jointly predict IQ from both age and the fMRI response, and test for a decrease in the remaining (direct) effect of age on IQ, relative to the total effect of age on IQ when modeled alone. In other words, we identify regions where variation in IQ is partly explained by an indirect path of age affecting the brain, which has a consequent effect on IQ. This also entails the relationship between the fMRI response and IQ remaining significant when controlling for age.

We first sought to replicate the recent finding that the responsiveness of frontoparietal regions is consistent with such mediation of age differences in fluid intelligence (Samu et al., 2017). For this analysis, we thus restricted ourselves to the independent subset of participants who were not included in the preceding study (N = 154), and we summarized the brain response according to major cortical networks (Ji et al., 2019) plus more focused definitions of core MDN (Assem et al., 2020a) and DMN. An initial test for an interaction between age and brain response in predicting IQ found no significant effects for any network after adjusting for multiple comparisons (all FDR-adjusted p > 0.64). Therefore, consistent with Samu et al. (2017), there was no evidence that the relationship between the neural response and fluid intelligence differed with age.

We then probed neural mediation of the age effect on fluid intelligence. Significant mediation was observed in four networks: Dorsal-attention, Secondary-visual, Frontoparietal, and the Core MDN (which is a subset of the Frontoparietal network), confirmed by joint significance of their a and b paths (Table 1) as well as bootstrapped CIs around their product (Fig. 3A). We therefore replicate the mediation observed by Samu et al. (2017), link it to particular functional networks, and extend it from mediation of concurrent task performance to mediation of participants' fluid intelligence more generally. The mediation effect size (ab) is plotted against the group-average fMRI response to

difficult versus easy problem-solving in Figure 3A. The same data are plotted a different way in Figure 3B, breaking down each mediation effect into the magnitude of the a and b coefficients. This shows that the core DMN is impacted by age (a) at least as strongly as the networks activated by the difficult condition (red); the reason that it does not significantly mediate the decline in IQ is because of its small effect on IQ after controlling for age (b), again replicating Samu et al. (2017). Overall, it is striking that evidence for mediation is specific to those networks responding most positively to task difficulty. This is consistent with correlation of the BOLD response with the a and b paths separately, reported across ICA components by Samu et al. (2017).

The mediation results considered so far are at a relatively coarse spatial scale, using combined ICA-based brain-wide spatial components (Samu et al., 2017) or functional/anatomic networks (our replication). Next, we combine both samples to increase power and precisely localize the mediation effect on a voxel-wise basis. This also allows us to quantify the degree to which mediation is associated with voxels that respond positively to task difficulty, separately for voxels activated and voxels suppressed by cognitive demand, and accounting for spatial autocorrelation.

**Voxel-wise localization of mediation effect, and its association with voxels most responsive to cognitive demand**
To identify mediating voxels in a conservative but efficient manner, we took a hierarchical approach where we first tested for significance of the a path, and used significant voxels (after FDR correction) to define an analysis mask within which to test for voxels where the b path was also significant. We further restricted the analysis mask to voxels where there was no evidence of an interaction between age and brain response in predicting IQ. Consistent with Samu et al. (2017) and with our independent replication at the network level, no voxels showed such an interaction after correcting for multiple comparisons (all FDR-adjusted p > 0.32). Nonetheless, to be conservative, we excluded voxels based on uncorrected significance (p < 0.05) or more than "small" effect size (Cohen's $f^2 > 0.02$) of the interaction term.

The resultant analysis mask is shown as the blue and red overlays in the second row of Figure 4A, where the color represents the sign of the difficulty contrast (red represents difficult >
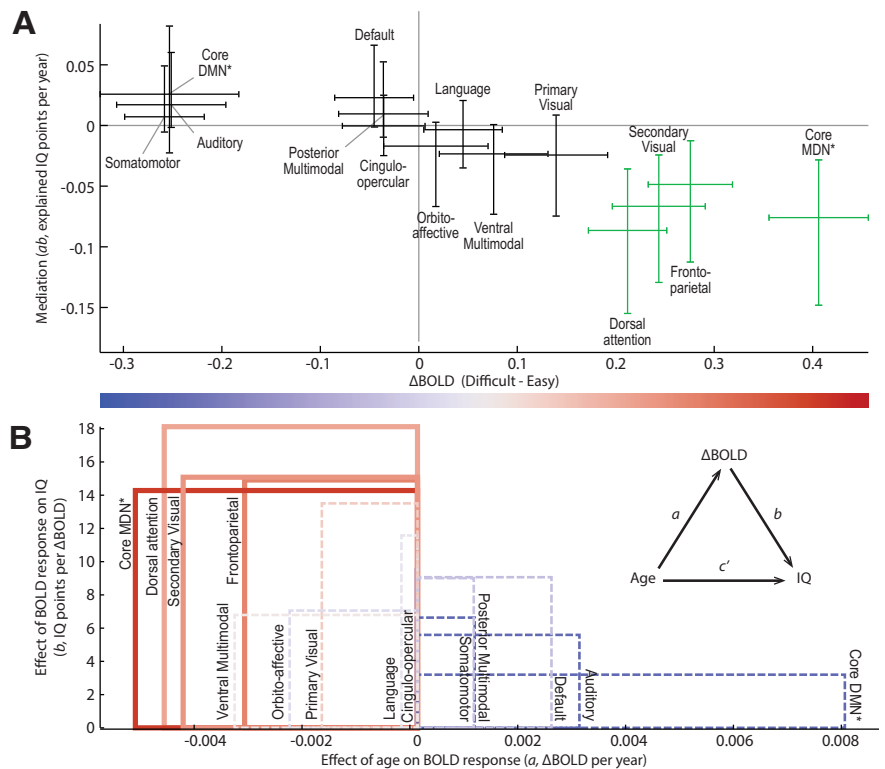
easy; blue represents easy > difficult). Green represents voxels with significant mediation after correcting for multiple comparisons within this mask (FDR < 0.05). Interestingly, >90% of voxels with significant mediation responded more to difficult than to easy problem-solving. Identified voxels show a close relationship to MDN regions, including key foci along the precentral sulcus and middle frontal gyrus, and within the intraparietal sulcus, anterior insula, and anterior cingulate cortex. Mediating voxels are also observed in lateral occipital cortex and subcortical structures, including the basal ganglia, thalamus, and cerebellum, which are often coactivated with the cortical MDN (Assem et al., 2020a).

The relationship between neural mediation of fluid intelligence and a voxel's average response to difficult problem-solving is illustrated in Figure 4B, where the mediation effect size is plotted against the response to difficult versus easy problems. The overall relationship across voxels is highly significant (Pearson correlation, $r = -0.57$; $p < 0.001$; Fig. 4B,C, purple) after using a permutation test with Moran spectral randomization to comprehensively account for spatial autocorrelation in the maps (Wagner and Dray, 2015; Vos de Wael et al., 2020). We also test the correlation separately within voxels significantly activated by difficult versus easy problem-solving, and within voxels significantly suppressed by difficult versus easy problem-solving. The correlation is again significant within positively activated voxels ($r = -0.41$; $p < 0.001$; Fig. 4B,C, red), whereas there is no evidence that mediation is correlated with the level of difficulty-induced suppression ($r = -0.04$; $p = 0.52$; Fig. 4B,C, blue).

