INTER AND INTRA-INDIVIDUAL BEHAVIORAL VARIABILITY PREDICTED BY NEURAL ACTIVITY IN THE MULTIPLE DEMAND NETWORK

by

Moataz Assem

MB BCh, Bachelor in Medicine and Surgery, Alexandria University, 2013

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2016

INTER AND INTRA-INDIVIDUAL BEHAVIORAL VARIABILITY PREDICTED BY NEURAL ACTIVITY IN THE MULTIPLE DEMAND NETWORK

APPROVED BY:

Prof. Dr. Ahmet Ademoğlu .....
(Thesis Advisor)

Assist. Prof. Dr. Evelina Fedorenko .....
(Thesis Co-advisor)

Assist. Prof. Dr. Burak Güçlü .....

Prof. Dr. Tamer Demiralp .....

Assoc. Prof. Dr. Ayşecan Boduroğlu .....

DATE OF APPROVAL: 6 June 2016

ACKNOWLEDGMENTS

First, I would like to express how deeply indebted I am to my advisor and mentor Ahmet Ademoglu for his never ending support since the first day I met him. Our coffee table discussions enriched me both academically and on personal life matters. He is also the best partner to run with.

I am extremely grateful for my co-advisor and friend Evelina Fedorenko for accepting me as a member in her amazing team. Despite working together from different continents, I never felt she could support me more. I would like to thank her lab members: Zach Mineroff(ah ah aahh), Brianna Pritchett, Caitlyn Hoeflin and Idan Blank for the good times we shared together.

I would like to thank Tamer Demiralp for his continuous support for always welcoming me to his lab. I would like to thank his lab members: Ani Kicik, Elif Kurt, Cigdem Ulasoglu and Emel for our crazy and fun times together.

My time at NeurAl couldn't have been happier without its present and past members: Duygu Sahin, Huden Nese, Sencer Melih, Pinar Adanali, Esma Ece and Ali Bayram. I am grateful to Esin Karahan, for her invaluable guidance and for sharing the long working lab nights with me. I am also grateful to Seda Dumlu for being the concon she is and turning me into one. I am also thankful to Adil Deniz Duru for mentoring me in EEG research and our kofte sandwiches.

The unforgettable fun at BME was only possible with Bora, Hakan, Nermin, Fatma, Hale Hatay, Sevim, Dilek Betul Arslan, Agah Karakuzu, Sedef, Deniz, Ismail Devecioğlu. I am happy I got to meet people like Bige Vardar, Sevgi Ozturk, Didar Talaat, Mustafa Ruhi, and Mohamed Avsar in my life. I am thankful to my friend Ahmed Emad for the good times we shared together.

There are no words that can describe my love towards my wife-to-be Melek Karadag. I wouldn't be where I am without you. Thank you for always being by my side, for introducing me to new things in life and for accepting to spend the rest of your life with me, despite all my weirdness.

Finally, none of this would have been possible without the support of my family. Despite their initial doubts about my career path, they never stopped supporting and believing in me. This thesis is dedicated to my mother Maha Moharram, my father, Ibrahim Assem and my sister Mariam Assem.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Moataz Assem, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name :

Signature:

Date:

ABSTRACT

INTER AND INTRA-INDIVIDUAL BEHAVIORAL VARIABILITY PREDICTED BY NEURAL ACTIVITY IN THE MULTIPLE DEMAND NETWORK

A wide range of cognitive tasks consistently identify a Multiple Demand (MD) network in frontal and parietal brain regions. Its activity is closely linked to executive functions (EFs) such as attention, task switching, solving novel problems and manipulating information in the working memory. We here investigate the relation between MD neural activity and EFs using a large fMRI dataset (n=120). We examine this relation through two approaches (1) inter-individual variability, addressing several methodological challenges: We find that MD activity - which varies substantially across individuals, but is consistent within individuals across time - can explain a substantial proportion of variance in individual performance on a spatial working memory task such that individuals who find the task challenging, increase their MD activity substantially to improve their performance. This suggests that MD activity tightly reflects the executive demand of an individual. In the second approach we examine (2) trial-by-trial variability by employing three different models to fuse reaction time (RT) data with the BOLD time-series. We find that BOLD amplitude increases with longer RTs. This is consistent with the findings from the first approach showing increased MD activity for slower individuals. Together both findings support the view that within and between individual differences are manifested in the same brain regions. These results have implications for (1) understanding brain processes of EFs through ID studies (2) given that ID in EF are largely genetically determined, genetic variability can be linked to the neural activity of the MD network as an intermediate stage to link genetics with behavior (3) using ID in fMRI responses as clinical biomarkers.

Keywords: Multiple Demand, Executive Functions, Individual Differences, Reaction time, Spatial Working Memory, Parametric modulation, fMRI.

ÖZET

BİREYLER ARASI VE BİREY İÇİ DAVRANIŞSAL FARKLILIKLARIN ÇOKLU TALEP SİSTEMİNDEKİ NÖRAL AKTİVİTE KULLANILARAK ÖNGÖRÜLMESİ

Geniş bir yelpazeye yayılan bilişsel görevler, Çoklu Talep (ÇT) ağını frontal ve paryetal beyin alanlarında tutarlı olarak tanımlamaktadır. Bu ağın aktivitesi dikkat, görev geçişi, alışılmamış problemleri çözme ve işler bellekte bilginin işlenmesi gibi yürütücü işlevlerle (Yİ) yakından ilişkilidir. Bu çalışmada ÇT nöral aktivitesi ve Yİ arasındaki ilişkiyi geniş bir fMRG veri öbeği kullanarak araştırdık (n=120). Bu ilişkiyi bireyler arası değişkenlik ve denemeler arası değişkenlik olmak üzere iki yaklaşımla inceledik; (1) Yöntemsel zorlukları göz önüne alarak incelediğimiz bireyler arası değişkenlikte, bireyler arasında önemli ölçüde değişen ancak zaman boyutunda bireyler içinde tutarlı gözüken ÇT aktivitesi, uzaysal işler bellek görevlerindeki bireysel performans (BF) farklılıklarında oluşan değişkenliğin önemli bir bölümünü açıklayabilmekte, görevi zorlu bulan bireyler performanslarını geliştirmek için aktivitelerini önemli ölçüde arttırmaktadırlar. Bu durum ÇT aktivitesinin bireyin yürütücü talebini bütünüyle yansıttığını düşündürtmektedir. (2) İkinci yaklaşımda ise tepki zamanı verisini BOLD zaman serileri ile birleştirmek için üç farklı model kullanarak denemeler arası değişkenliği inceledik. BOLD genliğinin uzayan tepki zamanıyla arttığını gösterdik. Bu sonuç ilk yaklaşımdan elde edilen daha yavaş bireylerde artan ÇT aktivitesi bulgusuyla tutarlıdır. İki yaklaşımdan elde edilen bulgular da bireyler arası ve bireyler içi farklılıkların aynı beyin bölgelerinde ortaya çıktığı görüşünü desteklemektedir. Bu sonuçlardan yapılan çıkarımlar (1) Yİ'deki beyin süreçlerinin bireysel farklılık çalışmaları üzerinden anlaşılması (2) Yürütücü işlevlerdeki bireysel farklılıkların büyük ölçüde genetik olarak belirlendiği göz önüne alındığında, ÇT nöral aktivitesinin genetik farklılıkla olan bağlantısının genetik ve davranışı ilişkilendiren ara bir aşama olarak ortaya konması (3) fMRG yanıtlarındaki bireysel farklılıkların klinik biyobelirteç olarak kullanılması üzerinedir.

Anahtar Sözcükler: Çoklu Talep, Yürütücü İşlevler, Bireysel Farklılıklar, Tepki Zamanı, Uzaysal İşler Bellek, Parametrik Modülasyon, fMRG

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$Y^{(i)}$	Time-series data vector for (i) model level
$X^{(i)}$	Design Matrix for (i) model level
$\beta^{(i)}$	Parameter estimates for (i) model level
$\epsilon^{(i)}$	Residual errors for (i) model level
C_ϵ	Error Covariance Matrix
V	Temporal Autocorrelation Matrix
σ^2	Error Variance
w	White Noise
λ	Hyper-parameter
Q	Covariance Component

LIST OF ABBREVIATIONS

fMRI	Functional MRI
SNR	Signal to noise ratio
BOLD	Blood Oxygen Level Dependent
MD	Multiple Demand
ID	Individual Differences
EFs	Executive functions
E	Easy
H	Hard
H-E	Hard minus Easy
H&E	sum of Hard and Easy
VarImp	Variable Impulse
VarEp	Variable Epoch
FIR	Finite Impulse Response
ReML	Restricted Maximum Likelihood Estimation

1. INTRODUCTION

1.1 The Science of Individual Differences from FMRI

Individuals differ from each other both behaviorally and physiologically. Underlying these differences are complex genetic, environmental and neural interactions. Executive functions (EFs), a set of cognitive abilities necessary for goal-directed behavior, is one area where individual differences (ID) are prominent. Although genetics account for a significant proportion of EFs variability [1, 2], it has failed to be linked directly to behavioral differences [3]. On the other hand, linking genes with neural activity has been more successful. Thus, it has been suggested that neural activity can serve as an intermediate stage in linking genetics with behavioral variability [3, 4].

A common method to investigate neural activity is functional MRI (fMRI). The prevalent approach in fMRI research is to average functional brain scans across individuals to improve the low signal to noise ratio (SNR). This approach is useful in revealing general brain processes that are common among individuals, as well as exposing differences between groups, as in comparing patients and healthy groups.

Yet neural activity is substantially variable from one individual to another and the averaging approach has limited the understanding of brain activity at the individual level [5]. Further, a number of methodological challenges held back the progress of investigating neural activity at the individual level. These are mainly concerned with the validity and reliability of comparing brain activity between individuals [4]-[6].

That said, studies examining neural activity and individual differences are increasing due to the great potential they carry for basic research and clinical applications. If functional brain scans can be acquired for a large population, then an individual's brain scan can be statistically compared against it. This has the potential of using individual scans as clinical bio-markers for early diagnosis and better classifi-

cation of patients. Moreover, since inter and intra-individual variances are statistically independent, studies of ID can be used to infer underlying brain processes as a complementary evidence to those revealed by experimental manipulation or difficult to investigate experimentally.

1.2 Objective

In this thesis, We aim to provide a detailed and comprehensive investigation of the relation between a Multiple Demand (MD) brain network and individual behavioral variability in a spatial working memory task.

We will examine both (i) trial-by-trial variability (i.e., dynamic changes in blood oxygen level dependent (BOLD) signal and behavior within an individual over time) (ii) individual-by-individual variability (i.e., stable characteristics of individuals).

We aim to investigate to what extent can behavioral variability be predicted by the MD network activity and whether inter and intra-individual differences findings from the same brain regions converge.

1.3 Thesis Outline

The thesis is organized as follows:

In Chapter 2, we examine inter-individual differences. We first introduce the Multiple Demand system and its relation to executive functions. Followed by a review of previous fMRI research on individual differences in executive functions focusing on working memory studies. In a separate section we outline the methodological shortcomings of several of the previous studies and set out a hypothesis space for the expected results. We then present the experiment and analysis through which we investigate the

relation between MD activity and behavioral performance on a spatial working memory task and illustrate how we address any methodological shortcomings. We then present the results.

In Chapter 3, we examine intra-individual differences. Using the same dataset from Chapter 2, we examine reaction time variability with BOLD signal from the MD network. First, We review previous studies examining fMRI activity and reaction time on a trial-by-trial basis. We then employ three different models to fuse fMRI and behavior.

In Chapter 4, we discuss both inter-individual findings in contrast with previous studies and explain the extent to which MD activity can account for individual behavioral variability. We then present intra-individual results and discuss the validity of examining inter and intra-individual variations in the same brain regions.

In the Conclusion, we summarize the main results and their implications for future neuroimaging research on individual differences.

In the Appendix, we included detailed results of performance on the working memory task as well as results concerning the validation of fMRI predictions.

2. INTER-INDIVIDUAL DIFFERENCES

2.1 Background and Literature Review

2.1.1 A Multiple Demand Brain Network for Executive Functions

Executive functions (EFs) is an umbrella term for a group of cognitive abilities, such as the ability to pay attention to a specific item while ignoring other distractions (selective attention), monitoring the surrounding environment (performance monitoring), detecting false outcomes from actions (error feedback), switching between tasks, manipulating items in working memory, withholding an unwanted response (response inhibition) and representing the current available information into an abstract set of rules to guide novel and flexible goal-oriented behavior [7].

EFs are tightly linked to individual differences in general cognitive abilities such as general fluid intelligence (gF) and working memory capacity (WMC) [8, 9].

The neural architecture underlying EFs has been extensively investigated. The focus started on the lateral prefrontal cortex (PFC) with Positron Emission Tomography (PET) studies finding its activity increases while performing tasks with high cognitive loads [10] and this activity predicted participants' gF. Improvements in fMRI SNR allowed studies to implicate a wider network beyond lateral PFC, mainly in the parietal cortices as well as medial PFC [11, 12].

A consistent distributed fronto-parietal network then started to emerge from several studies tapping into different aspects of EFs such as response inhibition [13, 14], working memory [15, 16], selective attention [17, 18], task switching [19] and flexible coding of task information [20, 21]. Meta-analysis of several of these studies showed a common set of frontal and parietal regions that are recruited in a diverse set of tasks [22]. This network is known in the literature by several names, including the

Multiple Demand (MD) system [22, 23]. A recent study showed that the difference between a hard and easy version of seven diverse tasks (spatial and verbal WM, three response inhibition tasks, arithmetics, Non-words > sentences) revealed a distributed set of focal regions, in the frontal and parietal lobes, at the individual subject level [24] (i.e the same set of voxels in each individual respond to each of the seven different tasks) providing yet the strongest and most comprehensive evidence for the existence of a core set of domain general brain regions in every individual brain (Figure 2.1) yet varying in pattern from one person to another (Figure 2.2 (B)). The same study further validated this conclusion by testing responses from control regions in the temporal lobe, outside of the MD system, on each of the seven tasks and found no voxels that respond commonly to all tasks.

