bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Neural Activity in the Fronto-Parietal Multiple Demand Network Robustly Predicts Individual Differences In Working Memory And Fluid Intelligence

Authors: Moataz Assem[1], Idan Blank[2], Zach Mineroff[2], Ahmet Ademoglu[3], and Evelina Fedorenko[4,5]

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, CB2 7EF, UK
[2] Brain & Cognitive Sciences Department, McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139
[3] Institute of Biomedical Engineering, Bogazici University, Istanbul, 34684, Turkey
[4] Department of Psychiatry, Harvard Medical School, Boston, MA 02115
[5] Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

A bilateral network of frontal and parietal domain-general brain regions – the multiple demand (MD) system (Duncan, 2010, 2013) – has been linked to our ability to engage in goal-directed behaviors, solve novel problems, and acquire new skills. Damage to this network leads to deficits in executive abilities and lower fluid intelligence (e.g., Woolgar et al., 2010), and aberrant functioning of this network has been reported in a variety of neurological and psychiatric disorders (e.g., Cole et al., 2014). However, prior attempts to link MD activity to behavior in neurotypical adults have yielded contradictory findings. In a large-scale fMRI study (n=140), we found that stronger up-regulation of the MD activity with increases in task difficulty, as indexed by larger differences between responses to the harder vs. easier condition, was associated with better behavioral performance on the working memory task performed in the scanner, and overall higher fluid intelligence measured independently. We further demonstrate how small samples, like those used in some earlier studies, could have led to the opposite patterns of results. Finally, the relationship we observed between MD activity and behavior was selective: neural activity in another large-scale network (the fronto-temporal language network) did not reliably predict working memory performance or fluid intelligence. Our study thus paves the way for using individual fMRI measures to link genetic and behavioral variation in executive functions in healthy and patient populations.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Significance statement

A distributed frontoparietal Multiple Demand (MD) network has long been implicated in intelligent behavior, and its damage has been associated with lower intelligence and difficulties in problem solving. Yet prior studies have not yielded a clear answer on how individual differences in MD activity translate into differences in behavior. Across a large number of participants, we find that stronger up-regulation of the MD network's activity robustly and selectively predicts higher intelligence scores and better task performance. We demonstrate how small samples, along with other shortcomings, could have led to contradictory results in previous studies. Thus, MD activity up-regulation can serve as a robust individual measure to link genetic and behavioral variation in executive functions in healthy and patient populations.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

A bilateral network of frontal and parietal domain-general brain regions – the multiple demand (MD) system (Duncan, 2010, 2013) – has been has been linked to our ability to engage in goal-directed behaviors, solve novel problems, and acquire new skills (Duncan and Owen, 2000; Fedorenko et al., 2013; Hugdahl et al., 2015). Damage to this network as a result of stroke, degeneration or head injury leads to poorer executive abilities (attention, working memory, and inhibitory control) and lower fluid intelligence (Glascher et al., 2010; Roca et al., 2010; Woolgar et al., 2010). Furthermore, aberrant functioning of this network, as measured with fMRI, has been reported in a variety of cognitive and psychiatric disorders (Cole et al., 2014).

Given the fundamental importance of flexible thought and behavior for humans, a deeper understanding of the individual differences in this network's activity would have critical implications for medicine, providing an *intermediate link* between behavioral and genetic variability, as well as yield a deeper understanding of basic cognitive and neural architecture (Braver et al., 2010; Dubois and Adolphs, 2016). Critically though, the potential usefulness of neural measures of MD activity is contingent on our ability to link such activity to behavior. Yet previous attempts to do so have yielded contradictory findings: some studies have reported stronger MD responses associated with worse behavioral performance and lower IQ (Haier et al., 1988; Rypma and Esposito, 2000; Rypma et al., 2006), others – with better performance and higher IQ (Gray et al., 2003; Lee et al., 2006; Tschentscher et al., 2017).

These discrepancies may be due to a number of shortcomings that characterize many prior studies that have probed the relationship between MD neural responses and behavioral measures. First, many studies have used small numbers of participants and/or transformed continuous behavioral measures into categorical variables (e.g., high- vs. low-performing

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

participants), both of which can produce inflated or spurious relationships (Rypma and Esposito, 2000; Wager et al., 2005; Lee et al., 2006; Rypma et al., 2006; Tschentscher et al., 2017). Second, some studies have used BOLD estimates based on contrasts of task relative to fixation, which may fail to isolate MD activity from general state (e.g., motivation, sleepiness, caffeine intake) or trait (e.g., brain vascularization) variables (Rypma and Esposito, 2000; Gray et al., 2003; Rypma et al., 2006). And third, most studies have failed to take into consideration inter-individual variability in the precise locations of MD regions, which leads to losses in sensitivity and functional resolution (Nieto-Castañón and Fedorenko, 2012). This latter problem is compounded by the proximity of MD regions to language-selective regions (Fedorenko et al., 2012), which are functionally distinct, showing no response to any demanding task other than language processing (Fedorenko et al., 2011; Monti et al., 2012). In addition to these limitations, prior studies have failed to establish, or even assess, the *selectivity* of the relationship between MD activity specifically (cf. any other neural measure) and behavior (Gray et al., 2003; Rypma et al., 2006; Dubois and Adolphs, 2016).

To circumvent these limitations and rigorously test the relationship between MD activity and behavior, we conducted a large-scale fMRI study, where participants (n=140) performed a spatial working memory (WM) task that included a harder and an easier condition (**Fig. 1**). We then examined the relationship between the size of the Hard>Easy (H>E) BOLD effect across the MD network (defined functionally in each participant individually (Fedorenko et al., 2013); **Fig. 3**), and a) behavioral performance on the task (including in an independent run of data), as well as b) a measure of fluid intelligence (in a subset of participants, n=63). We further evaluated the selectivity of this brain-behavior relationship by examining neural activity in another large-

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

scale brain network: the left fronto-temporal language network (Binder et al., 1997; Fedorenko et al., 2010).

Materials and Methods

Participants

140 right-handed participants (age 22.8 ± 5.4, 47 males) with normal or corrected-to-normal vision, students at Massachusetts Institute of Technology (MIT) and members of the surrounding community, participated for payment. All participants gave informed consent in accordance with the requirements of the Committee on the Use of Humans as Experimental Subjects at MIT.