**A summary mediation model, and diagnostic analyses: robustness to confounders and a test of causal direction**

To summarize and illustrate the localized mediation effect, we averaged across those voxels with significant mediation, data from all 252 participants, and a preferential response to difficult problem-solving. For these voxels, the mediation model is depicted in Figure 5A. Figure 5B illustrates the mediation effect in terms of the underlying relationships as suggested by MacKinnon et al. (2007), where the green triangles highlight the reduction of the total effect of age when modeling the neural response ($c-c'$).

Since the mediation effect rests on standard assumptions of linear regression, we next confirmed that the model is appropriate. We used the RESET test (Ramsey, 1969) to confirm correct specification of functional form (i.e., that nonlinear functions of the independent variables were not required to fit the data). For Equations 1–3, we found no evidence of misspecification (all $F_{(3, 246-7)} < 2.23$, all $p > 0.08$). Although the linearity of relationships with age may break down at the extremes of the lifespan, approximate linearity is common for accuracy-based performance measures (Salthouse, 2011a) and BOLD responses (Grady et al.,
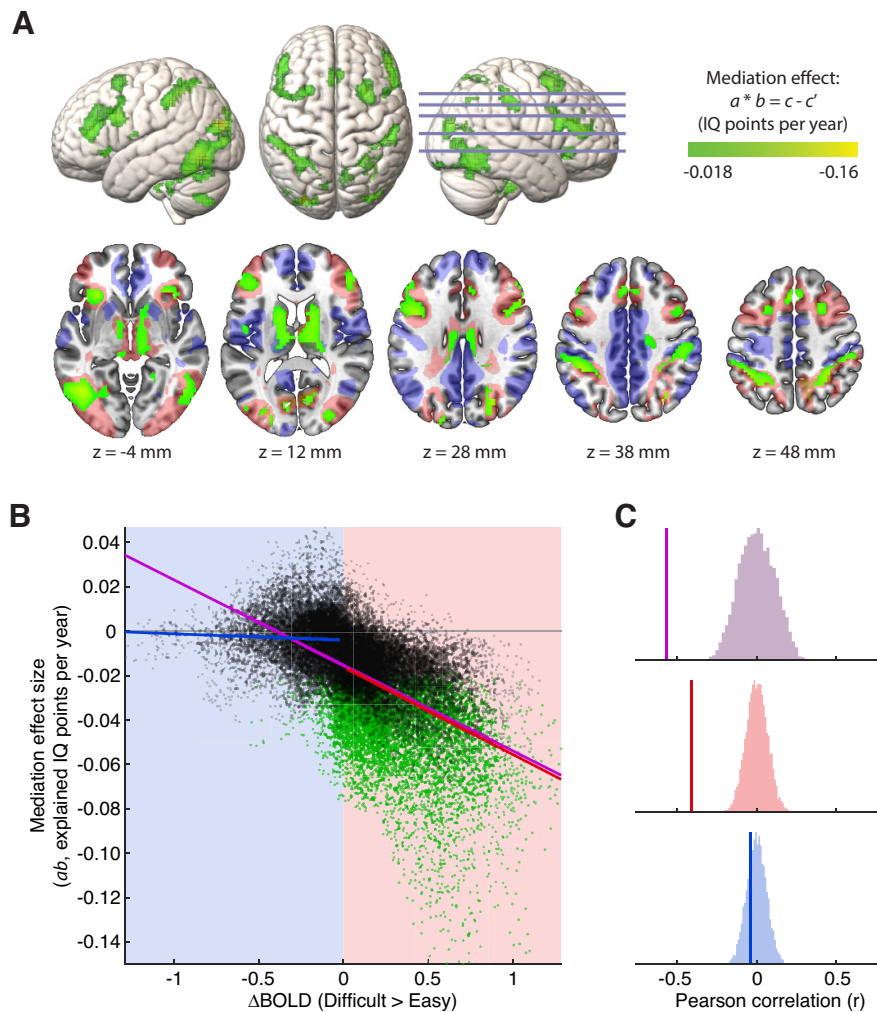


**Figure 3.** BOLD response to cognitive difficulty, and its mediation of age effects on fluid intelligence, for cortical network ROIs ($N = 154$). ***A***, Overall mediation effect size versus group average BOLD response. Error bars are 95% CIs derived from the associated $t$ test (BOLD contrast) or bias-corrected bootstrap (mediation). Green represents networks with significant mediation after FDR correction (adjusted $p < 0.05$). ***B***, The $b$ coefficient (reflecting the age-adjusted effect of the BOLD response on IQ) is plotted versus the $a$ coefficient (reflecting the effect of age on the BOLD response), with group-mean BOLD response represented by color. Inset, Mediation model. The area of each rectangle conveys the mediation effect size ($ab$) for the corresponding network. Solid and dashed lines indicate significant and nonsignificant mediation, respectively. Asterisks indicate networks defined from their conjunction response to easy/difficult conditions using data from three Human Connectome Project tasks as presented in Assem et al. (2020a); remaining networks are defined from functional connectivity as presented in Ji et al. (2019).

2006) during healthy aging. Next, we used the White test (Wooldridge, 2012) to assess whether the residual variance varied as a function of the independent variables. For Equations 1–3, we found no evidence against homoscedasticity (all $\chi^2_{(2)} < 4.36$, all $p > 0.11$). We also note that any violation of the assumption that variables are measured without error would underestimate the magnitude of the true mediation effect (Pieters, 2017).

The mediation analysis assumes that there are no confounding variables omitted from the model (MacKinnon and Pirlott, 2015). Potential unmeasured confounders, however, are plausible, if not inevitable, in practice. Assuming that chronological age must be a cause rather than a consequence, the main concern is that an unmodeled variable could produce a spurious mediation effect by covarying with both IQ and the neural measure (Salthouse, 2011a). Using a trait measure of fluid intelligence rather than concurrent performance avoids some potential sources of shared variance, but others will remain. Therefore, to infer a direct relationship between neural responsiveness and fluid intelligence, and to establish the validity of the mediation model as a potential mechanistic explanation, it is important that the mediation result is robust to possible unmodeled covariates. To address this, we used sensitivity analysis (Mauro, 1990; Valente et al., 2017) to estimate how strong any unmodeled confounders would need to be to fully explain the observed mediation relationship. Figure 5C plots the size of the true mediation effect as a function of an unknown confounder's correlation with the fMRI

**Figure 4.** Voxel-wise mediation analyses. ***A***, Regions with significant mediation (FDR < 0.05 within analysis mask) are shown in green, overlaid on surface renderings (top) and horizontal slices (bottom). Slice positions are labeled in MNI coordinates and marked on the right hemisphere rendering. In the bottom row, the analysis mask is shown by the blue/red overlay, colored by the sign of the difficulty contrast (red represents difficult > easy; blue represents easy > difficult). Surface renderings represent activations to a depth of 30 mm. ***B***, Scatterplot represents the relationship between the size of the mediation effect and the group-mean difficulty contrast. Each point is a voxel within the analysis mask, colored green where mediation is significant (FDR < 0.05). ***C***, Pearson correlations (vertical lines) between the size of the mediation effect and the difficulty contrast, compared with permutation null distributions constructed using Moran spectral randomization. The correlation is separately tested across all voxels (purple), only voxels with a positive response to task difficulty (red), and only voxels with a negative response to task difficulty (blue), corresponding to the lines of best fit in ***B***.

response and with IQ. This shows that, for there to be no true mediation (at points along the black line), the correlations of any confounder with both the fMRI response and IQ would need to exceed 0.43 on average (and at least one must exceed 0.45). While this cannot be ruled out, it provides reassurance that the mediation result is robust to the possibility of moderate confounding.