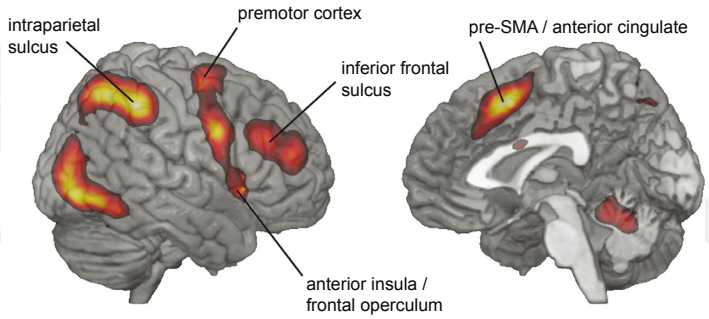


Figure 2.1: The Multiple Demand system. Adapted from [24].

Neuropsychology studies provided complementary support for the existence of a MD network. For example, patients with lesions to frontal and parietal cortices had deficiencies in a number of executive functions [25] and in a more spatially specific study, lesions to MD regions, but not outside of it, predicted loss of gF [26].

Studies tapping into the function of the MD system find that it encodes a broad range of information [23] and task rules [21, 27] as well as flexible and rapid transitions to encode different stimuli within the current context [28]-[30]. MD regions are also highly correlated in activity either during rest [31] or task [32].

This led to the proposal that the MD system main role is to guide complex

behavior through the construction of a series of attentional episodes [23]. The rapid and flexible encoding properties of its neurons support the idea of encoding the current attentional state. MD system then communicates with other brain networks during each episode, assembling and diffusing connections to solve the current problem [23].

That said, the neurophysiological processes underlying MD activity and the contribution of MD system to individual differences in EFs is still a matter of investigation. In the next section, we discuss previous fMRI studies on individual differences in EFs, focusing on working memory, and point out several of their limitations and the contribution that this thesis can add to the existing literature to inform both areas of individual differences in EFs and the neurophysiology of MD system.

2.1.2 Previous fMRI Studies on Inter-Individual Differences in Executive Functions

Individual differences (ID) studies relate brain activity to two types of behavioral measures (1) in-scanner task performance such as RT or accuracy (2) measures of general cognitive abilities from tests outside of the scanner such as tests of gF. The former approach is thought to give unreliable results due to the lack of control of confounding factors related to current state of the subject. While the latter approach can provide a better understanding of findings from the former approach, it is less commonly implemented. Nonetheless, in both approaches, most EFs studies use activity measures from a subset of the MD regions but never the whole set.

One study investigating ID in RT report that activity in fronto-parietal regions (from three different response inhibition tasks) could not predict individuals RT except for the bilateral insula, and inferior frontal gyrus (IFG), which are found to have higher activity for slower individuals [13]. Further, only when they used an averaged activity measure from a frontal-cingulate-opercular network, rather than separate regions, they report the same relation; that is slower subjects have a higher level of activity in this network.

Another study on ID in RT found that slower individuals had a larger number of voxels recruited during a WM task [33]. Outside of the MD network, one study used a 3-back WM task and reported opposite findings that faster subjects have increased activity in the amygdala [34].

To summarize the few studies investigating RT ID, they find that slower subjects have, on average, increased neural activity within MD regions. This activity could be reflecting increased neural processing as longer RTs usually accompany tasks with a higher cognitive load. However, some studies argue that this reflects a "time on task" effect. In other words, the relation between RT and fMRI activity could be reflecting basic properties of the BOLD signal. The longer the subject stays on a task, the larger the BOLD signal will be, simply because of linear summation of the BOLD signal across a longer time [35]. However, several studies show that regressing out RT variance from BOLD signal still shows significant activity in MD regions [36, 37] (See Discussion section for further details).

As for studies investigating ID in accuracy, their results are not consistent. Two large scale working memory (WM) studies report that better performing individuals have on average stronger activity in lateral PFC, ACC and cerebellum [11], all within the MD regions. While outside of the MD regions, one study reports increased activity in the medial posterior parietal cortex predicts performance accuracy [4].

On the other hand, Wager et al. report weaker neural activity for lower performers in IFG (on a go-no-go task) and in posterior parietal cortex (on a flanker task) [13]. Rypma et al. reports the same relation; less accurate individuals had weaker activity in the ventro- and dorso-lateral PFC (in a WM task) [15].

Another study using a 3-back WM task found both relationships in different regions. Better performing individuals had higher activity in the left MFG yet lower activity in left ACC and inferior parietal regions [38], again all are within the MD network.

In summary, ID studies on accuracy show a mixed picture. On one hand, researchers explain weaker activity for individuals who perform better in terms of *neural efficiency* (i.e., more efficient recruitment of their cognitive resources). However, it is argued that better performers have stronger activity for briefer periods, which is not detected by averaging over long tasks thus showing weaker activations [11]. One further explanation is that activity-difficulty relation in MD network might follow an inverted U-shape. It increases as the task cognitive load increase, reaches a peak, then decreases as the individuals give up if the task is too hard. The previously mentioned studies use different tasks with at most two levels of difficulties which might explain these mixed results.

As for studies investigating ID using more stable behavioral measures such as gF scores, they report that participants with higher gF scores have on average stronger activity in lateral PFC, dorsal anterior cingulate cortex (ACC) and lateral cerebellum [11], parietal regions [12] and left MFG [37]. These studies lend support to the view that better performers have stronger MD activity. Importantly, one of these studies show that activations in three MD regions *mediate* the relation between gF and accuracy (i.e. their neural activity could account for variance between the two behavioral measures) which is a more powerful measure than simple correlations [11]. Worth to note also is that this relation was only observed when using activity from trials with high cognitive load (non-lure trials in an n-back task). Easier trials (lure and target) showed no reliable correlations.

Working memory capacity (WMC) is another out of scanner cognitive measure that studies used. Individuals with higher WMC have stronger activity in ACC and IFG [16], and posterior parietal cortex (PPC), intra-parietal sulcus (IPS) and intra-occipital sulcus (IOS) [39]. These studies again lend support to the view that better performing individuals have stronger MD activations.

In summary, studies on ID using fMRI show increased activity in MD regions for (1) slower subjects (2) subjects with better cognitive abilities (higher gF and WMC scores) (3) better performing subjects (more supported by studies combining accuracy

and gF or WMC). There are two clear discrepancies here; (a) subjects with higher cognitive abilities are usually faster in response (b) not all studies show increased MD activity for those with better performance. In the next section, we illustrate several methodological limitations for many of the studies mentioned above which might explain such discrepancies. We then lay out a hypothesis space for expected results by considering only methodologically sound results.

2.1.3 Limitations of Previous Studies

Most of the previously mentioned studies suffer from a number of shortcomings that are of concern for fMRI studies of ID in general, these include:

1. **Small number of subjects**: Projecting inferences using a small number of participants can produce inflated statistics (n = 14 for [13] and n = 6 for [33], n = 17 [39]) casting doubt on the studies concluding decreased neural activity in better performers.

2. **Validity of neural measurements**: fMRI responses from a Task>Fix contrast may reflect state (current) conditions or trait characteristics of the individual which confound ID of interest. In other words, a subject might have better performance or higher activations on a task due to unrelated state conditions such as motivation, sleepiness, caffeine intake or differences in trait characteristics such as different levels brain vascularizations or age. This again increases doubt about studies, which used this contrast, showing decreased neural activity in better performers [13] or which showed mixed results from different regions [38] or the one showing increased activity with WMC [39]. Another limitation in [38] is using the same data for ROIs definition and analysis. This type of circular analysis can give biased results [40].

3. **Reliability of results**: The use of simple correlational analysis is less statistically powerful than other types of analysis such as prediction analysis [5]. In a prediction analysis, BOLD and behavioral measures should be from independent

sources. This is a common limitation for most of the studies mentioned above. However, studies which also correlated neural measures with independent measures (gF or WMC) are more reliable. This increases confidence in the findings of increased neural activity with better performers.

4. **Use of Extreme Groups** Categorizing participants into high and low groups can inflate correlations [6]. This result decreases confidence in the high- level of correlations reported by [12, 16].

5. **Lack of Intra-Individual differences analysis**: None of the previous studies complement their findings with analysis of trial-by-trial BOLD-behavior variability which complicates the interpretation of the results. Only one study performed this within-subjects analysis and found converging evidence for increased activity in the medial PPC for more accurate subjects and on accurate trials [4]. However, the analysis was reported for one region only. Nevertheless, it again reinforces the findings of increased neural activity in better performers.

2.1.4 Hypothesis Space

In the following analysis, we investigate the ID in MD fMRI responses with in-scanner behavioral measures: RT and accuracy. From the previous literature review and taking into consideration the limitations of each study, one can set out a hypothesis space for the expected results.

Concerning accuracy, we expect more accurate subjects to have, on average, increased neural activity. This is following the results by the most methodologically sound study [11]. Such a finding would lend further support to the findings that stronger MD network activity is related to having higher cognitive abilities [10, 11].

As for RT, although previous studies suggest increased neural activity for slower subjects, this is in contradiction with the classical evidence of faster responses for more accurate individuals or those with higher cognitive abilities (excluding tasks with speed-

accuracy trade-offs). Thus, we would expect that faster subjects would show increased MD activity, in line with the findings on accuracy.

2.2 Methodology

2.2.1 Participants

120 participants were recruited from the Massachusetts Institute of Technology (MIT). Data were collected by Ev Fedorenko's Language Lab (PI: Evelina Fedorenko) and Kanwisher Lab (PI: Nancy Kanwisher). Participants ages ranged from 18 to 50 and their vision was normal or corrected-to-normal. Informed consent was given by all participants in line with Internal Review Board at MIT.

2.2.2 Experiment Design

Participants performed a spatial working memory task. The task was block designed with four blocks for each of the easy, hard and fixation conditions. Each trial consisted of four successive 3x4 grids, with one (easy) or two (hard) square locations greyed out in each grid. For each trial, participants had to keep track of four (easy) or eight (hard) square locations in total. Then they had to choose the correct grid in a two-alternative forced-choice by pressing one of two buttons followed by a feedback screen. Two fixation screens at the start and end of each trial were also presented. (Figure 2.2 (A))

Before the spatial working memory runs, there were two functional localizer runs using a block design of sentences and non-words. These are used for validation analysis (Section 2.3.3).

2.2.3 fMRI data acquisition

Data were acquired using a 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. 176 sagittal T1-weighted scans were acquired. BOLD data were collected using an echo-planar imaging (EPI) sequence (Acquisition parameters: 31 4-mm thick near-axial slices acquired in the interleaved order, field of view in the phase encoding (A$\gg$ P) direction 200 mm and matrix size 96 mm $\times$ 96 mm, $TR = 2000$ ms, and $TE = 30$ ms). To allow for steady-state magnetization, the first 10 s of each run were disregarded .

2.2.4 Functional ROIs definition and Data analysis

Data preprocessing pipeline included motion correction, normalization to Montreal Neurological Institute (MNI) space and resampling into 2-mm isotropic voxels. Followed by smoothing using a 4-mm Gaussian filter and high-pass filtering (at 200 s). Boxcar regressors were created for the Easy and Hard conditions and convolved with a canonical Hemodynamic Response Function (HRF) and its temporal derivative. Motion regressors and a mean regressor were further added to the model.

Functional Regions of Interest (fROIs) were defined in each individual participant by i) creating a set of anatomical parcels corresponding to regions where MD activity has been previously reported [22] ii) intersecting the anatomical map with the functional activation t-map for the Hard-Easy contrast iii) The top 10 % of voxels that fell within the anatomical parcels were included in the analysis. This resulted in the definition of 18 fROIs, 9 in each hemisphere (Table 2.1 and Figure 2.2 (B))

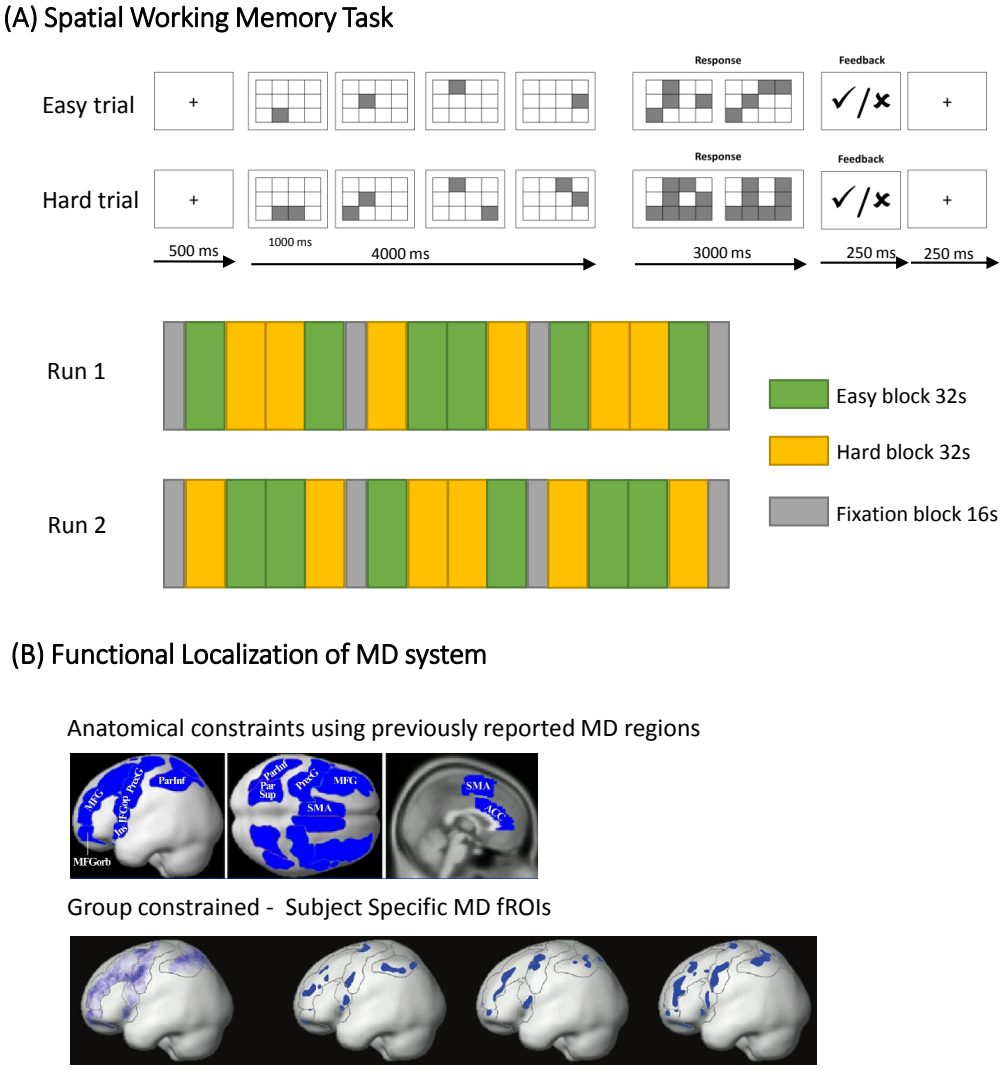


Figure 2.2: (A) Each condition had six blocks, each block lasted for 32 sec and consisted of four trials. Four fixation blocks each 16 sec (B) MD fROIs definition (*Top*) Anatomical ROIs constraining subject activations (adapted from [24]). (*Bottom*) Left-most brain: probabilistic H-E activations across a set of subjects. The other three brains are examples of single subject activations constrained by anatomical ROIs adapted from [31].