Experimental Design and Statistical Analysis

Participants performed two tasks in the scanner (the critical spatial working memory task, and a language processing task used here to assess the selectivity of the relationship between MD activity and behavior). A subset of participants performed a behavioral IQ test after the scanning session. Neural measures were statistically estimated using the standard General Linear Model (GLM) in SPM5 (see the *fMRI data preprocessing and first-level analysis* and *MD/Language fROIs definition and response estimation* sections for details). The following statistical tests were used: A two-sample paired *t*-test was used to compare behavioral performance between the easy and hard conditions of the spatial working memory (WM) task. A one-sample *t*-test was used to test the reliability of the H>E effect size in the spatial WM task for each ROI separately, and across ROIs, across participants. Pearson and Spearman correlations were used: 1) to assess the stability of behavioral or neural measures across runs (**Figures 2, 3C-E**; for the results reported in **Figure 3D**, a Bonferroni correction for the number of ROIs, i.e., 18, was applied), 2) to test

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the relationship between MD/language networks activities and behavioral measures **(Figures 4, 5B)**, and 3) to test the relationship between the BOLD H>E-behavior in different sample sizes (**Figure 6**). A partial correlation test was used to assess the unique variance in IQ scores that each network can predict after controlling for the other network's responses (see *The Selectivity of the MD BOLD predictions* section).

Experimental Paradigms

Participants performed a spatial working memory task in a blocked design (**Fig. 1**). Each trial lasted 8 seconds: within a 3x4 grid, a set of locations lit up in blue, one at a time for a total of 4 (easy condition) or two at a time for a total of 8 (hard condition). Participants were asked to keep track of the locations. At the end of each trial, they were shown two grids with some locations lit up and asked to choose the grid that showed the correct locations by pressing one of two buttons. They received feedback on whether they answered correctly. Each participant performed two runs, with each run consisting of four 32-second easy condition blocks, four 32-second hard blocks, and four 16-second fixation blocks. Condition order was counterbalanced across runs.

In addition to the spatial working memory task, all participants performed a language localizer task (Fedorenko et al., 2010), used here to test the selectivity of the relationship between MD network's activity and behavior. The majority (n=113, 81%) passively read sentences and lists of pronounceable nonwords in a blocked design. The Sentences>Nonwords contrast targets brain regions sensitive to high-level linguistic processing (Fedorenko et al., 2010, 2011). Each trial started with 100ms pre-trial fixation, followed by a 12-word-long sentence or a list of 12 nonwords presented on the screen one word/nonword at a time at the rate of 450ms per word/nonword. Then, a line drawing of a hand pressing a button appeared for 400ms, and

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

participants were instructed to press a button whenever they saw the icon, and finally a blank screen was shown for 100ms, for a total trial duration of 6s. The button-press task was included to help participants stay awake and focused. Each block consisted of 3 trials and lasted 18s. Each run consisted of sixteen experimental blocks (eight per condition), and five fixation blocks (14s each), for a total duration of 358s (5min 58s). Each participant performed two runs. Condition order was counterbalanced across runs. The remaining 27 participants performed similar versions of the language localizer with minor differences in the timing and procedure. (We have previously established that the localizer contrast is robust to such differences (e.g., Fedorenko et al., 2010; Scott et al., 2016).

Finally, most participants completed one or more additional experiments for unrelated studies. The entire scanning session lasted approximately 2 hours.

A subset of 63 participants performed the non-verbal component of KBIT (Kaufman and Kaufman, 2013) after the scanning session. The test consists of 46 items (of increasing difficulty) and includes both meaningful stimuli (people and objects) and abstract ones (designs and symbols). All items require understanding the relationships among the stimuli and have a multiple-choice format, requiring the participant to select the correct response. If a participant answers 4 questions in a row incorrectly, the test is terminated, and the remaining items are marked as incorrect. The test is scored following the formal guidelines to calculate each participant's IQ score.

fMRI data acquisition

Structural and functional data were collected on the whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

128 axial slices with 1mm isotropic voxels (TR=2530ms, TE=3.48ms). Functional, blood oxygenation level dependent (BOLD) data were acquired using an EPI sequence (with a 90° flip angle and using GRAPPA with an acceleration factor of 2), with the following acquisition parameters: thirty-one 4mm thick near-axial slices, acquired in an interleaved order with a 10% distance factor; 2.1mm x 2.1mm in-plane resolution; field of view of 200mm in the phase encoding anterior to posterior (A > P) direction; matrix size of 96mm x 96mm; TR of 2000ms; and TE of 30ms. Prospective acquisition correction (Thesen et al., 2000) was used to adjust the positions of the gradients based on the participant's motion one TR back. The first 10s of each run were excluded to allow for steady-state magnetization.

fMRI data preprocessing and first-level analysis

fMRI data were analyzed using SPM5 and custom MATLAB scripts. Each subject's data were motion corrected and then normalized into a common brain space (the Montreal Neurological Institute (MNI) template) and resampled into 2mm isotropic voxels. The data were then smoothed with a 4mm Gaussian filter and high-pass filtered (at 200s). The task effects in both the spatial WM task and in the language localizer task were estimated using a General Linear Model (GLM) in which each experimental condition was modeled with a boxcar function (corresponding to a block) convolved with the canonical hemodynamic response function (HRF).

MD fROIs definition and response estimation

To define the MD and language (see below) functional regions of interest (fROIs), we used the Group-constrained Subject-Specific (GSS) approach (Fedorenko et al., 2010). In particular, fROIs were constrained to fall within a set of "parcels", areas that corresponded to the expected gross locations of activations for the relevant contrast. For the MD fROIs, following Fedorenko et al. (Fedorenko et al., 2013) and Blank et al. (Blank et al., 2014), we used eighteen

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

anatomical parcels (Tzourio-Mazoyer et al., 2002) across the two hemispheres. These parcels covered the portions of the frontal and parietal cortices where MD activity has been previously reported, including bilateral opercular IFG (L/R IFGop), MFG (L/R MFG), orbital MFG (L/R MFGorb), insular cortex (L/R Insula), precentral gyrus (L/R PrecG), supplementary and presupplementary motor areas (L/R SMA), inferior parietal cortex (L/R ParInf), superior parietal cortex (L/R ParSup), and anterior cingulate cortex (L/R ACC) (**Figure 3A)**. Within each parcel, we selected the top 10% of most responsive voxels in each individual participant based on the *t*-values for the Hard>Easy spatial WM contrast. This top n% approach ensures that a fROI can be defined in every participant, and that the fROI sizes are identical across participants.