Although the mediation analyses confirm that the data are consistent with the hypothesized causal model, without experimental manipulation of the variables, the data are equally consistent with alternative causal models (Fiedler et al., 2011; Salthouse, 2011a; MacKinnon and Pirlott, 2015). The various possible causal orderings are statistically equivalent, having the same covariance matrix and global model fit, and one cannot adjudicate between them based on the size or significance of their mediation effect (Thoemmes, 2015). To assess causal directionality, we therefore used a LiNGAM (Shimizu et al., 2006). By

making additional assumptions that no more than one error term is perfectly Gaussian, and that the model is acyclic, LiNGAM uses ICA to estimate the generating causal model based on purely observational data (see Materials and Methods). Using the mean across mediating voxels, as defined above, LiNGAM estimated the causal directions to be in line with the assumed model (Fig. 5D, dashed red lines). Bootstrap resampling, however, revealed the sampling distribution to be biased and far from multivariate normality. We therefore split the mediating voxels into local clusters, applied LiNGAM to each cluster, discarded clusters where a causal model could not be reliably identified or whose undirected model coefficients differed in sign from the all-voxel model, and averaged directionality estimates across the remaining clusters. Again, estimated causal directions were all in line with the assumed model (Fig. 5D, solid red lines), now with approximately Gaussian bootstrapped sampling distributions (gray histograms) from which 95% bias-corrected percentile CIs were constructed (pink bands). The bootstrap distributions remain shifted with respect to their observed direction estimates (solid red lines), suggesting that the observed estimates are similarly biased with respect to their true population values; therefore, the bias-corrected CIs do not lie on the percentiles of the sampling distribution but are shifted in the opposite direction to compensate (Hesterberg, 2015). None of the intervals spans zero, suggesting confidence in the estimated directions under the assumptions of the model, although the interval for the brain-behavior relation is close to zero and the assumptions in this case are questionable (see Discussion).

## Moderation of the mediation model by a varied active lifestyle

Finally, in exploratory analyses, we tested whether any path in the observed mediation model might be moderated by self-reported regular physical activities, in terms of their variety (number of different activities; range 0-14), mean frequency per activity (episodes per year; range 12-365), mean duration per episode (range 4-415 min), or total duration (range 10-3240 h per year). We found that the variety of physical activities significantly moderated the path from the brain response to IQ, but not the effect of age on the brain response or the conditional direct effect of age on IQ (Table 2). This carried through to a significant moderation of the overall indirect effect of age on IQ, although moderation of the total effect of age was not significant. In contrast to the moderating effect of the variety of activities, the mean frequency and duration of each activity had no significant effect on any path in the model, nor did the total duration of activity.
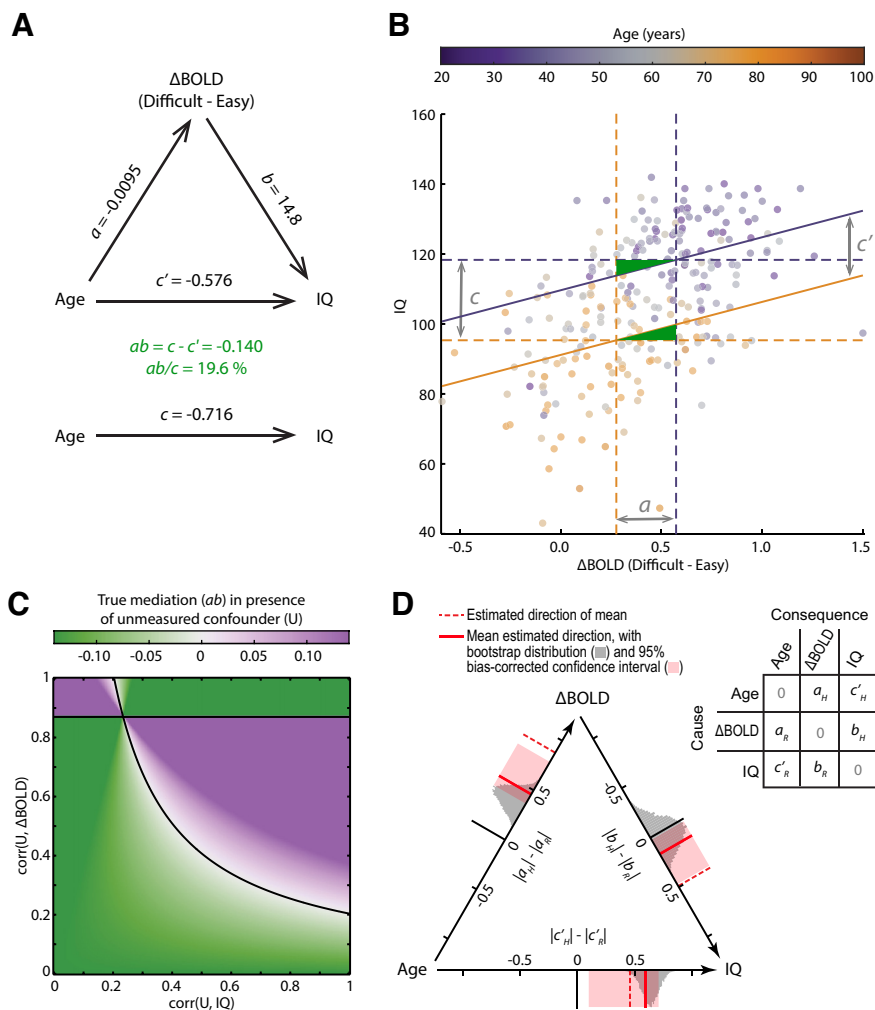
Some dependence between the four activity measures is expected. We therefore repeated the analyses using the residuals

for each measure after regressing out the other three. The results were largely similar, again with an effect of variety, but no effects of frequency, mean duration, or total duration. Now the total effect of age on IQ was also significantly moderated by the variety of activities, driven, as before, by reduction of the effect of brain response on IQ, although moderation of the compound indirect path no longer reached significance. The moderating effect of the variety of regular physical activities is illustrated in Figure 6, showing the change in simple slopes for each single and compound path (Edwards and Lambert, 2007). The substantial mediation effect for people engaging in relatively few regular physical activities (solid lines) is abolished for people engaging in a larger variety of activities (dashed lines), due primarily to decoupling of the relation between the neural response to task difficulty and fluid intelligence (blue lines).

In addition to physical exercise, engagement in socially and intellectually stimulating leisure activities is also thought to be beneficial for cognitive aging (Hughes et al., 2010; Kuiper et al., 2015; Yates et al., 2016; Borgeest et al., 2020). Therefore, given the significant moderating effect of the variety of regular physical activities, we next asked whether this generalizes to the variety of regular nonphysical activities, again measured by questionnaire, as well as in a home interview (number of social or intellectual activities; range 4-43). This time there was no moderating effect on any path in the model (Table 2).