Table 2.1: Identified MD ROIs at subject level.

	MD regions (bilateral)
1	Inferior Frontal Gyrus opercular (IFGop)
2	Middle Frontal Gyrus (MFG)
3	MFG Orbital (MFGOrb)
4	Inferior Parietal (ParInf)
5	Superior Parietal (ParSup)
6	Anterior Cingulate Cortex (ACC)
7	Insula
8	Supplementary Motor Area (SMA)
9	Pre-central Gyrus (PrecG)

For response estimation, the first run was used to define the fROIs while the second run was used for estimation then vice versa to ensure that the data used to identify the ROIs are independent from the one used for estimation [40]. Responses were averaged across voxels for each ROI and then further averaged across all ROIs for each individual.

We divided participants into two sets, 60 in each, and performed the analysis on both separately in order to examine to what extent results can be replicated. We will now present the results from the first set in details and then present the main results from the second set in a separate section.

2.3 Results

2.3.1 Behavioral Performance

Averaging across both runs, individuals are faster and more accurate on the easy (E) trials (reaction time (RT), 1.19 $\pm$ 0.22 sec (mean $\pm$ s.t.d)), accuracy (% correct responses) 91.7% $\pm$ 7.49) than hard trials (H) (RT, 1.47 $\pm$ 0.27 sec, accuracy 78.65% $\pm$ 11.11, $t_{59} = 14.19$ for RT and $t_{59} = -11.40$ for accuracy, $p < 0.0001$ for both) (Figure 2.3 (a)).

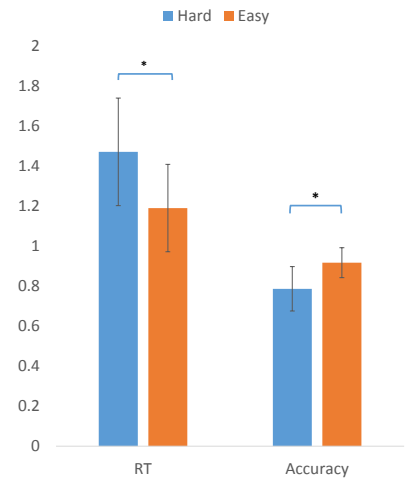
This relationship also holds between individuals, as those who respond faster are more accurate on either conditions ($r = -0.32, p = 0.013$) (Figure 2.3 (b)). These findings also hold when each run is examined separately and are stable between runs (correlation between run 1 and run 2 for RT(E) ($r = 0.87$), RT(H) $r = 0.81$, RT Hard minus Easy (H-E) $r = 0.48$, accuracy(E) ($r = 0.59$), accuracy(H) ($r = 0.60, p < 0.0001$ for all) accuracy(H-E) (% correct responses on hard trials minus % correct responses on easy trials) ($r = 0.36, p < 0.005$).(Figure 2.3 (c-d)).

Individuals who performed better (H&E) (average performance on all trials) showed smaller H-E differences for both accuracy and RT (accuracy(H&E) vs accuracy(H-E) ($r = 0.45, p < 0.001$); RT(H&E) vs RT(H-E) ($r = 0.34, p = 0.0003$).

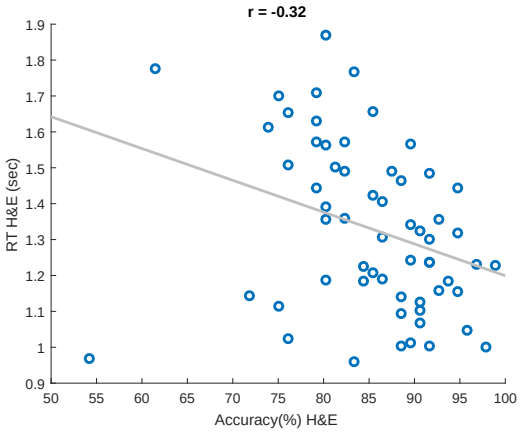
However, on examining correlation between accuracy(H-E) and RT(H-E) we found no significant relation ($r = -0.15, p = 0.26$). This could be because both low and high performers can show small H-E differences. This is indeed evident from the distribution of H-E accuracies when plotted against average performance accuracy(H&E) (Figure 2.4 (b)). Thus, when we excluded low performers (those with accuracy $<$75%, n=56), the expected relation between accuracy H-E and RT H-E was found ($r = -0.29, p = 0.030$) (Figure 2.4 (c)) and the positive relation between overall performance and H-E was more significant (accuracy(H&E) vs accuracy(H-E) $r = 0.47, p < 0.0001$; RT(H&E) vs RT(H-E) $r = 0.45, p < 0.0001$).

The same relation between RT(H-E) distribution against accuracy (H&E) could not be observed (Figure 2.4 (c)). This suggests that differences between RT in the H and E conditions do not reflect differences in overall individual performance.

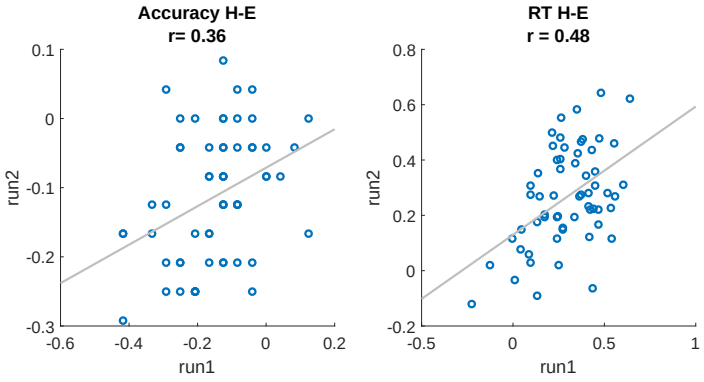
To test whether subjects might be performing better on the 2nd run (learning effect), we compared the means and variances between both runs for RT(H-E) ($F_{1,59} = 1.5, p = 0.4$ and $t_{59} = 1.2, p = 0.23$) and accuracy(H-E) ($F_{1,59} = 1.67$, $p = 0.025$ and $t_{59} = -2.27, p = 0.027$). These results show that subjects were indeed performing better on the 2nd run but their speed did not change.



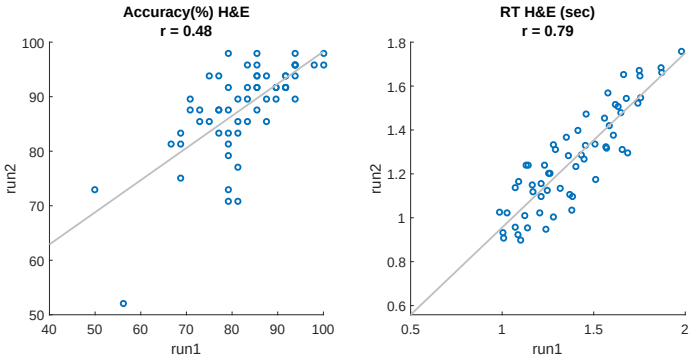
(a)



(b)



(c)



(d)

Figure 2.3: (a) Behavioral performance in spatial WM task (b) Individuals with better performance have shorter RTs (c) and (d) Stability of behavioral measures across runs within individuals.

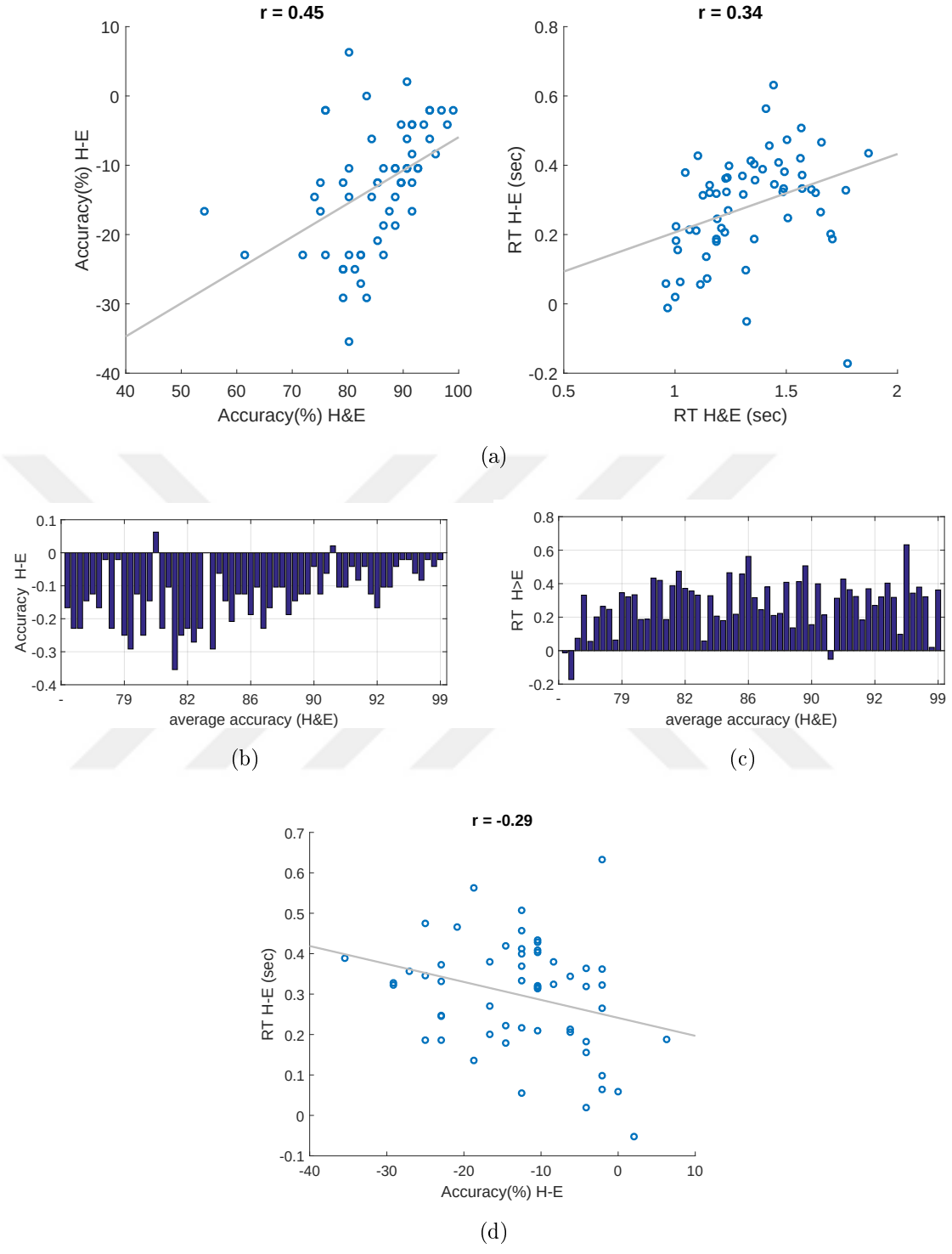


Figure 2.4: (a) Better performing subjects show small differences between H-E accuracy and RT, however, when looking at distribution of accuracy H-E scores (b) it is evident that low H-E scores can belong to both low and top performers. while (c) such a relation is not evident for the RT H-E distribution (d) Removing low performers (with accuracy <75) reveals the negative correlation between accuracy H-E and RT H-E, which was previously masked by low performers.

2.3.2 FMRI Predictions

2.3.2.1 Stability of MD fMRI responses within Individuals. MD fMRI responses are stable within individuals across both runs for the H condition and the H-E contrast for each MD ROI (Figure A.1) and on average across all ROIs (correlation between run 1 and run 2 for MD(H) $r = 0.67$ and MD(H-E) $r = 0.78$, $p < 0.0001$ for both). However, MD responses were less stable in the E condition for separate ROIs and on average (MD(E) $r = 0.15$, $p = 0.24$). (Figure 2.5)

2.3.2.2 Prediction of Behavior using MD fMRI responses. In the following analysis, we focused on using the contrast (H-E) as an index to compare BOLD and behavioral data. As mentioned before on the study limitations (section 2.1.3), the E>fix or H>fix contrasts could be affected by several confounding factors related to the subject's trait and state level properties such as their motivation to perform the task or different brain vascularizations. Thus a difference contrast would minimize such variability and mainly reflect change in task demand.

All MD effect sizes were averaged across all ROIs as the aim of this thesis is to understand relation between MD responses as a whole system. This is motivated by previously mentioned literature review on involvement of all ROIs in any individual task [24], the high task functional connectivity between all ROIs (i.e it is unlikely that some ROIs will be active while others don't) [32, 31] as well as the high correlation between individual ROIs effect sizes (Figure A.2).