To estimate the fROIs' responses to the Hard and Easy conditions, we used an across-run cross-validation procedure (Nieto-Castañón and Fedorenko, 2012) to ensure that the data used to identify the ROIs are independent from the data used to estimate their response magnitudes (Kriegeskorte et al., 2009). To do this, the first run was used to define the fROIs and the second run to estimate the responses. This procedure was then repeated using the second run to define the fROIs and the first run to estimate the responses. Finally, the responses were averaged across the left-out runs to derive a single response magnitude estimate for each participant in each fROI for each condition. Finally, these estimates were averaged across the 18 fROIs of the MD network to derive one value per condition for each participant. (An alternative approach would have been to examine fROI *volumes* – number of MD-responsive voxels at a fixed significance threshold – instead of effect sizes. However, first, effect sizes and region volumes are generally strongly correlated; and second, effect sizes tend to be more stable within participants than region volumes (Mahowald and Fedorenko, 2016).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Language fROIs definition and response estimation

To define the language fROIs, we used a set of functional parcels that were generated based on a group-level representation of data from a large set of participants. In particular, we used six parcels derived from a group-level representation of data for the Sentences>Nonwords contrast in 220 participants. These parcels included three regions in the left frontal cortex: two located in the inferior frontal gyrus, and one located in the middle frontal gyrus; and three regions in the left temporal and parietal cortices spanning the entire extent of the lateral temporal lobe and going posteriorly to the angular gyrus. Within each parcel, we selected the top 10% of most responsive voxels in each individual participant based on the *t*-values for the Sentences>Nonwords contrast. To estimate the fROIs' responses to the Sentences and Nonwords conditions, we used an across-run cross-validation, as for the MD fROIs.

Results

Behavioral measures

Behavioral performance on the spatial WM task was as expected: individuals were more accurate and faster on the easy trials (accuracy=92.86%; RT=1.19s) than the hard trials (accuracy=78.11%, $t_{(139)}$=-18.64, p<0.0001; RT=1.51s, $t_{(139)}$=23.45, p<0.0001). Behavioral measures were stable within individuals across runs for overall (averaging across the Hard and Easy conditions) accuracies (r=0.59, p<0.001) and RTs (r=0.83, p<0.001) (**Figure 2**), which suggests they can be meaningfully related to neural measures. Further, higher IQ scores, as measured by KBIT, correlated with overall accuracies (r=0.29, p=0.018) and RTs (r=-0.33, p=0.008), but not with the difference scores (accuracies H>E (r=-0.06, p=0.63); RTs H>E (r=0.08, p=0.53)).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In the critical brain-behavior analyses below, we used overall accuracies and RTs rather than the Hard>Easy measures, because the former i) are more stable within individuals (r=0.59 vs. r=0.28 for the accuracies, and r=0.83 vs. r=0.41 for the RTs) (**Figure 2**), ii) are more intuitively interpretable, and iii) correlate with IQ (the Hard>Easy measures do not). Furthermore, the Hard>Easy measures contain a non-linearity, such that smaller between-condition differences can be observed in both high performers (when performance is closer to ceiling) and low performers (when performance is closer to chance).

MD BOLD measure

As expected (Fedorenko et al., 2013), each of the eighteen MD fROIs individually, as well as the average across fROIs, showed a highly reliable Hard>Easy effect across participants ($ts_{(139)}$>11.5, *p*s<0.00001). MD Hard>Easy neural responses were also stable across runs for each MD ROI individually (*r*s=0.60–0.80, *p*s<0.0001) and collapsing across ROIs (r=0.73, p<0.0001) (**Figures 3C,E**). We used the Hard>Easy effect size for our neural measure (cf. task>fixation) to factor out variability due to state/trait differences and thus to hone in on the variability in the MD system's activity given its functional signature of sensitivity to difficulty (Duncan and Owen, 2000; Fedorenko et al., 2013; Hugdahl et al., 2015). For each participant, we averaged the size of the Hard>Easy effect across the 18 MD fROIs to derive a single measure because the MD network has been shown to be a highly functionally integrated system: the MD regions' time-courses show strong correlations during both rest and task performance (Dosenbach et al., 2007; Seeley et al., 2007; Hampshire et al., 2012; Blank et al., 2014). In line with these prior findings, the Hard>Easy effect sizes were strongly correlated across the 18 regions in the current dataset (**Figure 3D**; see also Mineroff et al., 2017).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

MD BOLD predicts task performance and fluid intelligence

For each participant, we used two behavioral measures from the spatial WM task (overall accuracies and RTs), and one neural measure (the size of the Hard>Easy effect averaged across the 18 MD fROIs). The critical analyses revealed that larger MD Hard>Easy responses were associated with better behavioral performance as reflected in both higher accuracies (r=0.33, p=0.0001) and fasters RTs (r=-0.23, p=0.0057; **Figure 4A**). To further test the predictive power of the MD Hard>Easy index, we cross-compared BOLD-behavior relationships across runs (Dubois and Adolphs, 2016) and found that the MD Hard>Easy effect size in Run 2 successfully predicted both accuracies (r=0.29, p=0.0004) and RTs (r=-0.21, p=0.012) in Run 1, and MD H>E effects size in Run 1 predicted accuracies (r=0.25, p=0.0027) and RTs (r=-0.21, p=0.013) in Run 2 (**Figure 4B**).

Next, to test the generalizability of the relationship between MD activity and behavior, we asked whether the Hard>Easy MD index could explain variance in fluid intelligence, as measured with the Kaufman Brief Intelligence Test (KBIT) (Kaufman and Kaufman, 2013) in a subset of participants (n=63). Indeed, larger MD Hard>Easy responses were associated with higher intelligence quotient (IQ) scores (r=0.36, p=0.0035) (**Figure 4A**). It is worth noting that the strength of the BOLD responses to the Hard or Easy condition relative to the fixation baseline did not correlate with IQ (H>fix: r=0.17, p=0.2; E>fix: r=0.01, p=0.9).

The selectivity of the MD BOLD predictions

To test whether the brain-behavior relationship we observed is selective to the MD network, we considered another large-scale neural network: the fronto-temporal language-selective network in the left hemisphere (Fedorenko et al., 2011) (**Figure 5A**). To the best of our

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

knowledge, none of the prior studies that have investigated the relationship between MD network's activity and behavior have assessed the spatial selectivity of the observed relationship.

Like the MD Hard>Easy responses, the size of the Sentences>Nonwords effect, used to define the language regions (Fedorenko et al., 2010), was highly stable across runs for each language ROI individually and collapsing across ROIs (r=0.82, p<0.0001), in line with prior work (Mahowald and Fedorenko, 2016). Critically, the Sentences>Nonwords effect only weakly correlated with the spatial WM task accuracies (r=0.19, p=0.02, cf. r=0.33, p<0.001 for the MD Hard>Easy effect), and not at all with RTs (r=-0.10, p=0.25) (**Figure 5B)**. In an analysis parallel to the one we performed on the MD Hard>Easy effect, we tested the predictive power of the Sentences>Nonwords index by cross-comparing BOLD-behavior relationships across runs. The Sentences>Nonwords effect size in Run 2 showed a non-significant relationship with accuracies in Run 1 (r=0.15, p=0.068) and did not predict RTs (r=-0.11, p=0.21). Similarly, the Sentences>Nonwords effect size in Run 1 did not predict accuracies (r=0.15, p=0.074) or RTs in Run 2 (r=-0.06, p=0.47).