## Control analyses adjusting for head motion

Individual differences in head motion can covary with both age and fluid intelligence (Siegel et al., 2017), and degrade task fMRI data (Siegel et al., 2014). Therefore, the potential contribution of head motion was assessed by repeating all analyses after regressing summary measures of mean FD across the fMRI scan, and the proportion of high-motion frames. These measures were highly related ($r = 0.91$) and they both correlated substantially with age ($r > 0.42$, $p < 1.4 \times 10^{-12}$) and IQ ($r < -0.39$, $p < 5.3 \times 10^{-13}$). After regressing both motion measures from all key variables (age, IQ, BOLD response per ROI/voxel, and lifestyle measures), conclusions were largely unchanged, except as follows. In the ROI analysis, the mediation effect in the Frontoparietal Network ROI became nonsignificant, although the effect in the Core Multiple-Demand network remained. The percentage of significantly mediating voxels that were more active in the difficult condition increased from 90% to 95%. The size of the voxel-



**Figure 5.** Summary and assessment of the localized mediation effect, averaged across voxels where mediation is significant, data are available from all participants ($N = 252$), and the difficult–easy contrast is positive. **A**, Fitted mediation model. **B**, Illustration of the mediation effect and underlying relationships. Vertical dashed lines indicate Equation 3 at the first and third quartiles of the age distribution (purple represents 39 years; orange represents 71 years). Horizontal dashed lines indicate Equation 1 at the same ages. Solid lines indicate Equation 2 at the same ages, with their slope equal to coefficient *b*. Gray arrows indicate coefficients *a*, *c*, and *c'*, when age is scaled by the interquartile range. Green triangles represent the standardized mediation effect size, relative to the area of the central rectangle, which can be seen to account for approximately one-fifth of the total effect. **C**, Left Out Variables Error (LOVE) plot indicating how large any correlations with an unmodeled confounder (U) would need to be to reduce the true mediation to zero. The observed mediation effect size lies at the origin. **D**, Estimated causal direction of each path, using ICA-LiNGAM after standardization of variables (see Materials and Methods). Axes represent the difference of absolute model coefficients for the hypothesized (H) minus reverse (R) causal directions, indicated by the matrix. Estimation of the model returns three non-zero coefficients in this matrix, one for each pair of variables, in either of the two possible directions. (Zeros on the diagonal reflect the fact that variables are assumed not to cause themselves.) Dashed red lines were calculated using the mean signal across all mediating voxels as plotted in **B**. Solid red lines indicate mean direction estimates, averaged across local clusters of mediating voxels whose undirected coefficient signs matched those from the mean signal. Gray histograms represent the bootstrap sampling distribution of the cluster-mean directionality. Pink bands represent bootstrapped 95% bias-corrected percentile CIs for the cluster-mean directionality.

summary mediation effect increased from 19.6% to 22.8% of the total effect. In the LiNGAM analysis, estimated mean causal directions remained in the hypothesized direction, although the age–IQ directionality no longer differed significantly from zero.

The largest differences were in the analysis of moderation by everyday activities. Here, the variety of physical activities continued to significantly moderate the second stage of the mediation path ($p = 0.010$ when modeled alone; $p = 0.033$ after covarying the other measures), but now alongside additional moderation by the residual duration of activity per episode ($p = 0.043$), and
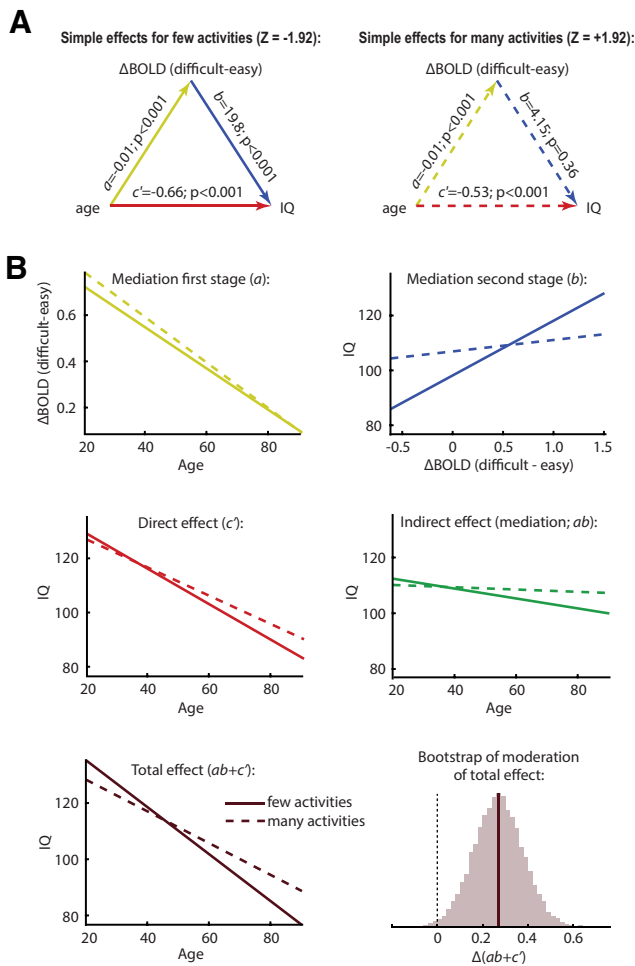
**Table 2. Moderated mediation results[a]**

| Moderator | N | First stage of mediation path | | Second stage of mediation path | | Direct effect | | Indirect effect (mediation) | | Total effect | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\Delta a$ | p | $\Delta b$ | p | $\Delta c'$ | p | $\Delta(ab)$ | p | $\Delta(ab+c')$ | p |
| Physical activities | | | | | | | | | | | |
| Variety | 239 | −0.000046 | 0.983 | −17.1 | 0.0061* | 0.0077 | 0.941 | 0.157 | 0.0095∗ | 0.165 | 0.146 |
| Frequency per activity | 231 | −0.00033 | 0.876 | 4.67 | 0.499 | −0.081 | 0.501 | −0.048 | 0.495 | −0.128 | 0.206 |
| Duration per episode | 220 | −0.0028 | 0.276 | −7.52 | 0.189 | −0.015 | 0.921 | 0.037 | 0.587 | 0.023 | 0.888 |
| Total duration | 220 | 0.00035 | 0.854 | −7.30 | 0.229 | −0.745 | 0.424 | 0.074 | 0.309 | −0.00072 | 0.811 |
| Variety[b] | 220 | −0.00093 | 0.701 | −15.7 | 0.022* | 0.134 | 0.244 | 0.138 | 0.076 | 0.272 | 0.011* |
| Frequency per activity[b] | 220 | −0.0013 | 0.531 | −4.88 | 0.502 | −0.020 | 0.857 | 0.031 | 0.678 | 0.010 | 0.930 |
| Duration per episode[b] | 220 | −0.0032 | 0.211 | −10.1 | 0.096 | 0.089 | 0.519 | 0.057 | 0.522 | 0.146 | 0.295 |
| Total duration[b] | 220 | 0.0013 | 0.499 | 5.42 | 0.373 | −0.087 | 0.348 | −0.036 | 0.586 | −0.123 | 0.342 |
| Nonphysical activities | | | | | | | | | | | |
| Variety | 237 | −0.0040 | 0.101 | −11.5 | 0.074 | −0.087 | 0.500 | 0.049 | 0.468 | −0.038 | 0.769 |

[a]Moderation of each single and compound path in the mediation model, by various moderator variables. Δ indicates the change in coefficients across 2 SDs of the moderator variable. Following Edwards and Lambert (2007), p values for single paths are based on standard errors from the regression model; p values for compound paths (indirect and total effect) are based on bias-corrected CIs derived from bootstrap resampling. Sample size (N) varies because not all participants answered all activities questions.
[b]Moderators that are residualized with respect to each other.
*Significant results at $p < 0.05$.