Averaging across both runs, MD(H-E) responses are positively correlated with RT(H-E) ($r = 0.44 p < 0.001$) (i.e, as the difference in MD activity between the E and H conditions increase, the difference between H RT and E RT also increase). (Figure 2.6 *Left*). MD(H-E) is also correlated with RT(E) ($r = -0.30$, $p = 0.02$). MD(H-E) is not correlated with overall RT in both conditions (RT H&E) ($r = -0.14$, $p = 0.29$). To further test the prediction power of the MD H>E index, we compared BOLD-Behavior relations across the different runs. MD(H-E) run 2 successfully predicted RT(H-E) run

1 ($r = 0.43p < 0.001$) while MD(H-E) for the first run failed to predict RT(H-E) for the second run ($r = 0.17$, $p = 0.18$).

Regarding accuracy, MD(H-E) responses did not correlate with accuracy(H-E) ($r = -0.097, p = 0.46$). As mentioned earlier could be because small MD(H-E) responses can reflect both top performers (Figure 2.4 (b-c)). Thus, we repeated this analysis after excluding bottom performers (with an accuracy <75, remaining n = 54) which revealed a trending negative correlation ($r = -0.20, p = 0.14$). Again to test the predictive power across runs, MD(H-E) of the second run could predict accuracy(H-E) of the first run ($r = -0.24, p = 0.08$) but not vice versa ($r = -0.10, p = 0.44$) (Figure 2.6).

Overall accuracy in both conditions (H&E) is correlated with MD(H-E) ($r = 0.49, p = 0.0001$) when averaged across both runs (Figure 2.6). This prediction was consistent across runs; second run MD H-E predicts first run accuracy(H&E) ($r = 0.49, p = 0.0001$) and first run MD(H-E) predicts second run accuracy(H&E) ($r = 0.40p = 0.002$). All other correlations can be found in Table 2.2. Correlations of separate runs can be found in Table A.1.

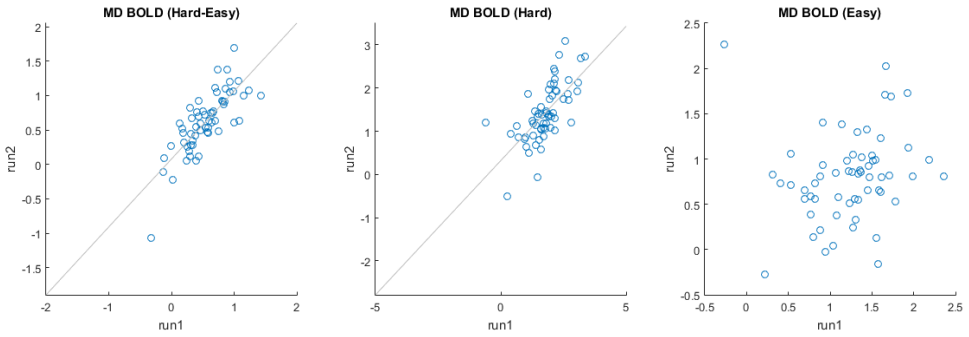


Figure 2.5: Stability of MD responses across runs. MD E response is not consistent across runs.

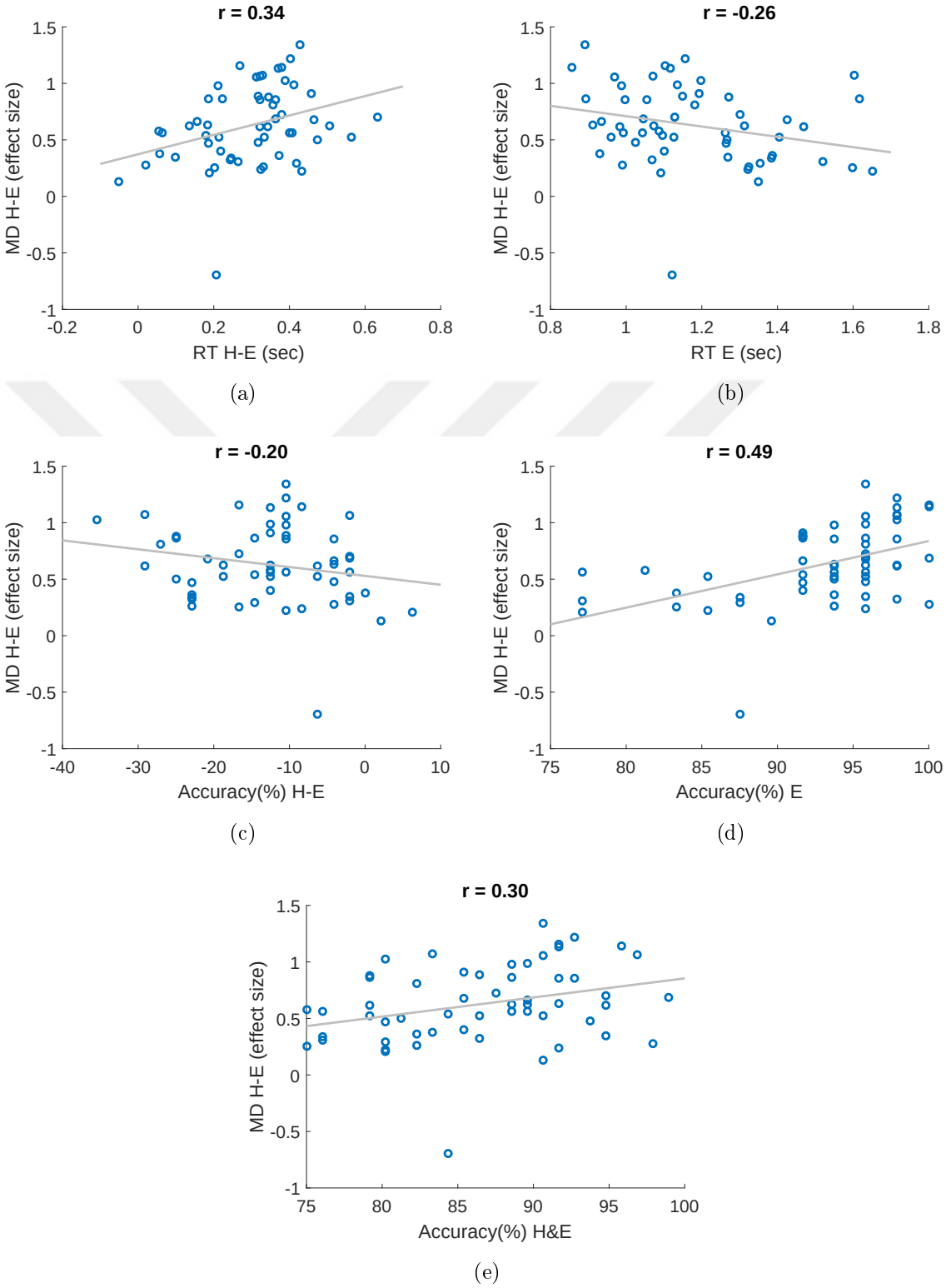


Figure 2.6: Prediction of behavioral performance by MD BOLD responses for the first dataset after excluding low performers (accuracy<75%) (n=56).

Table 2.2: MD BOLD prediction of behavioral performance.

		n = 60		(acc<75%) excluded (n = 56)	
		r	*p*	*r*	*p*
MD H-E vs	RT H-E	0.44	<0.01	0.34	0.01
	RT E	-0.30	0.02	-0.26	0.05
	RT H	0.01	0.97	-0.03	0.85
	RT H&E	-0.14	0.29	-0.13	0.33
	Acc H-E	-0.10	0.46	-0.20	0.14
	Acc E	0.60	<0.01	0.49	<0.01
	Acc H	0.33	0.01	0.11	0.42
	Acc H&E	0.49	0.00	0.30	0.02
MD H vs	RT H-E	0.35	0.01	0.20	0.13
	RT E	-0.18	0.16	-0.04	0.77
	RT H	0.05	0.71	0.08	0.58
	RT H&E	-0.06	0.66	0.03	0.85
	Acc H-E	0.00	0.98	-0.09	0.52
	Acc E	0.21	0.10	0.00	0.97
	Acc H	0.15	0.26	-0.08	0.55
	Acc H&E	0.19	0.14	-0.06	0.67
MD E vs	RT H-E	0.46	<0.01	0.33	0.01
	RT E	-0.29	0.03	-0.18	0.18
	RT H	0.03	0.81	0.03	0.83
	RT H&E	-0.12	0.38	-0.07	0.63
	Acc H-E	-0.06	0.68	-0.17	0.20
	Acc E	0.48	<0.01	0.30	0.03
	Acc H	0.28	0.03	0.02	0.90
	Acc H&E	0.40	<0.01	0.15	0.27

2.3.2.3 Replicability of Findings in a Second Dataset. We performed the same analysis on a second group of participants (n=59) who performed the same spatial WM task. The aim is to check the reliability of the findings from the first group of participants. We here only mention the main results that could be replicated with a strong or trending significance. The full set of the results is reported in (Table A.3)

For behavioral measures, RT H-E correlation with ACC H-E is trending for the second set ($r = -0.23$, $p = 0.08$). Correlation between accuracy(H&E) and accuracy(H-E) is also trending ($r = 0.22$, $p = 0.09$) but not significant for RT(H&E) correlation with RT(H-E) ($r = 0.18, p = 0.18$) though the trend is in the same direction as the first group results.

As for behavioral prediction using fMRI responses, MD(H-E) predicts RT(E) ($r = -0.4, p = 0.0017$) and RT(H-E) though with a trending significance ($r = 0.24, p = 0.06$). MD(H-E) could also predict accuracy(E) ($r = 0.29$, $p = 0.026$) and accuracy(H-E) ($r = -0.27$, $p = 0.04$) but not accuracy(H&E) ($r = 0.16$, $p = 0.24$) though again the trend is in the same direction (Figure 2.7).

In summary, all predictions were in line with the first group results. However, most are with a trending significance. This could be for several reasons such as different distribution of extreme performers in both groups. We elaborate on several interpretations in the Discussion section.

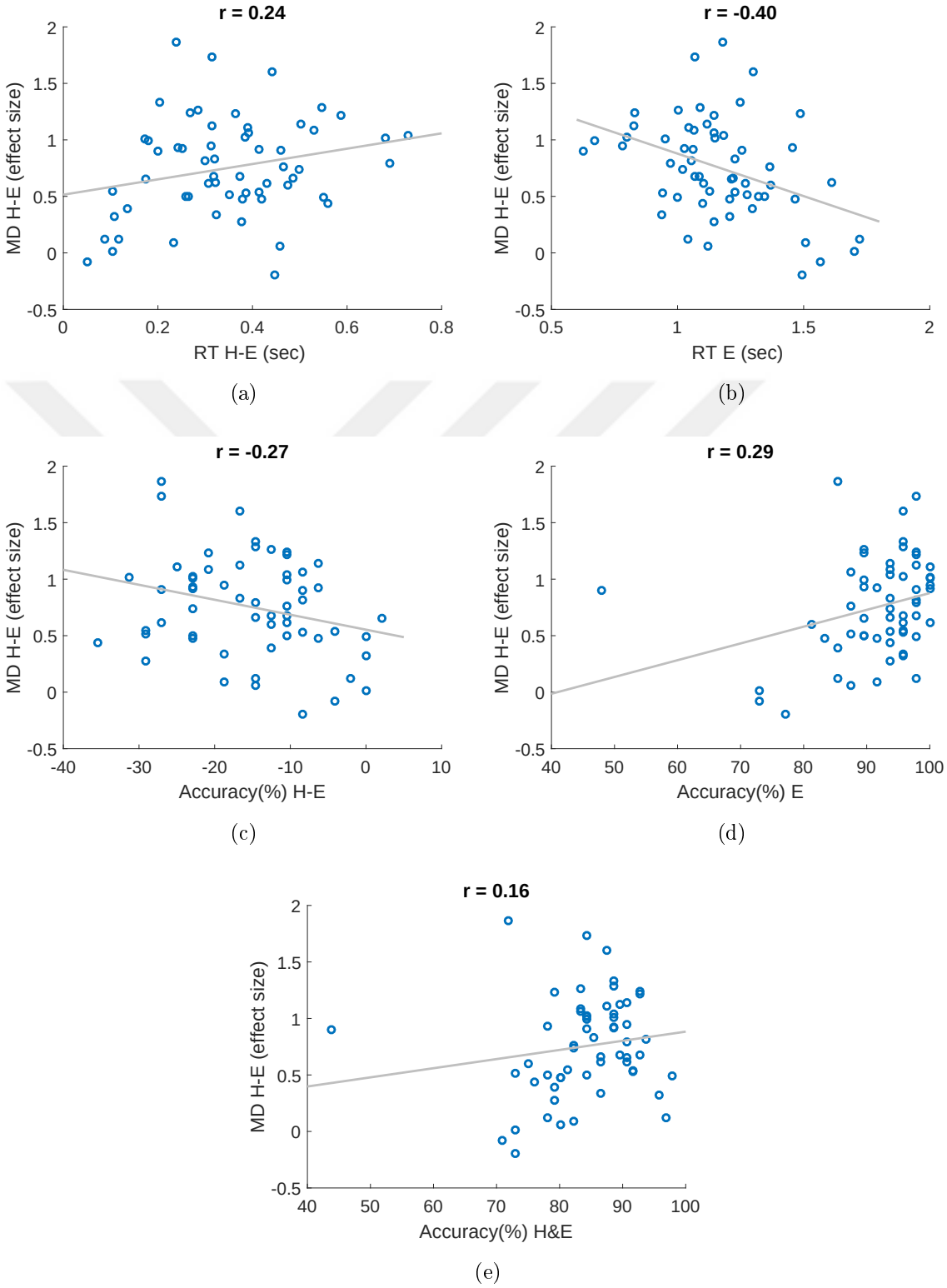
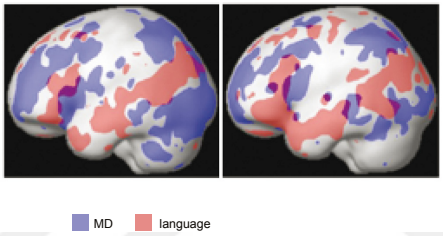


Figure 2.7: Prediction of behavioral performance by MD BOLD responses for the second dataset(n=59).