We also found a weak and non-significant relationship between the Sentences>Nonwords effect size and IQ scores (r=0.22, p=0.090) (**Figure 5B**). Importantly, replicating (Mineroff et al., 2017), the MD Hard>Easy effect sizes and the Sentences>Nonwords effect sizes were only weakly and non-significantly correlated (r=0.12, p=0.14), suggesting that neural activity in the two networks explain largely non-overlapping variance in the IQ scores. Indeed, even after controlling for Sentences>Nonwords responses, the MD Hard>Easy effect sizes still significantly predicted IQ scores (r_p=0.37, p=0.0035). Similarly, the relationship between the Sentences>Nonwords effect sizes and IQ was not affected by controlling for MD Hard>Easy responses (r_p=0.22, p=0.086).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

Across a large set of participants, we observed a robust relationship between neural activity in the domain-general fronto-parietal MD network and behavioral performance on a working memory (WM) task performed in the scanner, as well as an independent measure of fluid intelligence. A stronger up-regulation of the MD activity with increases in task difficulty (as indexed by larger Hard>Easy effect sizes) – a functional signature of this network (Duncan and Owen, 2000; Fedorenko et al., 2013) – was associated with higher accuracies, faster RTs, and overall higher intelligence. This relationship was selective to the MD network: neural activity in another large-scale network important for high-level cognition – the fronto-temporal language network – did not reliably predict WM performance or IQ scores, although larger Sentences>Nonwords effect sizes did exhibit weak associations with task accuracies and IQ scores.

A number of earlier studies have investigated the relationship between neural activity in the MD network and behavioral task performance and/or general fluid intelligence. Yet a clear answer has failed to emerge. In particular, some studies have reported increases in activation with better performance and higher intelligence (e.g., Gray et al., 2003; Lee et al., 2006; Tschentscher et al., 2017). But others have observed the opposite pattern of results: lower neural activity for better performers (e.g., Haier et al., 1988; Reuter-Lorenz et al., 2000; Rypma and D'Esposito, 2000; Rypma et al., 2006). As noted in the Introduction, these studies have relied on small sample sizes (range 6-16). To illustrate how small samples could produce misleading results (Gelman and Carlin, 2014), we performed an additional analysis where we examined the relationship between the Hard>Easy effect size (as well as Hard>Fixation, and Easy>Fixation effect sizes because those are the measures used in some prior studies) and the overall accuracies

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

on the spatial WM task in small samples. Different sample sizes (ranging from 10 to 130, in increments of 5) were randomly selected from our larger set of 140 participants. The correlation with totally accuracy was calculated for each sample size. This process was repeated 1000 times to produce 1000 correlations per sample size.

These correlations were then examined for their sign, size, and significance. The results (**Figure 6**) clearly show that in small samples, like those used in some of the earlier studies, it is possible to observe a significant correlation of the opposite sign to that observed in a larger population. For example, using a sample size of 20, 5 of the 368 significant (p<0.05) correlations (1.4%) had this property when considering the relationship between the Hard>Easy effect size and accuracies. This problem is exacerbated when using task>fixation contrasts (3.5% for Hard>fix, 15.6% for Easy>fix) (**Figure 6**).

In addition to the issue of small samples, some prior studies have used an overall measure of activation (e.g., response to some task vs. fixation) as their neural measure of interest instead of using the functional signature of the MD network, i.e., the difference between the response to a harder vs. an easier condition of a task (e.g., Duncan and Owen, 2000; Fedorenko et al., 2013; Hugdahl et al., 2015). Measures of overall activity relative to a low-level baseline incorporate variability related to general state (e.g., motivation or caffeine intake) or trait (e.g., brain vascularization) characteristics, and are thus necessarily noisier. Indeed, in our dataset, the latter measures are much less stable within individuals across runs compared to the measures that rely on the Hard vs. Easy contrast.

Finally, most prior studies have relied on the assumption of functional-anatomic correspondence across individuals in a common brain space (e.g., the MNI space), i.e., treating each voxel as functionally the same across participants. This assumption is problematic,

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

however, given the well-known inter-individual variability in the human association cortices (Frost and Goebel, 2012; Tahmasebi et al., 2012). Thus, prior studies risked "diluting" the neural measures by picking up signals from the nearby language regions, which have the opposite functional profiles (Fedorenko et al., 2012). We circumvented the issue of inter-individual variability in the precise locations of the MD regions by defining those regions functionally in each individual brain (e.g., Fedorenko et al., 2013).

It is worth noting that some have tried to explain the discrepancies in the literature by alluding to differences in the age of participants across studies (Reuter-Lorenz et al., 2000; Rypma and Esposito, 2000), with the hypothesis that the relationship between MD activity level and behavior may vary between younger and older participants. The age range in our sample (25th-75th percentile = 20-23) is not sufficient to evaluate this hypothesis rigorously. Studies with large samples of participants of varying ages will be needed to test this idea. That said, many relevant prior studies a) used small sample sizes, b) used overall measures of neural activity, and c) did not take into account inter-individual variability in MD regions, which may be especially important given the increased variability in the functional architecture of older adults (Geerligs et al., 2017). As a result, it is not currently clear whether there exists any support for the hypothesis about age-related changes in how neural measures of MD activity relate to behavior.

To briefly return to the issue of selectivity of the relationship between MD system's activity and task performance / intelligence: in our analysis of a control brain network (the fronto-temporal language network; e.g., Fedorenko et al., 2010), we observed a weak relationship between responses in that network and spatial WM task accuracies and intelligence. Importantly, in line with prior work, the language network is functionally distinct from the MD network (e.g., Fedorenko et al., 2013; Blank et al., 2014; Mineroff et al., 2017), and so the two

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

networks explain independent variance in accuracies and IQ scores, as confirmed by partial correlation analyses. The relationship between activity in the language system and fluid intelligence is surprising in light of the prior literature: in particular, language brain regions do not respond to demanding executive tasks as assessed in fMRI (e.g., Fedorenko et al., 2011; Monti et al., 2012), and patients with even severe damage to the language network can retain high intelligence and the ability to perform challenging cognitive tasks, like arithmetic (Varley et al., 2005) and causal reasoning (e.g., Varley and Siegal, 2000; see Fedorenko and Varley, 2016, for a review). It is also worth keeping in mind that the correlations we observed in our dataset are relatively small in size (substantially lower than those between the behavioral measures and the MD network's neural responses); further, we did not observe any relationship with the RT measure, and only the correlation with task accuracies, not IQ scores, reached significance at $p<0.05$ level. So, this observed suggestive relationship bears replication in another large set of participants before theorizing too much about its interpretation and significance.