**Figure 6.** Moderated mediation. The moderator variable (Z) is the variety of regular physical activities, residualized with respect to the other physical activity measures. **A**, Simple mediation effects, at low (left, solid lines) and high (right, dashed lines) levels of the moderator (±1 SD from mean). **B**, The effect of the moderator on each individual and compound path in the model. Solid and dashed lines again indicate low and high levels of the moderator, respectively, as in **A**, with corresponding colors for the single paths. Final histogram represents the distribution of bootstrap samples used to calculate the p value for the moderation of the overall effect of age on IQ.

without significant moderation of the total effect ($p = 0.137$). Most interestingly, the first stage of the mediation path was now significantly moderated by the number of nonphysical activities ($p = 0.008$). This latter effect reflected an exacerbated age effect for people reporting more activities: the BOLD response to difficult problems increased with the number of activities for younger people (<~50 years old), whereas it decreased for older people who reported more activities. This effect of nonphysical activities therefore differed from the effect of physical activities in two ways: it was associated with a different stage of the mediation pathway, and instead of flattening the age effect (Fig. 6, blue lines), it steepened it (see https://osf.io/xgw56/ for results tables from these analyses).

In summary, the key conclusions were unchanged, suggesting that head motion is unlikely to explain the mediation of age-related IQ decline by MDN responsiveness, the association of mediation with voxels responding positively to cognitive challenge, or its moderation by the variety of physical recreation. The emergence of additional lifestyle moderators in this control analysis suggests that head motion may act as a suppressor variable in these cases (MacKinnon et al., 2000), being associated with multiple other variables in opposite directions that can cancel out if not statistically adjusted. This suggests, in turn, that future studies would benefit from careful control of head motion, and that to further understand the complicated relationships that underlie the moderated mediation results it would be useful to replicate them, ideally via direct experimental manipulation.

## Discussion

This paper presents five key results. First, we provide an independent replication that responsiveness of multiple-demand cortex to cognitive demand partially mediates age-related decline in fluid intelligence. Second, strongest mediation is specific to voxels most activated by cognitive demand, and not those suppressed by cognitive demand. Third, the summarized mediation effect is robust to moderate confounding by unmodeled variables. Fourth, assuming unidirectional causality, differences in brain response more likely drive IQ differences than vice versa. Finally, diversity of physical activity moderates the summarized mediation effect, through decoupling of IQ from neural responsiveness to cognitive demand.

The specificity of mediation to the MDN and associated regions is notable, supporting previous hypotheses that "deterioration in DLPFC may (at least partially) underlie the relationship between adult age and abstract reasoning ability" (Phillips and Della Sala, 1998). However, MDN and DMN responses often anticorrelate (Fox et al., 2005), and this anticorrelation can positively or negatively relate to intelligence (Hearne et al., 2016; Santarnecchi et al., 2017b). Task-induced DMN suppression also reduces with age (Grady et al., 2006; Turner and Spreng, 2015), as does its coupling with frontoparietal attention networks (Spreng et al., 2016), while maintenance of DMN deactivation may explain cross-domain differences in cognitive aging (Samu et al., 2017). Although DMN suppression in this task indeed reduced with age, we found, matching Samu et al. (2017), no evidence that this mediated age-related IQ differences. With mediation strength selectively associated with voxels most activated, but not those most suppressed, by cognitive demand, reduced MDN function appears more important than altered DMN function in explaining age-related differences in IQ, while links between DMN activity and IQ may reflect confounding variables, such as age. It remains possible that altered DMN activity may mediate age-related differences in other cognitive domains for which this network is specialized.

We observed no reliable indication that age moderates the association between neural responsiveness and IQ. This is perhaps surprising given the wide age range examined and the expected plasticity of neural recruitment and cognitive strategy (e.g., Cabeza, 2002; Davis et al., 2008; Park and Reuter-Lorenz, 2009), but matches previous observations using this task (Samu et al., 2017). Plausibly, neurocognitive shifts might primarily occur in more domain-specific tasks, when recruitment of domain-general MDN regions is not essential, but offers an optional compensatory strategy. Conversely, fluid intelligence may be especially susceptible to normal aging because of limited capacity for functional plasticity of the networks involved.

Multiple mechanisms undoubtedly link aging and fluid intelligence (Kievit et al., 2016). Indeed, the observed partial mediation explains ∼20% of the relationship between age and IQ, leaving much room for additional mechanisms, alongside neural responsiveness to cognitive demand. For example, various aspects of structural brain integrity may play mediating roles (Salthouse, 2011a; Kievit et al., 2014). Network connectivity is also associated with fluid intelligence (Cole et al., 2012; Barbey, 2018; Dubois et al., 2018; Hilger et al., 2020) and changes with age (Tsvetanov et al., 2016; Bethlehem et al., 2020). Lifespan differences in cerebral vascularization may further influence neural function and neurovascular coupling (West et al., 2019; Tsvetanov et al., 2021), although blood flow variation explains relatively little covariance between age, BOLD response, and performance on this task (Wu et al., 2021). Future work could usefully address the relative importance of multiple mediators, and the relationships between them (Hedden et al., 2016).

While mediation analysis can suggest a mechanistic explanation for an observed relationship, it tests consistency with a hypothesized causal model, rather than inferring the causal structure itself (Fiedler et al., 2011; MacKinnon and Pirlott, 2015; Thoemmes, 2015). Although the LiNGAM method estimated directionality matching the hypothesized model, its assumption of nonreciprocal influences is tenuous in practice, and significance was sensitive to control analyses. Longitudinal data would provide another means to strengthen causal interpretations; if correlated changes are temporally separated, the preceding change may be the more likely cause. Although cross-sectional age-related differences sometimes mirror within-individual longitudinal changes, especially in healthy populations (Salthouse, 2011a), this need not be so (Raz and Lindenberger, 2011), and cross-sectional samples of an underlying longitudinal process may either overestimate or underestimate mediation effect sizes (Maxwell and Cole, 2007). While longitudinal studies present their own challenges (Salthouse, 2011b), the cross-sectional nature of the current dataset imposes interpretational limitations, which would benefit from examination in longitudinal cohorts.

Even assuming a causal role of age, questions remain regarding the relationship between neural responsiveness and IQ. Sensitivity analysis indicated robustness to moderate confounders, suggesting a direct effect, although with uncertain direction. We chose to treat IQ as the outcome variable because this is what we would ultimately hope to improve and because lesion and neuro-stimulation studies suggest that the integrity and function of frontoparietal networks do causally influence IQ (Glascher et al., 2010; Woolgar et al., 2010; Barbey et al., 2014; Momi et al., 2020; Smith et al., 2022). Although it is hard to imagine experimental manipulations of IQ that are not causally dependent on neural responses, we would not claim that the neural response to cognitive demand is unaffected by IQ. Reduced IQ might either increase MDN responsiveness, if puzzles are experienced as more challenging, or reduce responsiveness, if people confidently select incorrect responses or disengage from the task, making the sign of any reverse relationship difficult to predict. Nonetheless, some combination of reciprocal relationships between IQ and MDN function remains likely, and challenging to disentangle.