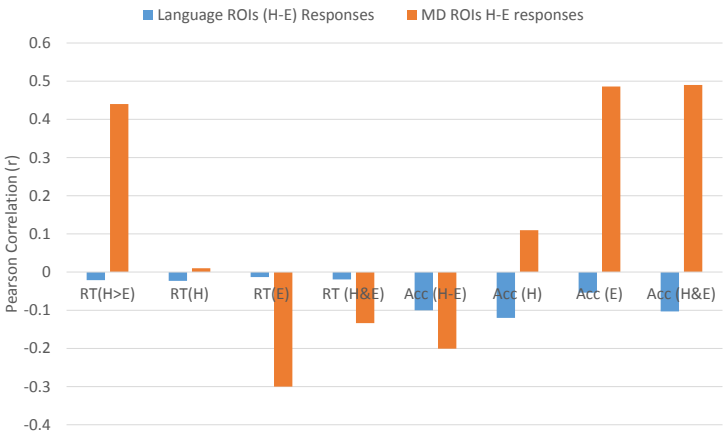
2.3.3 Validation using Language network fMRI responses

To test the validity of the predictions by MD regions, we repeated the same prediction analysis but using fMRI responses of the left hemisphere language network ROIs. The language ROIs were identified using the Sentences>Non-Words contrast. Their responses to the H and E trials of the spatial WM task were estimated as described before.

Language ROIs H-E contrast responses failed to predict any of the behavioral measures of the spatial WM task (Figure 2.8). Thus supporting the view that MD responses are indeed spatially localized to this network.



(a)



(b)

Figure 2.8: (a) Example of Language and MD ROIs from two participants adapted from [24] (b) Comparing correlations between responses of Language and MD ROIs vs Behavioral measures on the spatial WM task. None of the Language ROIs correlations is significant.

3. INTRA-INDIVIDUAL DIFFERENCES

3.1 Literature Review

3.1.1 Previous studies on Trial by Trial BOLD-Reaction Time Variability

Several brain regions have been consistently found to vary in activity with RT. However, the nature of this variation and the direction of causality is not well understood. The hemodynamic response (HDR) can vary with RT in one or more of the following ways: (1) change in HDR onset (2) change in HDR amplitude (3) change in HDR width.

One expected BOLD-RT relation is for longer RTs to be accompanied by a delayed HDR onset in sensorimotor brain regions because trials with longer RTs are accompanied by a delay in responses initiation. This finding has been observed repeatedly in several studies [35, 41, 42]. Interestingly, a recent study observed a delayed HDR onset in not only sensori-motor regions but also a wider network which includes MD regions [35]. However, the delay effects were more pronounced in parietal, visual and cerebellar regions. The relation between RT and HDR amplitude/width is more complex. Two different phenomena can be expected

1. As RT increases

   a) HDR amplitude increases: Reported in several MD regions, specifically lateral and medial PFCs and parietal, insular, SMA as well as primary sensory and motor areas [35, 36, 43]

   b)HDR width increases: Also observed in several MD regions especially the parietal lobes [41, 44]

2. As RT increases

   a) HDR amplitude decreases: This relation was observed primarily in ACC,

precuneus, PCC, angular and temporal regions [36, 42, 45]. This was also reported for the BOLD signal before trials with long RT in ACC and PFC [46] or right after the RT onset, within 2 seconds, in almost all MD regions [35]. It is worth noting that in the latter study these regions proceed to exhibit a positive correlation between RT and HDR amplitude after 2 sec from RT onset.

b) No studies report a decrease in HDR width

There are several possible interpretation of the previous findings. A positive correlation between RT and BOLD responses could be reflecting basic properties of the BOLD signal rather than cognitive demands. This is because BOLD signal sums approximately linearly over short durations [48]. Thus, the longer a person sustains attention on a task, the larger the expected BOLD response should be. This falls in line with the positive correlations observed across the fronto-parietal attention network, which is a part the MD network, supporting its role as an attention control network that is engaged regardless of the type of the task at hand and should covary with RT [47]. However, even after regressing out RT, activations are minimally affected and remain significant in this network [36, 37].

A negative correlation between BOLD and RT could indicate increased allocation of cognitive resources to solve the task at hand and respond faster. In this case, the negative correlation should be observed at task specific regions. However, as mentioned above, several of these regions are domain general such as the ACC or belong to the default mode network (DMN) such as precuneus and angular gyrus.

In an effort to reconcile this complex picture, an engagement-effort framework has been proposed to interpret RT-BOLD relations [36]. The framework suggests that if a region is tuned to respond to specific stimulus property, then its activity should not be expected to covary with RT. Further, if the region is less tuned yet still recruited, then this constitutes an "effort" for that region and such activity should covary with RT.

Through such a framework, MD regions covariation with RT could be interpreted: As the MD system encodes task-relevant properties, the easier the task, the faster the MD regions will represent it. While harder tasks would require to be "decomposed" [23] before being encoded. In this case, the decomposition (effort) period will be related with RT. Thus regressing out RT would remove the decomposition related activity but leave out the task representation activity. This interpretation has implications against the practice of regressing out of RT to remove "time on task" effects. While it might be true to some extent that BOLD-RT positive correlation could be reflecting basic BOLD properties, regressing out RT could be eliminating processing related to task decomposition.

3.2 Methodology

3.2.1 BOLD Time-series Extraction

Experiment and ROIs definition details are the same as explained in section 3.1.1. However, for this analysis, fROIs were defined using both runs (i.e most stable voxels in both runs). BOLD time series were then extracted from each voxel for each run separately. Each time series was highpass filtered (200s), de-trended, and white matter and CSF mean signals were regressed out.

3.2.2 Fusion of BOLD and Reaction Time

There are two main methods to model RT with respect to the BOLD signal, one which is based on assumptions of how RT should influence the BOLD signal and the other is an empirical approach with no assumptions.

For all the following models, RT onset is considered as the first appearance of the two-alternative forced-choice screen (Figure 2.2 (A)).

3.2.2.1 Variable Impulse Model. Variable Impulse (VarImp) (also known as parametric modulation) is based on the assumption that RT is related to BOLD amplitude. RT is modeled as a series of impulse functions at RT onset then convolved with the Hemodynamic Response Function (HRF) (Figure 3.1). The length of RT modulates the amplitude of the impulse function such that longer RTs will be modeled with a larger BOLD amplitude when convolved with HRF. The RTs are z-scored for each individual across all trials of the same run.

The VarImp approach is believed to sufficiently model short RTs (<4s) [48], which is the case in this experiment (max RT 2-3s). However, the approach is criticized as it only assumes RT modulates HRF amplitude while it ignores its effect on HRF width as will be described in the next section [49].

3.2.2.2 Variable Epoch Model. The variable epoch (VarEp) approach was proposed as an alternative to the VarImp to take into account the effect of RT on HRF width. RTs are modeled as a series of box cars (epochs) starting at the RT onset. Each epoch varies in width corresponding to the trial's RT duration. The RTs are not z-scored as negative RTs would shift the epoch before the onset time on interest. The epochs are then convolved with the HRF (Figure 3.1).Thus the VarEp regressor will model the linear relation between RT and HRF width.

Grindband et al demonstrated that the VarEp approach has better statistical power and more consistent results [49]. However, it is based on the assumption the cognitive processes differ in length but not in amplitude. Further, over stimuli with short period (<4s), it might decrease the variance in the modeled signal (as HRFs will be summed up) and lead to loss of statistical power.

3.2.2.3 Finite Impulse Response Model. Finite Impulse Response (FIR) is an empirical approach with no assumption on how RT relates to BOLD signal. Each scan time point is modeled with a single impulse function. Regressors are constructed as

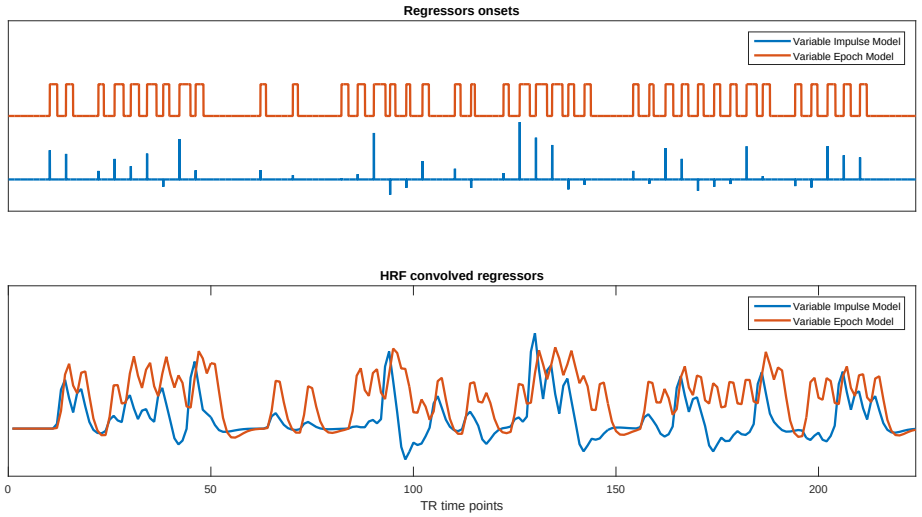


Figure 3.1: VarImp (blue) and VarEp (orange) regressor onsets (top) and after convolving with canonical HRF(bottom).

a series of impulse functions for the time point of interest. Since each time point is modeled, the regressors are not convolved with the HRF. In this case, each regressor parameter estimate corresponds to one time point.

In this experiment, each trial is 4 TRs in length thus each trial is modeled using 4 FIR "windows". The RT onset mostly coincides with the 3rd FIR point. Thus to incorporate RT information into the model, we modulated the height of 3rd FIR time point on each trial with z-scored RTs of that run. (Figure 3.2)

3.2.3 Model Estimation

3.2.3.1 Design Matrix. Data is modeled and estimated using the General Linear Model (GLM) approach [50, 51]. GLM in matrix formulations is:

$$\overrightarrow{Y}_{i\times 1} = X_{i\times j}\overrightarrow{\beta}_{j\times 1} + \overrightarrow{\epsilon}_{i\times 1} \quad for \quad \epsilon \sim N(0,\sigma^2) \tag{3.1}$$

where Y is a vector of BOLD values at each TR (i), X is the design matrix representing the model regressors, β is the vector of the model parameters to be estimated (one for

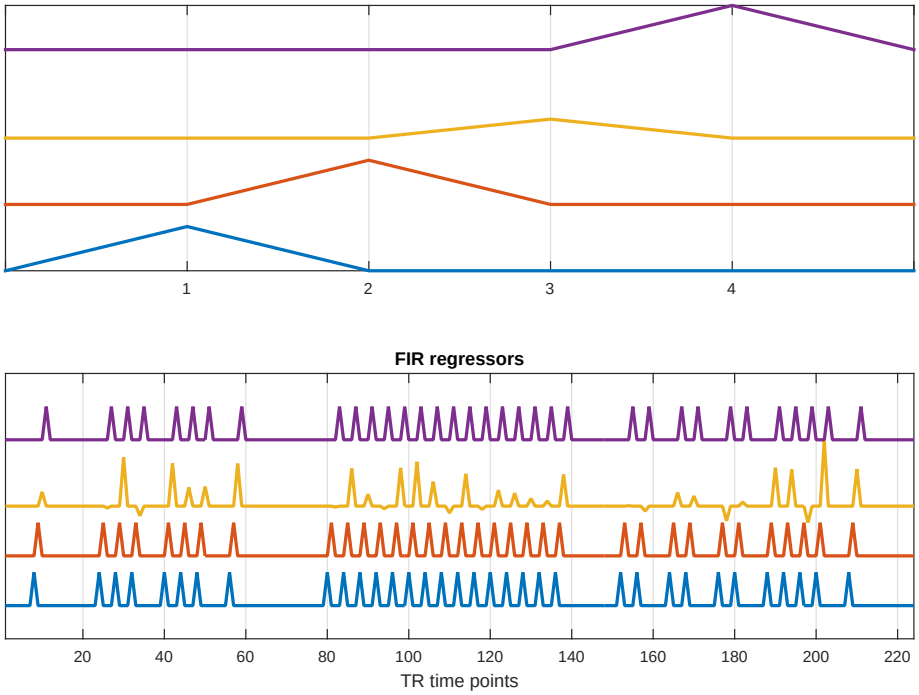


Figure 3.2: (Top) FIR modeling of a single trial (Bottom) four FIR regressors. Third regressor from below corresponds to third TR and is modulated by RT. Regressors are not convolved with HRF.

each regressor) and ϵ represents the residual errors.

For the VarImp and VarEp models the design matrix consists of four regressors: The first two are boxcar regressors convolved with HRF and its time derivative (TD) and the second two are RT regressors (VarImp or VarEp convolved with HRF and its TD. The TD of the HRF regressor models voxels which vary in temporal onset. Though parameter estimates from the TD regressor are not the focus of this study, it was included to improve model estimation and avoid large error variance that can undermine the statistical power of parameters of interest.

The regressors were then serially orthogonalized as follows Task reg > Task TD reg > RT reg > RT TD reg so that task regressors capture mean task(also mean RT) variance leaving RT-related variance to be captured by RT regressors.

As for FIR model, the design matrix consisted of four regressors, each corresponding to one of the four time points of a single trial. Regressors were already orthogonal and uncorrelated as they represent separate time points and no HRF was convolved.

In all models, correct trials only were included. Incorrect trials were not modeled. six motion regressors were further added to the design matrix to capture signal related to head movement and a mean regressor of ones to capture mean BOLD signal.

3.2.3.2 Parameters Estimation. Under the assumption that the residual errors are independent and identically distributed (i.i.d.), model parameters can be estimated using the Ordinary Least Squares (OLS) approach from Eq. 3.1:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y \tag{3.2}$$

However, since fMRI data points are correlated from one scan to the other, residual errors are not i.i.d. such that

$$\epsilon \sim N(0, \sigma^2 V) \quad V \neq I$$

where V is the temporal autocorrelation matrix $N_t \times N_t$ (N_t is the total number of TRs). Thus, β_{OLS} estimates will be biased as the degrees of freedom (df) will be lower than in the case of i.i.d. errors. To correct for this bias, the error covariance matrix C_ϵ is estimated and used to find Maximum Likelihood (ML) parameter estimates through pre-whitening of the data and design matrix as follows:

$$\hat{\beta}_{ML} = (X^T C_\epsilon^{-1} X)^{-1} X^T C_\epsilon^{-1} Y \quad where \quad C_\epsilon = \sigma^2 V \tag{3.3}$$

The resulting errors are i.i.d and parameter estimates do not require df correction.