To conclude, the magnitude of the Hard>Easy effect for a working memory task in the MD network appears to be a stable measure, within an individual, that can be used to further probe variability in executive abilities between individuals both in the typical population and among individuals with cognitive and psychiatric disorders, many of which are characterized by decreases in fluid intelligence. Better task performers, and individuals with higher intelligence showed a larger Hard>Easy response. Thus, this marker can serve as a promising neural bridge (Braver et al., 2010) between behavioral variability and genetic variability associated with differences in fluid intelligence (Plomin and Spinath, 2004; Deary et al., 2006).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

Binder JR, Frost J a, Hammeke T a, Cox RW, Rao SM, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. J Neurosci 17:353–362.

Blank IA, Kanwisher N, Fedorenko E (2014) A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. J Neurophysiol:1105–1118 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24872535.

Braver TS, Cole MW, Yarkoni T (2010) Vive les differences! Individual variation in neural mechanisms of executive control. Curr Opin Neurobiol 20:242–250.

Cole MW, Repov G, Anticevic A (2014) The Frontoparietal Control System: A Central Role in Mental Health. Neurosci 20:652–664 Available at: http://nro.sagepub.com/cgi/doi/10.1177/1073858414525995.

Deary IJ, Spinath FM, Bates TC (2006) Genetics of intelligence. Eur J Hum Genet 14:690–700.

Dosenbach NUF, Fair D a, Miezin FM, Cohen AL, Wenger KK, Dosenbach R a T, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104:11073–11078 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17576922.

Dubois J, Adolphs R (2016) Building a Science of Individual Differences from fMRI. Trends Cogn Sci 20:425–443 Available at: http://linkinghub.elsevier.com/retrieve/pii/S1364661316300079.

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 (2013) The Structure of Cognition: Attentional Episodes in Mind and Brain. Neuron 80:35–50 Available at: http://dx.doi.org/10.1016/j.neuron.2013.09.015.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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

Fedorenko E, Behr MK, Kanwisher N (2011) Functional specificity for high-level linguistic processing in the human brain. Proc Natl Acad Sci U S A 108:16428–16433.

Fedorenko E, Duncan J, Kanwisher N (2012) Language-selective and domain-general regions lie side by side within Broca's area. Curr Biol 22:2059–2062 Available at: http://dx.doi.org/10.1016/j.cub.2012.09.011.

Fedorenko E, Duncan J, Kanwisher N (2013) Broad domain generality in focal regions of frontal and parietal cortex. Proc Natl Acad Sci U S A 110:16616–16621 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3799302&tool=pmcentrez&rendertype=abstract.

Fedorenko E, Hsieh P-J, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N (2010) New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J Neurophysiol 104:1177–1194.

Fedorenko E, Varley R (2016) Language and thought are not the same thing: Evidence from neuroimaging and neurological patients. Ann N Y Acad Sci 1369:132–153.

Frost MA, Goebel R (2012) Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. Neuroimage 59:1369–1381 Available at: http://dx.doi.org/10.1016/j.neuroimage.2011.08.035.

Geerligs L, Tsvetanov KA, Cam-CAN, Henson RN (2017) Challenges in measuring individual differences in functional connectivity using fMRI: The case of healthy aging. Hum Brain Mapp 0 Available at: http://doi.wiley.com/10.1002/hbm.23653.

Gelman A, Carlin J (2014) Beyond Power Calculations: Assessing Type S (Sign) and Type M

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(Magnitude) Errors. Perspect Psychol Sci 9:641–651 Available at: http://journals.sagepub.com/doi/10.1177/1745691614551642.

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 107:4705–4709 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0910397107.

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

Haier RJ, Siegel B V, Nuechterlein KH, Hazlett E, Wu JC, Paek J, Browning HL, Buchsbaum MS (1988) Correlates of Abstract Reasoning and Attention Studied with Positron Emission Tomography. Intelligence 12:199–217.

Hampshire A, Highfield RR, Parkin BL, Owen AM (2012) Fractionating Human Intelligence. Neuron 76:1225–1237 Available at: http://dx.doi.org/10.1016/j.neuron.2012.06.022.

Hugdahl K, Raichle ME, Mitra A, Specht K (2015) On the existence of a generalized non-specific task-dependent network. Front Hum Neurosci 9:430 Available at: http://journal.frontiersin.org/article/10.3389/fnhum.2015.00430/abstract%5Cnhttp://journal.frontiersin.org/article/10.3389/fnhum.2015.00430/pdf.

Kaufman AS, Kaufman NL (2013) Kaufman Brief Intelligence Test, Second Edition. In: Encyclopedia of Special Education. John Wiley & Sons, Inc. Available at: http://dx.doi.org/10.1002/9781118660584.ese1325.

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540.

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–

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

586.

Mahowald K, Fedorenko E (2016) Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. Neuroimage 139:74–93 Available at: http://dx.doi.org/10.1016/j.neuroimage.2016.05.073.

Mineroff Z, Blank I, Mahowald K, Fedorenko E (2017) A robust dissociation among the language, multiple demand, and default mode networks: evidence from inter-region correlations in effect size Zachary. bioRxiv Available at: http://dx.doi.org/10.1101/140384.

Monti MM, Parsons LM, Osherson DN (2012) Thought Beyond Language: Neural Dissociation of Algebra and Natural Language. Psychol Sci 23:914–922 Available at: http://journals.sagepub.com/doi/10.1177/0956797612437427.

Nieto-Castañón A, Fedorenko E (2012) Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. Neuroimage 63:1646–1669 Available at: http://dx.doi.org/10.1016/j.neuroimage.2012.06.065.

Plomin R, Spinath FM (2004) Intelligence: genetics, genes, and genomics. J Pers Soc Psychol 86:112–129 Available at: http://www.ncbi.nlm.nih.gov/pubmed/14717631.

Reuter-Lorenz PA, Jonides J, Smith EE, Hartley A, Miller A, Marshuetz C, Koeppe RA (2000) Age Differences in the Frontal Lateralization of Verbal and Spatial Working Memory Revealed by PET. J Cogn Neurosci 12:174–187 Available at: http://www.mitpressjournals.org/doi/10.1162/089892900561814.

Roca M, Parr A, Thompson R, Woolgar A, Torralva T, Antoun N, Manes F, Duncan J (2010) Executive function and fluid intelligence after frontal lobe lesions. Brain 133:234–247.

Rypma B, Berger JS, Prabhakaran V, Bly M, Kimberg DY, Biswal BB, Esposito MD (2006)

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Neural correlates of cognitive efficiency. 33:969–979.