Finally, we consider the association between more varied physical activity and attenuation of the dependence of IQ on BOLD responsiveness. While this was consistent across control analyses, the other moderated mediation results should be viewed cautiously as they were affected by covarying head motion. Again, the nature of the causal links is difficult to determine. First, there are various biological mechanisms by which an active lifestyle might moderate effects of age on the brain, or brain function on cognition (Cotman and Berchtold, 2002; Cotman et al., 2007; Barnes, 2015). Here, the observed moderation altered the second of these links: that is, although engaging in diverse exercise did not reduce the effect of age on the MDN response, it did buffer its cognitive impact, so the reduced MDN response was less strongly associated with reduced IQ. This is consistent with evidence that exercise may support cognition by improving the efficiency of neural function (Neubauer and Fink, 2009; Voss et al., 2011). Indeed, while people with higher IQ may plausibly seek out more activities, better handle a busier routine, or recall more activities, the literature generally supports a causal effect of exercise on cognition (Smith et al., 2010; Liu-Ambrose et al., 2018).

Assuming a causal benefit of physical activity, it is interesting that its variety appears more important than its frequency or duration, consistent with a previous report (Angevaren et al., 2007), and echoing a role of task-novelty in protecting against cognitive decline (Oltmanns et al., 2017). An additional residual effect of exercise duration emerged after covarying head motion, although this was weaker and less consistent. This finding could provide a principle for designing activity-based interventions with a focus on variety, which may help older adults who are unable

or unwilling to perform single, intensive physical activities. Increasing the variety of activities would represent a lifestyle modification that could be made relatively easily, regardless of specific interests or abilities. Since we examined recent activities, within the preceding year, increasing their variety might be beneficial at any age (but see Chan et al., 2018).

Interestingly, we did not find that the variety of more intellectual and social activities moderated the same brain–IQ pathway. This null result could reflect insensitivity of the particular nonphysical measures, or a unique role of physical activities in boosting neural efficiency, perhaps via growth-factor-mediated neurogenesis and angiogenesis (Cotman and Berchtold, 2002; Cotman et al., 2007), and consistent with mouse models (Kobilo et al., 2011). Other evidence in humans suggests that the conjunction of physical exercise and mental engagement may especially benefit cognitive function during healthy aging (Fabre et al., 2002). Assuming that monotonous activity is less cognitively stimulating, our measure of variety in exercise may be one way to capture this combination of physical and mental engagement. An effect of nonphysical activities did emerge after covarying head motion; however, this moderated the preceding age–brain pathway and so may reflect a different mechanism to varied physical exercise. More nonphysical activities were associated with stronger brain responses at younger ages but weaker responses at older ages (crossing at ~50 years). Both observations support previous suggestions that physical and nonphysical activities differentially impact late-life cognition, with the latter benefitting more at younger ages (Gow et al., 2017).

In conclusion, we confirm and characterize a neural mechanism that partially explains age-related differences in fluid intelligence, namely, reduced responsiveness of the frontoparietal multiple-demand network to cognitive challenge. Specification of such a neuro-cognitive mechanism may facilitate design of targeted interventions to maintain fluid intelligence into healthy old age. As one example, we identify a widely applicable candidate lifestyle strategy: variety of regular physical activity, which might buffer age-related cognitive decline by decoupling one link in this putative causal pathway.

## References

Alvarez IC, Barbero J, Zofio JL (2017) A panel data toolbox for MATLAB. J Stat Soft 76:1–27.

Angevaren M, Vanhees L, Wendel-Vos W, Verhaar HJ, Aufdemkampe G, Aleman A, Verschuren WM (2007) Intensity, but not duration, of physical activities is related to cognitive function. Eur J Cardiovasc Prev Rehabil 14:825–830.

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38:95–113.

Assem M, Glasser MF, Van Essen DC, Duncan J (2020a) A domain-general cognitive core defined in multimodally parcellated human cortex. Cereb Cortex 30:4361–4380.

Assem M, Blank IA, Mineroff Z, Ademoglu A, Fedorenko E (2020b) Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in working memory and fluid intelligence. Cortex 131:1–16.

Barbey AK (2018) Network neuroscience theory of human intelligence. Trends Cogn Sci 22:8–20.

Barbey AK, Colom R, Paul EJ, Grafman J (2014) Architecture of fluid intelligence and working memory revealed by lesion mapping. Brain Struct Funct 219:485–494.

Barnes JN (2015) Exercise, cognitive function, and aging. Adv Physiol Educ 39:55–62.

Baron RM, Kenny DA (1986) The moderator mediator variable distinction in social psychological-research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51:1173–1182.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188.

Bethlehem RA, Paquola C, Seidlitz J, Ronan L, Bernhardt B, Consortium CC, Tsvetanov KA (2020) Dispersion of functional gradients across the adult lifespan. Neuroimage 222:117299.

Borgeest GS, Henson RN, Shafto M, Samu D, Cam-CAN, Kievit RA (2020) Greater lifestyle engagement is associated with better age-adjusted cognitive abilities. PLoS One 15:e0230077.

Brown BM, Peiffer JJ, Sohrabi HR, Mondal A, Gupta VB, Rainey-Smith SR, Taddei K, Burnham S, Ellis KA, Szoeke C, Masters CL, Ames D, Rowe CC, Martins RN, AIBL Research Group (2012) Intense physical activity is associated with cognitive performance in the elderly. Transl Psychiatry 2: e191.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann NY Acad Sci 1124:1–38.

Cabeza R (2002) Hemispheric asymmetry reduction in older adults: the HAROLD model. Psychol Aging 17:85–100.

Carpenter PA, Just MA, Shell P (1990) What One Intelligence-Test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. Psychol Rev 97:404–431.

Cattell RB (1963) Theory of fluid and crystallized intelligence: a critical experiment. J Educ Psychol 54:1–22.

Cattell RB, Cattell AK (1973) Measuring intelligence with the culture fair tests. Champaign, IL: Institute for Personality and Ability Testing.

Chan D, Shafto M, Kievit R, Matthews F, Spink M, Valenzuela M, Cam-CAN, Henson RN (2018) Lifestyle activities in mid-life contribute to cognitive reserve in late-life, independent of education, occupation, and late-life activities. Neurobiol Aging 70:180–183.

Ciric R, Rosen AF, Erus G, Cieslak M, Adebimpe A, Cook PA, Bassett DS, Davatzikos C, Wolf DH, Satterthwaite TD (2018) Mitigating head motion artifact in functional connectivity MRI. Nat Protoc 13:2801–2826.

Colcombe S, Kramer AF (2003) Fitness effects on the cognitive function of older adults: a meta-analytic study. Psychol Sci 14:125–130.

Cole MW, Yarkoni T, Repovs G, Anticevic A, Braver TS (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. J Neurosci 32:8988–8999.

Colom R, Garcia-Lopez O (2003) Secular gains in fluid intelligence: evidence from the Culture-Fair intelligence test. J Biosoc Sci 35:33–39.

Comon P (1994) Independent component analysis, a new concept. Signal Processing 36:287–314.