The first step to estimate C_ϵ is to find an appropriate model for it. A commonly used model is the first order autoregressive process + white noise ($AR(1) + Wn$) and

has been demonstrated to sufficiently model the error autocorrelation process [51]

$$\overrightarrow{\epsilon} = A\overrightarrow{\epsilon} + \overrightarrow{w} \tag{3.4}$$

Thus the estimation of C_ϵ can be divided into two parts: (1) Estimate autocorrelation matrix V reflecting the AR(1) part and (2) error variance σ^2 of the white noise process.

To estimate V, a common assumption is that it has a similar structure across all voxels of interest. That is, the pattern of autocorrelation is the same but the error variance is different for each voxel (i.e the amplitude of autocorrelation is different). Based on this assumption, all time series from voxels of interest are pooled together to estimate V.

However, based on calculations of data from this experiment, we found that this assumption does not hold for all ROIs. So time series was pooled from each ROI to estimate a ROI specific V. The pooled data are used in the form of their temporal covariance matrix to estimate V as follows

$$C_Y = \frac{YY^T}{N_v}$$

where Y is the $N_t \times v$ data matrix (v is the number of voxels, N_t is the number of time points). N_v is the total number of voxels in the ROI. The data covariance matrix C_Y is composed of two variances, the experiment related variance and the error variance component.

V can be expressed using a linear combination of known covariance components scaled by unknown hyperparameters, as follows

$$V = \sum_i \lambda_i Q_i$$

where Q_i are the $N \times N$ covariance components modeling the autocorrelation and λ_i are the hyperparameters to be estimated. Assuming the autocorrelation process follows

the $AR(1) + Wn$ model,

$$V = \lambda_1 Q_1 + \lambda_2 Q_2 \tag{3.5}$$

where component Q_1 represents white noise variance (I_N) and component Q_2 represents the $AR(1)$ process with an unspecified 1st order coefficient.

The use of covariance components allows V to be estimated using a one step Restricted Maximum Likelihood (ReML) to give an unbiased estimator of the two hyperparameters λ hence giving the estimated V (cf. Eq. 3.5).

$$V = ReML(C_Y, Q_I, X)$$

The ReML method estimates the hyperparameters using an Expectation-Maximization algorithm described in [52].

The next step is to estimate the error variance σ^2 which is also called the voxel specific hyperparameter. From equations 3.1 and 3.2

$$\overrightarrow{\epsilon} = [I - X(X^T X)^{-1} X^T]Y = RY \tag{3.6}$$

and from equation 3.4

$$\overrightarrow{\epsilon} = (I - A)^{-1}\overrightarrow{w} \tag{3.7}$$

$$\epsilon\epsilon^T = (I - A)^{-1} w w^T (I - A)^{-T} = (I - A)^{-1}\sigma^2 I (I - A)^{-T}$$

$$C_\epsilon = \sigma^2 (I - A)^{-1}(I - A)^{-T}$$

and since $C_\epsilon = \sigma^2 V$, thus

$$V = (I - A)^{-1}(I - A)^{-T}$$

$$(I - A) = V^{-1/2} \tag{3.8}$$

From Eq. 3.6, 3.7 and 3.8

$$\overrightarrow{w} = (I - A)\overrightarrow{\epsilon} = V^{-1/2}RY$$

$$\sigma_v^2 = variance(\overrightarrow{w})$$

To summarize, C_ϵ was assumed to follow an AR(1)+Wn model, which can be modeled using covariance components. Thus, its estimation was done in two steps: (1) Estimation of the autocorrelation matrix V, corresponding to the AR(1) part, for each ROI using ReML (to estimate the hyperparameters of the covariance components) (2)Estimation of the error variance or voxel specific hyperparameter σ_v^2. C_ϵ was then used to pre-whiten the data and design matrix returning ML parameter estimates as shown in Eq. 3.3.

3.2.4 Hierarchical Modeling

We then performed a random effects analysis (RFX) using the summary statistic approach [52] which applies a two stage model as follows

$$Y = X^{(1)}\beta_i^{(1)} + \epsilon^{(1)} \quad 1st \quad level$$

$$\beta_{all}^{(1)} = X^{(2)}\beta^{(2)} + \epsilon^{(2)} \quad 2nd \quad level$$

$\beta_{all}^{(1)}$ is a vector of each participant's β_i of interest from the 1st level model, $X^{(2)}$ is the 2nd level design matrix, $\beta^{(2)}$ are the 2nd level parameter estimates, $\epsilon^{(2)}$ represent 2nd level residual errors.

The β_i of interest in each model is the β corresponding to VarImp regressor, the VarEp regressor or the 3rd regressor in the FIR model. These βs where averaged across ROIs to give a single value for each individual.

The summary statistic approach assumes that the second level non-sphericity

(variance components) has the same form as the first level. This approach is valid in this experiment because (1) We used the same first level model for all participants (2) only one contrast is imported from the 1st level (3) We will perform a one-sample t-test (i.e. using only one group). Under these conditions, the contribution of non-sphericity at the second level can be ignored [52].

By using dummy codes (ones) as a single regressor in the design matrix $X^{(2)}$, $\beta^{(2)}$ can be estimated using OLS,

$$\hat{\beta}^{(2)} = (X^{(2)^T} X^{(2)})^{-1} X^{(2)^T} \beta^{(1)}_{all}$$

then the one sample T test,

$$T = \frac{\hat{\beta}^{(2)}}{\sqrt{var_{pop}}} \quad for \quad var_{pop} = var_{with} + var_{bet}$$

where var_{pop} is the population variance which consists of within-subject variance var_{with} from the first level,

$$var_{with} = \sigma_v^2 (V X^{(1)^-})(V X^{(1)^{-T}}) \quad for \quad V X^{(1)^-} = (X^{(1)^T} V X^{(1)})^{-1} X^{(1)^T}$$

and between-subject variance var_{bet} of the second model,

$$var_{bet} = \sigma_b^2 (X^{(2)^T} X^{(2)})^{-1} \quad for \quad \sigma_b^2 = \frac{\epsilon^{(2)^T} \epsilon^{(2)}}{S - rank(X^{(2)})}$$

where S is the length of vector $\beta^{(1)}_{all}$ corresponding to number of participants.

4. RESULTS AND DISCUSSION

In this thesis, we investigated the neural basis underlying individual differences (ID) in executive functions (EFs). Specifically, we investigated fMRI responses of a Multiple Demand (MD) brain network, closely related to EFs [23], and behavioral variability during a spatial working memory (WM) task. we investigated this relation through two approaches: (1) inter-individual differences (variation on an individual-by-individual basis) (2) intra-individual variations (variation on a trial-by-trial basis). This study investigated a large fMRI dataset (n=120). The spatial WM task consisted of easy and hard trials presented in a block design. Each subject was scanned for two runs. We used the same dataset for both approaches analysis.

The strength of this study stems from (1) overcoming several methodological challenges related to the reliability and validity of fMRI responses and its prediction power of the behavioral measures (Table 4.1) (2) complementing inter-individual findings with intra-individual variability analysis. For this thesis, we only investigated the intra-individual variability between BOLD and RT.

4.1 Inter-Individual differences

The results demonstrate a high degree of stability of MD activity within individuals across time. Specifically the H-E and H-fix responses are stable even at the individual ROIs level (Figures A.2 and A.3). However, E-fix responses are not consistent across runs. This could be due to the small variance of the E-fix responses, yet 9 out of the 18 ROIs still show consistent responses (Figure A.4). This consistency serves as the first confirmation of the validity of MD fMRI responses.

Behavioral measures, both RT and accuracy, were also relatively stable within individuals over time. The stability of neural and behavioral measures allowed the averaging of both measures across runs to give more robust results. More accurate

Table 4.1: Methodological challenges in fMRI studies of individual differences addressed in this thesis.

Limitations of previous studies	**Solutions implemented in this study**
Small sample sizes (n=15-20)	Large dataset (n=120)
Validity of MD fMRI responses	1) Used Hard-Easy contrast eliminates subjects' states confounds 2)Used Individual-specific fROIs (Figure 2.2 (B)) 3) Stability of MD responses across runs (Figures 2.5 and A.2) 4) Validation using language ROIs responses (Figure 2.8)
Reliability of findings	1) Replicability across both runs (Table A.1) 2) Prediction of behavioral measures of run 2 using BOLD measures of run 1 (Table A.2) 3) Testing results replicability on a second dataset (Table A.3)
Use of Extreme groups	Participants were not grouped into high/low performers
Lack of intra-individual analysis	Complementary intra-individual analysis performed supporting findings from inter-individual analysis

participants were faster on both the easy or hard conditions.

In this study, we focus on the H-E contrast for the validity concerns mentioned earlier (see Section 2.1.3). Thus, the use of this contrast allows a better interpretation of the results from previous fMRI studies; such that instead of discussing whether a participant has, on average, higher or lower brain activity the interpretation now examines whether a participant needs to engage their MD system more or less on a current task of high or low cognitive demand.

That said, the H-E is a difference index which means it is influenced by its two conditions. Thus, for a valid interpretation, it is necessary to determine whether one condition contributes to the H-E index more than the other. Figure 4.1 demonstrates that the hard condition of the task is the main driving factor behind the H-E index for all: MD BOLD, RT and accuracy measures

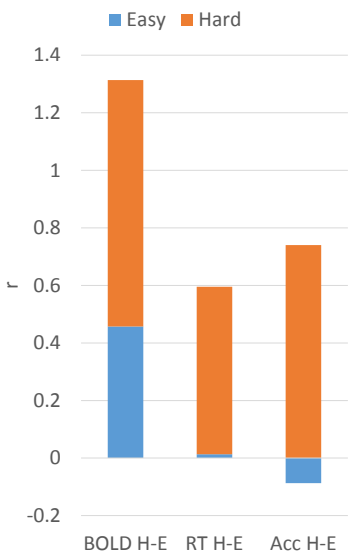


Figure 4.1: Correlations of H-E MD BOLD, RT and accuracy with each of their Easy (blue) and Hard (orange) measures.

Examining MD responses prediction of performance, we find that better performers on the easy task are also the ones who have the largest increase in MD activity (MD(H-E) vs ACC(E) positively correlated in both datasets). In other words, participants who found the hard task cognitively challenging are the ones who had the largest demand in MD activity.

This also suggests that low performers are not the ones with the strongest neural activity. This is further supported by the observation that participants who performed better on the hard task relative to the easy had the largest increase in MD activity (MD(H-E) vs accuracy(H-E) is negatively correlated in both datasets i.e., as the gap in accuracy between the easy and hard conditions gets closer, the larger is the demand in MD activity).

The same relation holds for RT variability. Faster participants on the easy task are the ones who have the largest surge in MD activity (MD(H-E) vs RT(E) negatively correlated in both datasets).

Moreover, participants with strongest demand in MD activity also have a large RT H-E (MD(H-E) and RT(H-E) are positively correlated in both datasets). One interpretation could be that this reflects the "time on task" effect previously mentioned [35] (i.e. that BOLD increase is reflecting sustained attention during duration of the task rather than task-relevant cognitive processing). If this was the case, one would expect the same positive correlation between BOLD and RT for their separate conditions. This is only true for the easy condition (Table 2.2, though the effect is abolished when low performers are excluded) and since the H-E index is largely driven by the H scores then it renders the "time on task" interpretation less likely. Moreover, there is no correlation between MD H-E and overall RT (H&E) lending further support against the time on task effect.

To further validate these findings, we tested the prediction power of MD BOLD and behavioral measures of opposite runs (Table A.2). MD BOLD measures of second run successfully predicted behavioral measures of the first run with similar results as

before. However, first run BOLD failed to predict behavior of the second run, except for the MD H-E vs overall accuracy (H&E) correlation, which is consistent with previous results. The non-replicablity of the accuracy H-E index could be attributed to its smaller variance on the second run (see Section 2.3.1) as subjects learn the task better and become more accurate. However, the RT H-E variance on the second run is not smaller than the first run. Thus this non-replicability remains to be explained.

Together these findings suggest that the MD system activity highly reflects the cognitive demand of the participants for the spatial WM task. The main goal, however, is to extrapolate these findings to explain more stable individual differences in EFs. While this can be best revealed by contrasting independent measures of EFs with neural or behavioral activity from this task, we speculate about this relation based on our results.

We start with the assumption that MD H-E belongs to individuals with better EFs abilities. The change in cognitive demand for them requires minimal change in neural activity. One opposing finding is the positive correlation between MD H-E and overall accuracy on the task. If smaller MD H-E values reflect better performers, then a negative correlation should have been expected. However, as previously mentioned, large MD H-E difference is less likely to reflect a bad performer but individuals who find this task challenging. Thus, individuals with a large MD(H-E) will engage their MD system more to achieve the same accuracy.

One supporting finding is the opposite correlation between MD(H-E) vs accuracy(H&E) and MD(H-E) vs accuracy(H-E). The positive correlation between MD H-E and overall accuracy should make us expect a positive correlation as well between MD(H-E) and accuracy(H-E) yet we find a negative correlation. If accuracy(H-E) taps into EFs better than accuracy(H&E) then this negative correlation predicts that individuals with small MD(H-E) will have better executive abilities. It could be argued that the positive correlation between accuracy(H&E) and accuracy(H-E) discredits the idea that they tap into different variance of behavior. However, it follows that a small accuracy(H-E) will reflect more correct trials solved on the task but it doesn't reveal

whether an individual struggled or not to achieve this high accuracy. Thus, what sets these two measures apart is their opposite correlation with MD(H-E).

In summary, BOLD responses in the MD system are consistently stable within individuals but variable between them. This variability explains a significant proportion of their performance on the spatial WM task. Specifically, BOLD responses of the MD system are sensitive to the challenging cognitive demands as participants who find the task difficult are the ones with largest increase in MD activity. The larger the leap in MD activity, the better the individual performs on the hard condition of the task yet the slower they become. The findings also suggest that MD activity is sensitive to learning effects. However, this is beyond the scope of this thesis.