Rypma B, Esposito MD (2000) Isolating the neural mechanisms of age-related changes in human working memory. Nat Neurosci 3:509–515.

Scott TL, Gall?e J, Fedorenko E (2016) A new fun and robust version of an fMRI localizer for the frontotemporal language system. Cogn Neurosci 8:167–176 Available at: http://www.tandfonline.com/action/journalInformation?journalCode=pcns20 [Accessed June 10, 2017].

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349–2356 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17329432.

Tahmasebi AM, Davis MH, Wild CJ, Rodd JM, Hakyemez H, Abolmaesumi P, Johnsrude IS (2012) Is the link between anatomical structure and function equally strong at all cognitive levels of processing? Cereb Cortex 22:1593–1603.

Thesen S, Heid O, Mueller E, Schad LR (2000) Prospective Acquisition Correction for head motion with image-based tracking for real-time fMRI. Magn Reson Med 44:457–465.

Tschentscher N, Mitchell D, Duncan J (2017) Fluid Intelligence Predicts Novel Rule Implementation in a Distributed Frontoparietal Control Network. J Neurosci 37:4841–4847 Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2478-16.2017.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289.

Varley R, Siegal M (2000) Evidence for cognition without grammar from causal reasoning and

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

"theory of mind" in an agrammatic aphasic patient. Available at: http://www.sciencedirect.com/science/article/pii/S0960982200005388 [Accessed June 10, 2017].

Wager TD, Sylvester CYC, Lacey SC, Nee DE, Franklin M, Jonides J (2005) Common and unique components of response inhibition revealed by fMRI. Neuroimage 27:323–340.

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 U S A 107:14899–14902.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figures

Figure 1. Spatial working memory task

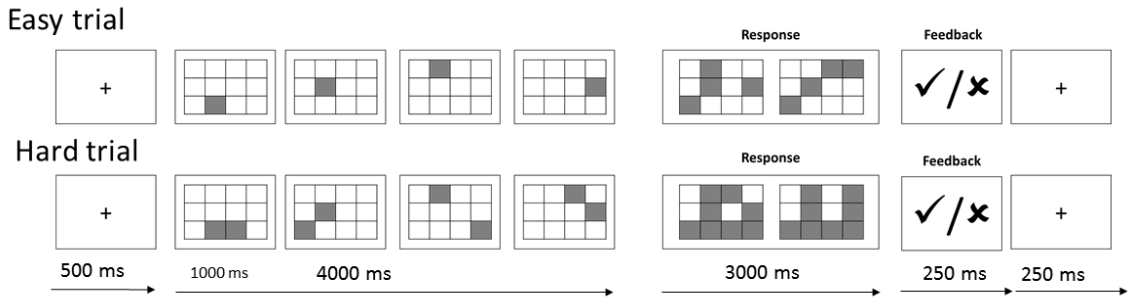


Figure 1 caption: Sample easy and hard trials.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 2. Stability of behavioral responses within individuals across runs

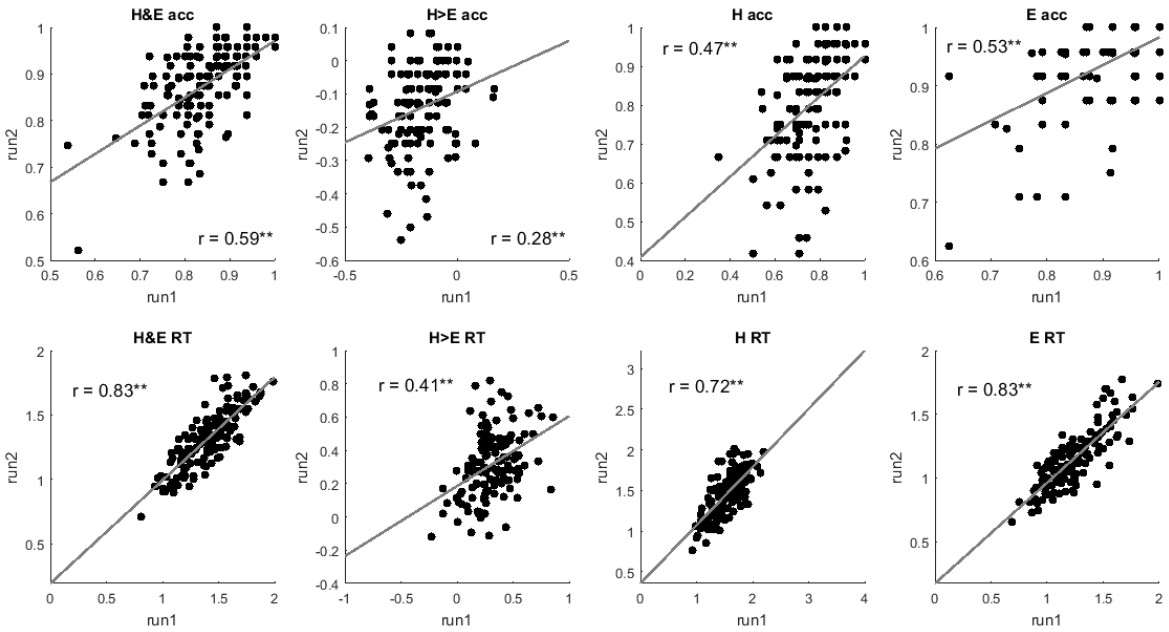
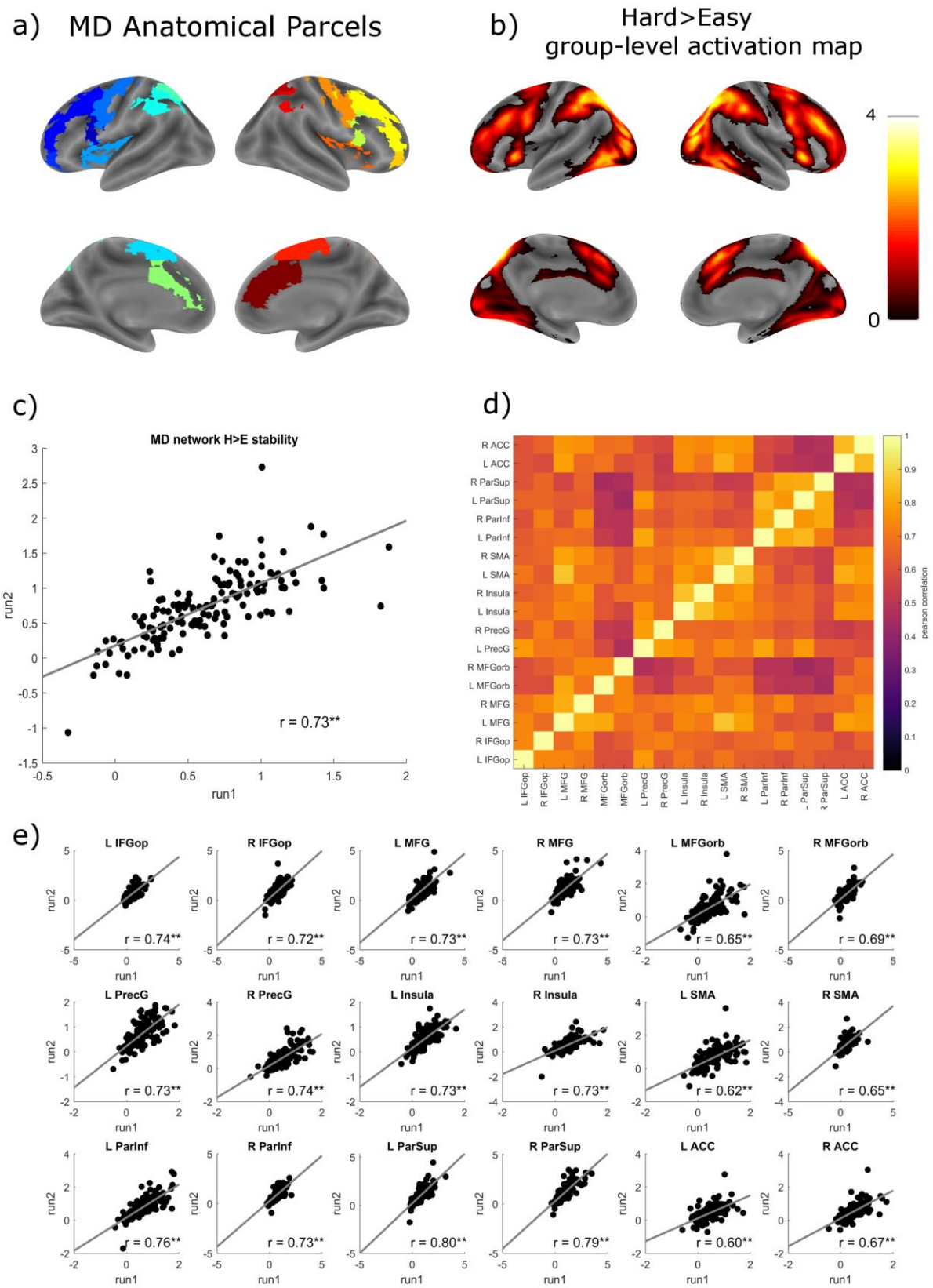