Cotman CW, Berchtold NC (2002) Exercise: a behavioral intervention to enhance brain health and plasticity. Trends Neurosci 25:295–301.

Cotman CW, Berchtold NC, Christie LA (2007) Exercise builds brain health: key roles of growth factor cascades and inflammation. Trends Neurosci 30:464–472.

Cusack R, Vicente-Grabovetsky A, Mitchell DJ, Wild C, Auer T, Linke AC, Peelle JE (2015) Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using MATLAB and XML. Front Neuroinform 8:90.

Davis SW, Dennis NA, Daselaar SM, Fleck MS, Cabeza R (2008) Que PASA? The posterior-anterior shift in aging. Cereb Cortex 18:1201–1209.

Dennis NA, Cabeza R (2008) Neuroimaging of healthy cognitive aging. In: The handbook of aging and cognition, 3rd ed, pp 1–54. New York, NY, US: Psychology Press.

Diehl M, Willis SL, Schaie KW (1995) Everyday problem solving in older adults: observational assessment and cognitive correlates. Psychol Aging 10:478–491.

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Modelling 196:483–493.

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.

Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483.

Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, Newell FN, Emslie H (2000) A neural basis for general intelligence. Science 289:457–460.

Duncan J, Chylinski D, Mitchell DJ, Bhandari A (2017) Complexity and compositionality in fluid intelligence. Proc Natl Acad Sci USA 114:5295–5299.

Edwards JR, Lambert LS (2007) Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. Psychol Methods 12:1–22.

Eyler LT, Sherzai A, Kaup AR, Jeste DV (2011) A review of functional brain imaging correlates of successful cognitive aging. Biol Psychiatry 70:115–122.

Fabre C, Chamari K, Mucci P, Masse-Biron J, Prefaut C (2002) Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. Int J Sports Med 23:415–421.

Fiedler K, Schott M, Meiser T (2011) What mediation analysis can (not) do. J Exp Soc Psychol 47:1231–1236.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA 102:9673–9678.

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.

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC (2016) A multi-modal parcellation of human cerebral cortex. Nature 536:171–178.

Gottfredson LS, Deary IJ (2004) Intelligence predicts health and longevity, but why? Curr Dir Psychol Sci 13:1–4.

Gow AJ, Pattie A, Deary IJ (2017) Lifecourse activity participation from early, mid, and later adulthood as determinants of cognitive aging: the Lothian Birth Cohort 1921. J Gerontol B Psychol Sci Soc Sci 72:25–37.

Grady CL (2012) The cognitive neuroscience of ageing. Nat Rev Neurosci 13:491–505.

Grady CL, Springer MV, Hongwanishkul D, McIntosh AR, Winocur G (2006) Age-related changes in brain activity across the adult lifespan. J Cogn Neurosci 18:227–241.

Gray JR, Chabris CF, Braver TS (2003) Neural mechanisms of general fluid intelligence. Nat Neurosci 6:316–322.

Griffith DA, Peres-Neto PR (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology 87:2603–2613.

Hearne LJ, Mattingley JB, Cocchi L (2016) Functional brain networks related to individual differences in human intelligence at rest. Sci Rep 6:32328.

Hedden T, Schultz AP, Rieckmann A, Mormino EC, Johnson KA, Sperling RA, Buckner RL (2016) Multiple brain markers are linked to age-related variation in cognition. Cereb Cortex 26:1388–1400.

Heller R, Golland Y, Malach R, Benjamini Y (2007) Conjunction group analysis: an alternative to mixed/random effect analysis. Neuroimage 37:1178–1185.

Hesterberg TC (2015) What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. Am Stat 69:371–386.

Hilger K, Fukushima M, Sporns O, Fiebach CJ (2020) Temporal stability of functional brain modules associated with human intelligence. Hum Brain Mapp 41:362–372.

Horn JL, Cattell RB (1967) Age differences in fluid and crystallized intelligence. Acta Psychol (Amst) 26:107–129.

Hughes TF, Chang CC, Vander Bilt J, Ganguli M (2010) Engagement in reading and hobbies and risk of incident dementia: the MoVIES project. Am J Alzheimers Dis Other Demen 25:432–438.

Hyvarinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 10:626–634.

Ji JL, Spronk M, Kulkarni K, Repovs G, Anticevic A, Cole MW (2019) Mapping the human brain's cortical-subcortical functional network organization. Neuroimage 185:35–57.

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.

Kievit RA, Davis SW, Mitchell DJ, Taylor JR, Duncan J, Henson RN, Cam-CAN (2014) Distinct aspects of frontal lobe structure mediate age-related differences in fluid intelligence and multitasking. Nat Commun 5:5658.

Kievit RA, Davis SW, Griffiths J, Correia MM, Cam-CAN, Henson RN (2016) A watershed model of individual differences in fluid intelligence. Neuropsychologia 91:186–198.

Kobilo T, Liu QR, Gandhi K, Mughal M, Shaham Y, van Praag H (2011) Running is the neurogenic and neurotrophic stimulus in environmental enrichment. Learn Mem 18:605–609.

Kuiper JS, Zuidersma M, Oude Voshaar RC, Zuidema SU, van den Heuvel ER, Stolk RP, Smidt N (2015) Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 22:39–57.

Lee KH, Choi YY, Gray JR, Cho SH, Chae JH, Lee S, Kim K (2006) Neural correlates of superior intelligence: stronger recruitment of posterior parietal cortex. Neuroimage 29:578–586.

Liu-Ambrose T, Barha CK, Best JR (2018) Physical activity for brain health in older adults. Appl Physiol Nutr Metab 43:1105–1112.

MacKinnon DP, Pirlott AG (2015) Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. Pers Soc Psychol Rev 19:30–43.

MacKinnon DP, Krull JL, Lockwood CM (2000) Equivalence of the mediation, confounding and suppression effect. Prev Sci 1:173–181.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V (2002) A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 7:83–104.

MacKinnon DP, Fairchild AJ, Fritz MS (2007) Mediation analysis. Annu Rev Psychol 58:593–614.

Mauro R (1990) Understanding LOVE (Left out Variables Error): a method for estimating the effects of omitted variables. Psychol Bull 108:314–329.

Maxwell SE, Cole DA (2007) Bias in cross-sectional analyses of longitudinal mediation. Psychol Methods 12:23–44.

Momi D, Neri F, Coiro G, Smeralda C, Veniero D, Sprugnoli G, Rossi A, Pascual-Leone A, Rossi S, Santarnecchi E (2020) Cognitive enhancement via network-targeted cortico-cortical associative brain stimulation. Cereb Cortex 30:1516–1527.

Naghavi HR, Nyberg L (2005) Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration? Conscious Cogn 14:390–425.

Neubauer AC, Fink A (2009) Intelligence and neural efficiency. Neurosci Biobehav Rev 33:1004–1023.

Oltmanns J, Godde B, Winneke AH, Richter G, Niemann C, Voelcker-Rehage C, Schomann K, Staudinger UM (2017) Don't lose your brain at work: the role of recurrent novelty at work in cognitive and brain aging. Front Psychol 8:117.

Park DC, Reuter-Lorenz P (2009) The adaptive brain: aging and neurocognitive scaffolding. Annu Rev Psychol 60:173–196.

Parkes L, Fulcher B, Yucel M, Fornito A (2018) An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. Neuroimage 171:415–436.