4.2 Intra-Individual differences

In this analysis, we investigated the trial-by-trial variability between BOLD and RT. The aim is to test whether MD-RT findings between individuals is also replicable within individuals. For the scope of this thesis, we am only investigating RT at the intra-individual level but not accuracy.

To perform this analysis, we tested three different models to fuse RT with the BOLD time-series. Two models are based on assumptions: one expects RT to be related to BOLD amplitude (VarImp model), while the other expects RT to be related to BOLD HDR width (VarEp model). The third model is not based on assumptions but rather estimates relation of RT at a each time point of the BOLD signal(FIR model). We tested the three models on each of the 60 participants and on each run separately, to replicate the findings in both runs lending more confidence in the results.

4.2.1 Results of RT Regressors

1. **Variable Impulse model** There is a positive relation between BOLD amplitude and RT which is replicable across runs (t_{run1} = 2.21, p=0.015, t_{run2}=2.68, p= 0.0048).

2. **Variable Epoch model** No significant relation between BOLD width and RT for both the first run (t=-0.077, p = 0.53) and the second run (t=-0.22, p = 0.59).

3. **Finite Impulse Response model** Modulating the amplitude of the third FIR time point also showed no significant relation between BOLD amplitude and RT for the first run (t=0.046, p=0.48) and the second run (t=0.004, p=0.50)

4.2.2 Discussion

The main result comes from the analysis using the VarImp model: the amplitude of the BOLD signal increases as a function of RT. This result is consistent across both runs. This is in line with previous studies reporting positive correlations between fronto-parietal regions and trials RT [35, 36].

The VarEp model results were not significant on either of the two runs. In other words, RT is not significantly related to BOLD HDR width. One reason could be the short RTs in this study. BOLD width is thought to increase for trials with RTs longer than 4 seconds [48, 49]. Further, modeling close trials with box-cars can decrease the variance of the regressor. In other words, due to the HDR latency, there is not enough time for the modeled signal to decrease in value before the next stimulus, thus the signal has more sustained "activations" (cf. Figure 3.1). This might cause a considerable loss of statistical power which might explain why the results of this model were not significant.

The FIR model results were also non significant across both runs. This was not

an expected result. It could be however due to the small duration between successive RT onsets which limited the FIR model to only four windows. If longer FIR windows could be used, perhaps differences due to RT could be detected later. This is indeed plausible as a large scale study on RT-BOLD relation using FIR modeling only finds significant RT effects after the first 3 seconds of RT onset [35]. The time window we modulated in this study is only 2 seconds after RT onset.

In summary, within-individuals, trial-by-trial BOLD signal increases in amplitude but not in width on trials with longer RTs.

4.3 Convergence of Inter- and Intra-individuals differences

The results from both approaches regarding RT converge. At an individual-by-individual level, participants who require a large increase in their MD activity for a difficult task have the longest RTs. At a trial-by-trial level, slower trials were correlated with large BOLD amplitude.

This convergence serves two purposes:

1. It reinforces the use of **inter-individual differences (ID) findings to infer brain processes underlying EFs**. This is useful for (1) research topics where experimental manipulations are difficult to implement or (2) complementary support for a brain-behavior hypothesis based on experimental manipulations (3) as a bio-marker for patients diagnosis, specifically in the cases where patients are not able to reliably undertake an experiment.

2. **Inter and intra-individual effects are not spatially dissociable, at least for the MD network.** There are concerns on whether inter-and intra-individual findings should converge in the first place. On one hand, it is logical to expect that differences within an individual in one region should also be observed across individuals. This follows the observation that BOLD signal variations showing

intra-individual differences are more reliable in revealing inter-individual differences [53]. On the other hand, there is no reason not to expect a different picture. For example, Yarkoni et al. report that the fronto-parietal attention control network showed inconsistent inter-individual differences in response accuracy during a WM task. However, activity in a region not related to WM, the medial posterior parietal cortex, predicted inter-individual differences reliably [6]. In fact, one study recommended focusing on regions that do not show within-individual variability to detect inter-individual differences [54]. However, the convergence of findings in our study provides support to using ROIs that are identified based on intra-individual variability.

5. CONCLUSION AND FUTURE RECOMMENDATIONS

In this study, we investigated the relation between neural activity in a Multiple Demand (MD) network and individual differences (ID) on an executive functions (EFs) task.

A number of conclusions can be drawn:

1. MD activity varies substantially across individuals but is consistent within an individual across time.
2. MD activity can explain a significant proportion of variance in individual performance on the spatial WM task, such that individuals who find the task challenging, increase their MD activity substantially to improve their performance. This suggests that MD activity is sensitive to the executive demand of an individual.
3. RT variability within and across individuals shares the same relation with MD activity such that longer RT is correlated with stronger MD activity.

To these conclusions, this study addressed several methodological challenges concerning the reliability and validity of using fMRI for individual differences research. This paves the way for the use of fMRI in individual differences in clinical settings.

This study could be further improved by investigating yet a larger sample, employing more complex statistics and relating MD activity with individual measures of executive functions such as fluid intelligence scores.

Given that a significant proportion of individual differences in EFs is genetically determined, future work should attempt to link genetic variability to variability in the neural activity of the MD system.

APPENDIX A. DETAILED RESULTS

A.1 Individual MD ROIs responses

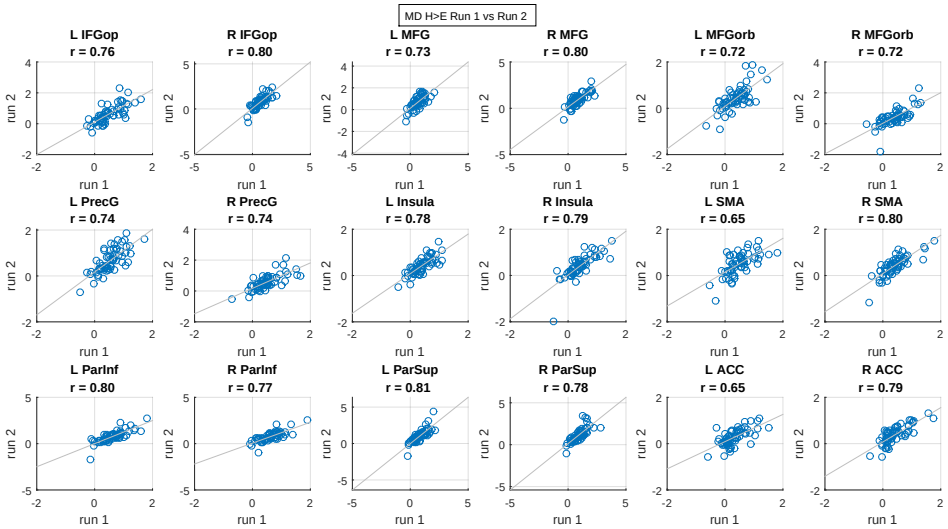


Figure A.1: Stability of individual ROIs responses H-E contrast.

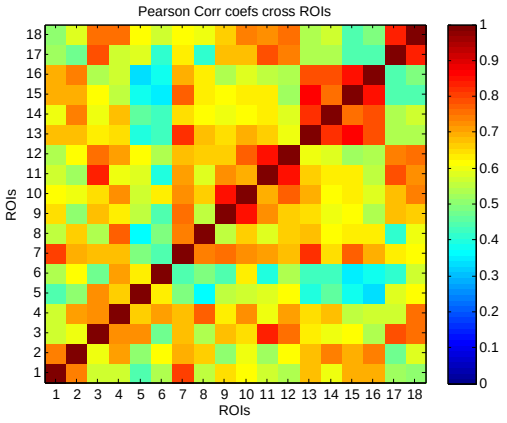


Figure A.2: High correlations between ROIs.

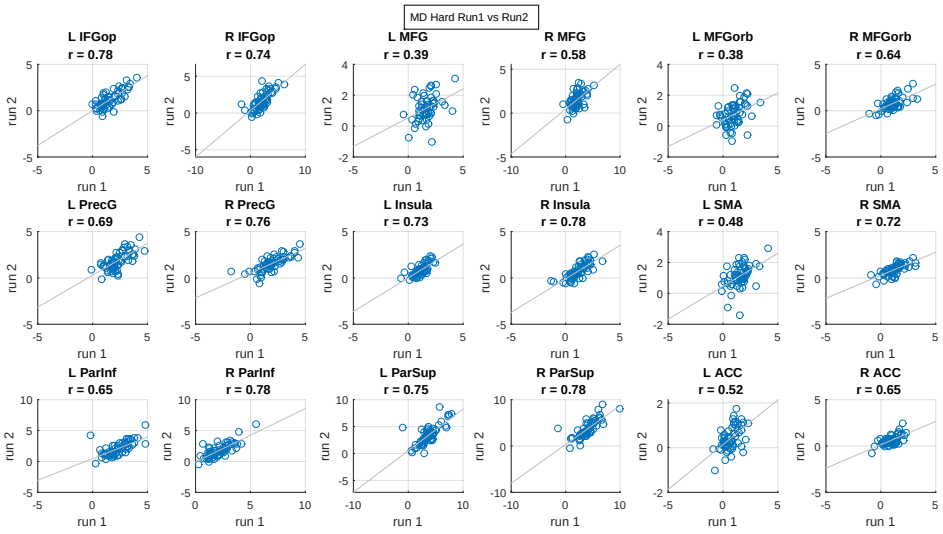


Figure A.3: Stability of individual ROIs responses across both runs (H-fix contrast).

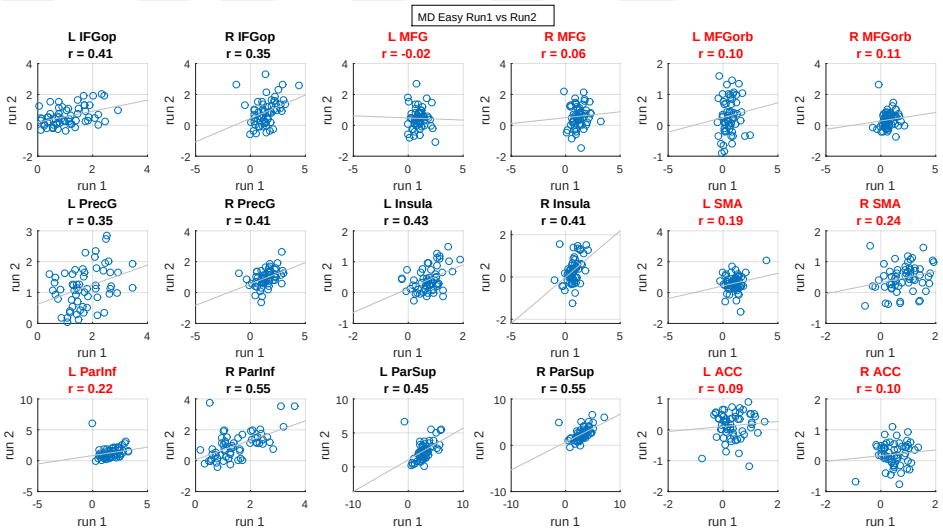


Figure A.4: Inconsistency of 9/18 individual ROIs responses across both runs for E-fix contrast.

A.2 MD BOLD prediction of behavioral performance

Table A.1: MD BOLD - Behavior predictions on separate runs (first n=60 group).

		run1		run 2		avg both runs	
		r	*p*	*r*	*p*	*r*	*p*
MD H-E vs	RT H-E	0.51	<0.01	0.30	0.02	0.44	<0.01
	RT E	-0.31	0.02	-0.22	0.08	-0.30	0.02
	RT H	0.06	0.64	0.02	0.86	0.01	0.97
	RT H&E	-0.11	0.40	-0.09	0.49	-0.14	0.29
	Acc H-E	-0.15	0.24	0.06	0.67	-0.10	0.46
	Acc E	0.62	<0.01	0.31	0.02	0.60	<0.01
	Acc H	0.28	0.03	0.26	0.05	0.33	0.01
	Acc H&E	0.48	<0.01	0.31	0.01	0.49	<0.01
MD H vs	RT H-E	0.34	0.01	0.20	0.13	0.35	0.01
	RT E	-0.20	0.12	-0.12	0.37	-0.18	0.16
	RT H	0.05	0.73	0.04	0.79	0.05	0.71
	RT H&E	-0.07	0.60	-0.03	0.80	-0.06	0.66
	Acc H-E	-0.10	0.43	0.09	0.51	0.00	0.98
	Acc E	0.38	<0.01	-0.06	0.65	0.21	0.10
	Acc H	0.16	0.21	0.03	0.81	0.15	0.26
	Acc H&E	0.29	0.02	-0.01	0.96	0.19	0.14
MD E vs	RT H-E	0.48	<0.01	0.34	0.01	0.46	<0.01
	RT E	-0.28	0.03	-0.23	0.07	-0.29	0.03
	RT H	0.06	0.64	0.04	0.75	0.03	0.81
	RT H&E	-0.10	0.45	-0.08	0.52	-0.12	0.38
	Acc H-E	-0.15	0.27	0.10	0.45	-0.06	0.68
	Acc E	0.56	<0.01	0.16	0.21	0.48	<0.01
	Acc H	0.25	0.06	0.19	0.14	0.28	0.03
	Acc H&E	0.43	<0.01	0.20	0.12	0.40	<0.01

Table A.2: MD BOLD - Behavior predictions across opposite runs (first n=60 group).