Figure 2 caption: acc (accuracy), RT (reaction time), H (Hard), E (Easy), H&E (average across Hard and Easy conditions), H>E (difference between Hard and Easy conditions). r (Pearson correlation), ** *p<0.001*

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 3. Stability of neural responses within individuals across runs



bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 3 caption:

a) MD anatomical parcels used to constrain individual fROI definition.

b) Group-level representations of neural activity for the Hard>Easy contrast. Individual unthresholded *t*-maps were averaged across 140 participants.

c) MD Network (average of all ROIs) Hard > Easy responses stability across runs (n = 140), r (Pearson correlation), ** $p<0.001$.

d) H>E effect size correlation across MD ROIs. Range of correlations (0.44 – 0.87). Average correlation (0.66). All correlations are significant ($p<0.05$ Bonferroni-corrected for 153 tests).

e) MD individual ROIs Hard > Easy responses stability across runs (n = 140).

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 4. The Multiple Demand neural responses-behavior relationship

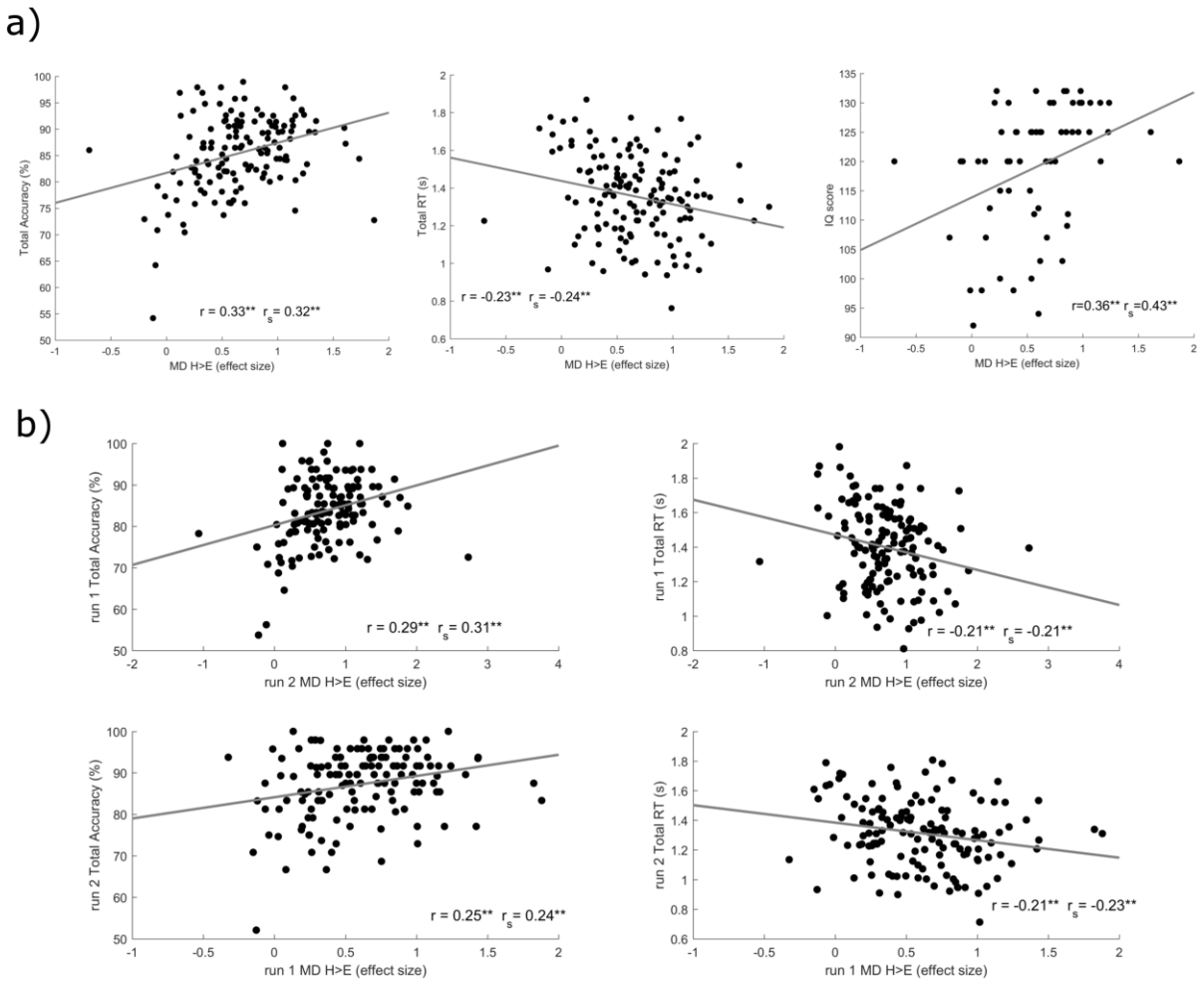


Figure 4 caption:

MD-behavior relationships: MD H>E effect size predicts (a) overall (average across both runs) accuracy (n = 140) (*left*) and reaction time (n = 140) (*middle*) in the spatial WM task, as well as IQ (n = 63) (*right*) as measured by an independent test (KBIT), and (b) performance for independent runs. r (Pearson correlation), r_s (Spearman correlation), ** *p<0.001*, * *p<0.05*.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 5. The left hemisphere language network and its neural responses-behavior relationship

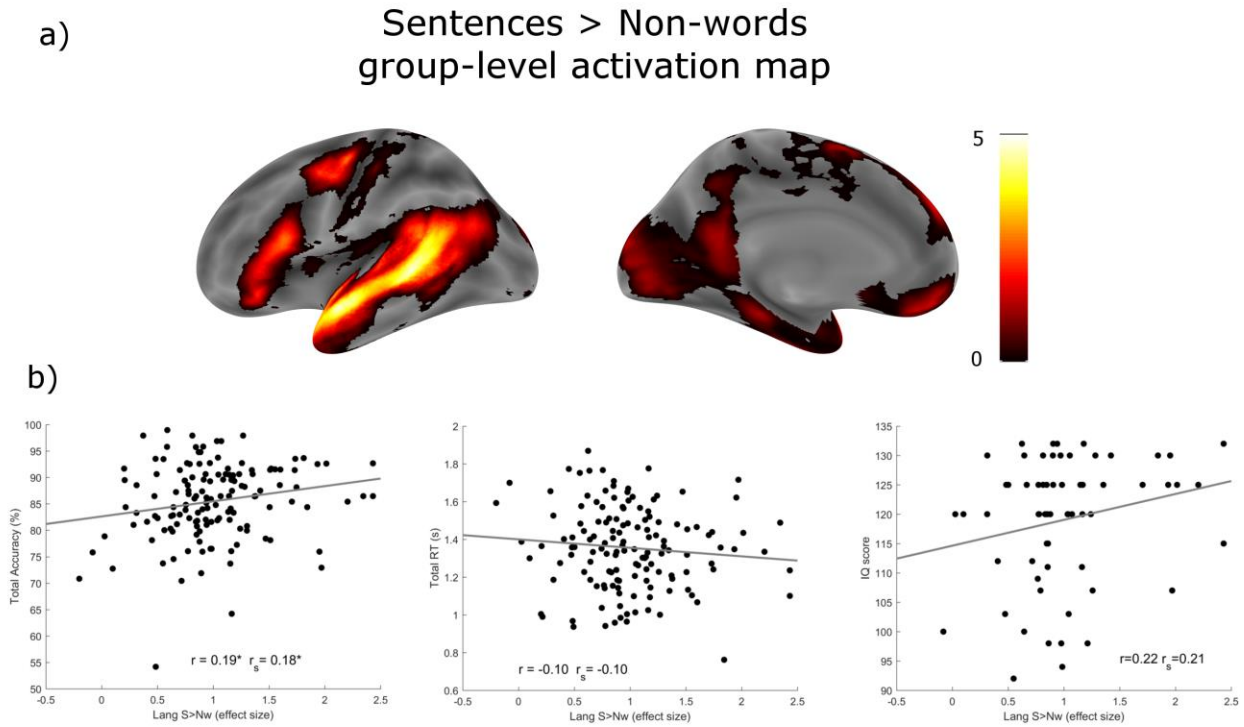


Figure 5 caption:

Language-behavior relationships: Language S>N relationship with overall accuracy (n=140) *(left)* RT (n=140) *(middle)* and IQ scores (n=63) *(right)*. r (Pearson correlation), r_s (Spearman correlation), ** *p<0.001*, * *p<0.05*.

bioRxiv preprint doi: https://doi.org/10.1101/110270; this version posted June 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 6: The effect of sample size on brain-behavior correlation

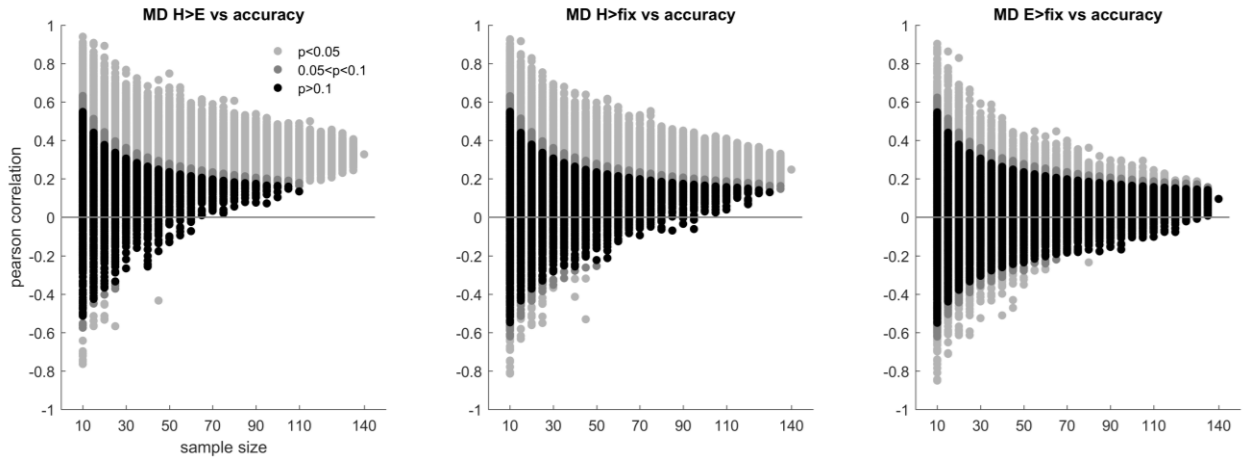


Figure 6 caption:

MD-total accuracy relationship for H>E (left), H>fix (middle) and E>fix (right) MD responses are shown on different sample sizes. Each sample size correlation analysis was repeated 1000 times, randomly sampled from the original data (n=140) each time.