Parkin AJ, Java RI (1999) Deterioration of frontal lobe function in normal aging: influences of fluid intelligence versus perceptual speed. Neuropsychology 13:539–545.

Pernet C (2021) SPM Utility Plus 1.0.0. Available at https://github.com/CPernet/spmup/.

Phillips LH, Della Sala S (1998) Aging, intelligence, and anatomical segregation in the frontal lobes. Learn Individ Diff 10:217–243.

Pieters R (2017) Meaningful mediation analysis: plausible causal inference and informative communication. J Consumer Res 44:692–716.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154.

Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84:320–341.

Prabhakaran V, Smith JA, Desmond JE, Glover GH, Gabrieli JD (1997) Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. Cogn Psychol 33:43–63.

Ramsey JB (1969) Tests for specification errors in classical linear least-squares regression analysis. J R Stat Soc B Stat Methodol 31:350–371.

Raz N, Lindenberger U (2011) Only time will tell: cross-sectional studies offer no solution to the age-brain-cognition triangle: comment on Salthouse (2011). Psychol Bull 137:790–795.

Richardson MT, Leon AS, Jacobs DR Jr, Ainsworth BE, Serfass R (1994) Comprehensive evaluation of the Minnesota Leisure Time Physical Activity Questionnaire. J Clin Epidemiol 47:271–281.

Salthouse TA (2011a) Neuroanatomical substrates of age-related cognitive decline. Psychol Bull 137:753–784.

Salthouse TA (2011b) All data collection and analysis methods have limitations: reply to Rabbitt (2011) and Raz and Lindenberger (2011). Psychol Bull 137:796–799.

Salthouse TA, Atkinson TM, Berish DE (2003) Executive functioning as a potential mediator of age-related cognitive decline in normal adults. J Exp Psychol Gen 132:566–594.

Samu D, Campbell KL, Tsvetanov KA, Shafto MA, Tyler LK, Cam-CAN (2017) Preserved cognitive functions with age are determined by domain-dependent shifts in network responsivity. Nat Commun 8:14743.

Santarnecchi E, Emmendorfer A, Pascual-Leone A (2017a) Dissecting the parieto-frontal correlates of fluid intelligence: a comprehensive ALE meta-analysis study. Intelligence 63:9–28.

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

Shafto MA, Tyler LK, Dixon M, Taylor JR, Rowe JB, Cusack R, Calder AJ, Marslen-Wilson WD, Duncan J, Dalgleish T, Henson RN, Brayne C, Matthews FE, Cam-CAN (2014) The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. BMC Neurol 14:204.

Shimizu S, Hoyer PO, Hyvarinen A, Kerminen A (2006) A linear non-Gaussian acyclic model for causal discovery. J Mach Learn Res 7:2003–2030.

Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE (2014) Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Hum Brain Mapp 35:1981–1996.

Siegel JS, Mitra A, Laumann TO, Seitzman BA, Raichle M, Corbetta M, Snyder AZ (2017) Data quality influences observed links between functional connectivity and behavior. Cereb Cortex 27:4492–4502.

Smith PJ, Blumenthal JA, Hoffman BM, Cooper H, Strauman TA, Welsh-Bohmer K, Browndyke JN, Sherwood A (2010) Aerobic exercise and neurocognitive performance: a meta-analytic review of randomized controlled trials. Psychosom Med 72:239–252.

Smith V, Pinasco C, Achterberg J, Mitchell DJ, Das T, Roca M, Duncan J (2022) Fluid intelligence and naturalistic task impairments after focal brain lesions. Cortex 146:106–115.

Spreng RN, Stevens WD, Viviano JD, Schacter DL (2016) Attenuated anti-correlation between the default and dorsal attention networks with aging: evidence from task and rest. Neurobiol Aging 45:149–160.

Thoemmes F (2015) Reversing arrows in mediation models does not distinguish plausible models. Basic Appl Soc Psychol 37:226–234.

Tschentscher N, Mitchell D, Duncan J (2017) Fluid intelligence predicts novel rule implementation in a distributed frontoparietal control network. J Neurosci 37:4841–4847.

Tsvetanov KA, Henson RN, Tyler LK, Razi A, Geerligs L, Ham TE, Rowe JB, Cam-CAN (2016) Extrinsic and intrinsic brain network connectivity maintains cognition across the lifespan despite accelerated decay of regional brain activation. J Neurosci 36:3115–3126.

Tsvetanov KA, Henson RN, Rowe JB (2021) Separating vascular and neuronal effects of age on fMRI BOLD signals. Philos Trans R Soc Lond B Biol Sci 376:20190631.

Turner GR, Spreng RN (2015) Prefrontal engagement and reduced default network suppression co-occur and are dynamically coupled in older adults: the default-executive coupling hypothesis of aging. J Cogn Neurosci 27:2462–2476.

Valente MJ, Pelham WE, Smyth H, MacKinnon DP (2017) Confounding in statistical mediation analysis: what it is and how to address it. J Couns Psychol 64:659–671.

Valenzuela MJ, Sachdev P (2007) Assessment of complex mental activity across the lifespan: development of the Lifetime of Experiences Questionnaire (LEQ). Psychol Med 37:1015–1025.

Vaupel JW (2010) Biodemography of human ageing. Nature 464:536–542.

Vos de Wael R, Benkarim O, Paquola C, Lariviere S, Royer J, Tavakol S, Xu T, Hong SJ, Langs G, Valk S, Misic B, Milham M, Margulies D, Smallwood J, Bernhardt BC (2020) BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. Commun Biol 3:103.

Voss MW, Chaddock L, Kim JS, Vanpatter M, Pontifex MB, Raine LB, Cohen NJ, Hillman CH, Kramer AF (2011) Aerobic fitness is associated with greater efficiency of the network underlying cognitive control in preadolescent children. Neuroscience 199:166–176.

Wagner HH, Dray S (2015) Generating spatially constrained null models for irregularly spaced data using Moran spectral randomization methods. Methods Ecol Evol 6:1169–1178.

Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE (2002) Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. Int J Epidemiol 31:168–174.

West KL, Zuppichini MD, Turner MP, Sivakolundu DK, Zhao Y, Abdelkarim D, Spence JS, Rypma B (2019) BOLD hemodynamic response function changes significantly with healthy aging. Neuroimage 188:198–207.

Wooldridge JM (2012) Introductory econometrics: a modern approach, Ed 5. Mason, OH: South-Western Cengage Learning.

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.

Woolgar A, Bor D, Duncan J (2013) Global increase in task-related fronto-parietal activity after focal frontal lobe lesion. J Cogn Neurosci 25:1542–1552.

Woolgar A, Duncan J, Manes F, Fedorenko E (2018) The multiple-demand system but not the language system supports fluid intelligence. Nat Hum Behav 2:200–204.

Wu S, Tyler LK, Henson RN, Rowe JB, Cam-CAN, Tsvetanov KA (2021) Cerebral blood flow predicts multiple demand network activity and fluid intelligence across the lifespan. bioRxiv 468042. https://doi.org/10.1101/2021.11.10.468042.

Yates LA, Ziser S, Spector A, Orrell M (2016) Cognitive leisure activities and future risk of cognitive impairment and dementia: systematic review and meta-analysis. Int Psychogeriatr 28:1791–1806.