		BOLD run 1 vs BHV run2		BOLD run 2 vs BHV run 1	
		r	*p*	*r*	*p*
MD H-E vs	RT H-E	0.18	0.18	0.43	<0.01
	RT E	-0.23	0.08	-0.33	0.01
	RT H	-0.06	0.63	-0.01	0.96
	RT H&E	-0.14	0.27	-0.16	0.22
	Acc H-E	0.01	0.96	-0.17	0.20
	Acc E	0.43	<0.01	0.64	<0.01
	Acc H	0.30	0.02	0.28	0.03
	Acc H&E	0.40	<0.01	0.49	<0.01
MD H vs	RT H-E	0.25	0.06	0.12	0.35
	RT E	-0.08	0.55	-0.14	0.30
	RT H	0.10	0.46	-0.04	0.78
	RT H&E	0.02	0.87	-0.09	0.51
	Acc H-E	0.10	0.43	-0.03	0.80
	Acc E	0.14	0.30	0.09	0.49
	Acc H	0.18	0.17	0.03	0.82
	Acc H&E	0.18	0.17	0.06	0.63
MD E vs	RT H-E	0.25	0.05	0.38	<0.01
	RT E	-0.16	0.21	-0.32	0.01
	RT H	0.04	0.79	-0.03	0.82
	RT H&E	-0.06	0.67	-0.17	0.19
	Acc H-E	0.07	0.58	-0.14	0.30
	Acc E	0.30	0.02	0.49	<0.01
	Acc H	0.27	0.04	0.21	0.11
	Acc H&E	0.31	0.01	0.37	0.00

Table A.3: MD BOLD-Behavior predictions (second dataset n=59).

		r	p
MD H-E vs	RT H-E	0.24	0.06
	RT E	-0.4	0.0017
	RT H	-0.21	0.10
	RT H&E	-0.32	0.014
	Acc H-E	-0.27	0.041
	Acc E	0.29	0.026
	Acc H	0.01	0.93
	Acc H&E	0.16	0.24
MD H vs	RT H-E	0.22	0.1
	RT E	-0.2	0.13
	RT H	-0.05	0.71
	RT H&E	-0.13	0.33
	Acc H-E	-0.29	0.03
	Acc E	0.26	0.053
	Acc H	-0.03	0.81
	Acc H&E	0.11	0.41
MD E vs	RT H-E	0.11	0.40
	RT E	0.07	0.61
	RT H	0.13	0.33
	RT H&E	0.1	0.43
	Acc H-E	-0.19	0.15
	Acc E	0.12	0.36
	Acc H	-0.06	0.64
	Acc H&E	0.02	0.86

REFERENCES

1. Friedman, N. P., A. Miyake, S. E. Young, J. C. Defries, R. P. Corley, and J. K. Hewitt, "Individual Differences in Executive Functions Are Almost Entirely Genetic in Origin," *Journal of Experimental Psychology*, Vol. 137, no. 2, pp. 201–225, 2009.

2. Ando, J., Y. Ono, and M. J. Wright, "Genetic structure of spatial and verbal working memory.," *Behavior genetics*, Vol. 31, no. 6, pp. 615–624, 2001.

3. Meyer-Lindenberg, A., and D. R. Weinberger, "Intermediate phenotypes and genetic mechanisms of psychiatric disorders," *Nat Rev Neurosci*, Vol. 7, no. 10, pp. 818–827, 2006.

4. Braver, T. S., M. W. Cole, and T. Yarkoni, "Vive les differences! Individual variation in neural mechanisms of executive control," *Current opinion in neurobiology*, Vol. 20, no. 2, pp. 242–250, 2010.

5. Dubois, J., and R. Adolphs, "Building a Science of Individual Differences from fMRI," *Trends in Cognitive Sciences*, Vol. xx, pp. 1–19, 2016.

6. Matthews, G., J. S. Warm, L. E. Reinerman, and L. K. Langheim, *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control*, 2010.

7. Gazzinga, M. S., ed., *The Cognitive Neurosciences*, CAMBRIDGE, MASSACHUSETTS: The MIT press, 2009.

8. Engle, R. W., S. W. Tuholski, J. E. Laughlin, and a. R. Conway, "Working memory, short-term memory, and general fluid intelligence: a latent-variable approach.," *Journal of experimental psychology. General*, Vol. 128, no. 3, pp. 309–331, 1999.

9. Kane, M. J., and R. W. Engle, "The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective," *Psychonomic Bulletin& Review*, Vol. 9, no. 4, pp. 637–671, 2002.

10. Duncan, J., "A Neural Basis for General Intelligence," *Science*, Vol. 289, no. 5478, pp. 457–460, 2000.

11. Gray, J. R., C. F. Chabris, and T. S. Braver, "Neural mechanisms of general fluid intelligence.," *Nature neuroscience*, Vol. 6, no. 3, pp. 316–322, 2003.

12. Lee, K. H., Y. Y. Choi, J. R. Gray, S. H. Cho, J. H. Chae, S. Lee, and K. Kim, "Neural correlates of superior intelligence: Stronger recruitment of posterior parietal cortex," *NeuroImage*, Vol. 29, no. 2, pp. 578–586, 2006.

13. Wager, T. D., C. Y. C. Sylvester, S. C. Lacey, D. E. Nee, M. Franklin, and J. Jonides, "Common and unique components of response inhibition revealed by fMRI," *NeuroImage*, Vol. 27, no. 2, pp. 323–340, 2005.

14. Nee, D. E., T. D. Wager, and J. Jonides, "Interference resolution: insights from a meta-analysis of neuroimaging tasks.," *Cognitive, affective & behavioral neuroscience*, Vol. 7, no. 1, pp. 1–17, 2007.

15. Rypma, B., J. S. Berger, and M. D'Esposito, "The influence of working-memory demand and subject performance on prefrontal cortical activity.," *Journal of cognitive neuroscience*, Vol. 14, no. 5, pp. 721–31, 2002.

16. Osaka, N., M. Osaka, H. Kondo, M. Morishita, H. Fukuyama, and H. Shibasaki, "The neural basis of executive function in working memory: An fMRI study based on individual differences," *NeuroImage*, Vol. 21, no. 2, pp. 623–631, 2004.

17. duncan, J., "Brain mechanisms of attention," *Quarterly Journal of Experimental Psychology*, Vol. 59, pp. 2–27, Jan 2006.

18. Hon, N., "Frontoparietal Activity with Minimal Decision and Control," *Journal of Neuroscience*, Vol. 26, no. 38, pp. 9805–9809, 2006.

19. Dove, A., S. Pollmann, T. Schubert, C. J. Wiggins, and D. Yves Von Cramon, "Prefrontal cortex activation in task switching: An event-related fMRI study," *Cognitive Brain Research*, Vol. 9, no. 1, pp. 103–109, 2000.

20. Li, S., D. Ostwald, M. Giese, and Z. Kourtzi, "Flexible coding for categorical decisions in the human brain.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, Vol. 27, no. 45, pp. 12321–30, 2007.

21. Woolgar, a., a. Hampshire, R. Thompson, and J. Duncan, "Adaptive Coding of Task-Relevant Information in Human Frontoparietal Cortex," *Journal of Neuroscience*, Vol. 31, no. 41, pp. 14592–14599, 2011.

22. Duncan, J., "The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour," *Trends in Cognitive Sciences*, Vol. 14, no. 4, pp. 172–179, 2010.

23. Duncan, J., "The Structure of Cognition: Attentional Episodes in Mind and Brain," *Neuron*, Vol. 80, no. 1, pp. 35–50, 2013.

24. Fedorenko, E., J. Duncan, and N. Kanwisher, "Broad domain generality in focal regions of frontal and parietal cortex.," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110, no. 41, pp. 16616–21, 2013.

25. Roca, M., A. Parr, R. Thompson, A. Woolgar, T. Torralva, N. Antoun, F. Manes, and J. Duncan, "Executive function and fluid intelligence after frontal lobe lesions," *Brain*, Vol. 133, no. 1, pp. 234–247, 2010.

26. Woolgar, A., A. Parr, R. Cusack, R. Thompson, I. Nimmo-Smith, T. Torralva, M. Roca, N. Antoun, F. Manes, and J. Duncan, "Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex.," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 107, no. 33, pp. 14899–14902, 2010.

27. Dosenbach, N. U. F., K. M. Visscher, E. D. Palmer, F. M. Miezin, K. K. Wenger, H. C. Kang, E. D. Burgund, A. L. Grimes, B. L. Schlaggar, and S. E. Petersen, "A Core System for the Implementation of Task Sets," *Neuron*, Vol. 50, no. 5, pp. 799–812, 2006.

28. Sigala, N., M. Kusunoki, I. Nimmo-Smith, D. Gaffan, and J. Duncan, "Hierarchical coding for sequential task events in the monkey prefrontal cortex.," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 105, no. 33, pp. 11969–11974, 2008.

29. Kusunoki, M., N. Sigala, D. Gaffan, and J. Duncan, "Detection of fixed and variable targets in the monkey prefrontal cortex," *Cerebral Cortex*, Vol. 19, no. 11, pp. 2522–2534, 2009.

30. Stokes, M. G., M. Kusunoki, N. Sigala, H. Nili, D. Gaffan, and J. Duncan, "Dynamic coding for cognitive control in prefrontal cortex," *Neuron*, Vol. 78, no. 2, pp. 364–375, 2013.

31. Blank, I. A., N. Kanwisher, and E. Fedorenko, "A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations.," *Journal of neurophysiology*, no. May, pp. 1105–1118, 2014.

32. Dosenbach, N. U. F., D. a. Fair, F. M. Miezin, A. L. Cohen, K. K. Wenger, R. a. T. Dosenbach, M. D. Fox, A. Z. Snyder, J. L. Vincent, M. E. Raichle, B. L. Schlaggar, and S. E. Petersen, "Distinct brain networks for adaptive and stable task control in humans.," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104, no. 26, pp. 11073–8, 2007.

33. Rypma, B., and M. D'Esposito, "The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences," *Proceedings of the National . . .*, Vol. 96, no. May, pp. 6558–6563, 1999.

34. Schaefer, A., T. S. Braver, J. R. Reynolds, G. C. Burgess, T. Yarkoni, and J. R. Gray, "Individual Differences in Amygdala Activity Predict Response Speed during Working Memory," *Journal of Neuroscience*, Vol. 26, no. 40, pp. 10120–10128, 2006.

35. Yarkoni, T., D. M. Barch, J. R. Gray, T. E. Conturo, and T. S. Braver, "BOLD correlates of trial-by-trial reaction time variability in gray and white matter: A multi-study fMRI analysis," *PLoS ONE*, Vol. 4, no. 1, 2009.

36. Taylor, J. S. H., K. Rastle, and M. H. Davis, "Interpreting response time effects in functional imaging studies," *NeuroImage*, Vol. 99, pp. 419–433, 2014.

37. Geake, J. G., and P. C. Hansen, "Neural correlates of intelligence as revealed by fMRI of fluid analogies," *NeuroImage*, Vol. 26, no. 2, pp. 555–564, 2005.

38. Shamosh, N. A., C. G. DeYoung, A. E. Green, D. L. Reis, M. R. Johnson, A. Conway, R. W. Engle, T. S. Braver, and J. R. Gray, "Individual differences in delay discounting," *Psychological Science*, Vol. 19, no. 9, pp. 904–911, 2008.

39. Todd, J. J., and R. Marois, "Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity.," *Cognitive, affective & behavioral neuroscience*, Vol. 5, no. 2, pp. 144–155, 2005.

40. Kriegeskorte, N., W. K. Simmons, P. S. F. Bellgowan, and C. I. Baker, "Circular analysis in systems neuroscience: the dangers of double dipping.," *Nature neuroscience*, Vol. 12, no. 5, pp. 535–540, 2009.

41. Richter, W., K. Ugurbil, A. Georgopoulos, and S. G. Kim, "Time-resolved fMRI of mental rotation," *Neuroreport*, Vol. 8, no. 17, pp. 3697–3702, 1997.

42. Bellgowan, P. S. F., Z. S. Saad, and P. a. Bandettini, "Understanding neural system dynamics through task modulation and measurement of functional MRI amplitude, latency, and width.," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 100, no. 3, pp. 1415–1419, 2003.

43. Mohamed, M. a., D. M. Yousem, A. Tekes, N. Browner, and V. D. Calhoun, "Correlation between the amplitude of cortical activation and reaction time: A functional MRI study," *American Journal of Roentgenology*, Vol. 183, no. 3, pp. 759–765, 2004.

44. Richter, W., R. Somorjai, R. Summers, M. Jarmasz, R. S. Menon, J. S. Gati, C. Tegeler, K. Ugurbil, and S.-g. Kim, "Motor Area Activity During Mental Rotation Studied by Time-Resolved Single-Trial fMRI," *Journal of cognitive neuroscience*, Vol. 12, no. 2, pp. 310–320, 2000.

45. Baayen, R. H., and P. Milin, "Analyzing Reaction Times," *International Journal of Psychological Research*, Vol. 3, no. 2, pp. 12 – 28, 2010.

46. Weissman, D. H., K. C. Roberts, K. M. Visscher, and M. G. Woldorff, "The neural bases of momentary lapses in attention.," *Nature neuroscience*, Vol. 9, no. 7, pp. 971–8, 2006.

47. Yarkoni, T., J. R. Gray, E. R. Chrastil, D. M. Barch, L. Green, and T. S. Braver, "Sustained neural activity associated with cognitive control during temporally extended decision making," *Cognitive Brain Research*, Vol. 23, no. 1, pp. 71–84, 2005.

48. Dale, A. M., and R. L. Buckner, "Selective Averaging of Rapidly Presented Individual Trials Using fMRI," *Hum. Brain. Mapp.*, Vol. 5, pp. 329–340, 1997.

49. Grinband, J., T. D. Wager, M. Lindquist, V. P. Ferrera, and J. Hirsch, "NIH Public Access," Vol. 43, no. 3, pp. 509–520, 2009.

50. Friston, K. J., a. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: A general linear approach," *Human Brain Mapping*, Vol. 2, no. 4, pp. 189–210, 1995.

51. Friston, K. J., *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, pp. 101–125. Academic Press Inc, 2007.

52. Friston, K. J., K. E. Stephan, T. E. Lund, A. Morcom, and S. Kiebel, "Mixed-effects and fMRI studies," *NeuroImage*, Vol. 24, no. 1, pp. 244–252, 2005.

53. Aron, A. R., M. a. Gluck, and R. a. Poldrack, "Classification Learning Task," *Neuroimage*, Vol. 29, no. 3, pp. 1000–1006, 2006.

54. Omura, K., A. Aron, and T. Canli, "Variance maps as a novel tool for localizing regions of interest in imaging studies of individual differences.," *Cognitive, affective & behavioral neuroscience*, Vol. 5, no. 2, pp. 252–261, 2005.