Connecting the Dots: Rethinking the Relationship between Code and Prose Writing with Functional Connectivity

Zachary Karas
University of Michigan
Ann Arbor, MI, USA
zackar@umich.edu

Andrew Jahn
University of Michigan
Ann Arbor, MI, USA
ajahn@umich.edu

Westley Weimer
University of Michigan
Ann Arbor, MI, USA
weimerw@umich.edu

Yu Huang
University of Michigan
Ann Arbor, MI, USA
yhhy@umich.edu

ABSTRACT

Medical imaging studies of software engineering have risen in popularity and may reveal the neural underpinnings of coding activities. To date, however, all studies in computer science venues have treated brain regions independently and in isolation. Since most complex neural activity involves coordination among multiple regions, previous analyses may overlook neural behavior.

We propose to apply *functional connectivity analysis* to medical imaging data from software engineering tasks. Informally, this analysis treats the brain as a graph, rather than a series of independent modules, and statistically infers relevant edges. We present a functional connectivity analysis of existing data, which elucidates the interconnections between code writing and prose writing, especially regarding higher mathematics and semantic processing. First, we found a significant link between Broca's Area (language) and the Number Form Area (higher mathematics) for coding. This both refines previous interpretations that code writing and natural language are distinct from each other, and may also contribute to the understanding of the Number Form Area in the Psychology literature. Second, we identify an area with important functional connectivity for both prose writing and coding, unlike previous analyses that associated it with coding. This advances our neural understanding of coding and prose writing, and was only exposed by using functional connectivity analysis. Third, for coding, we find a strong functional connectivity result for a brain region involved in semantic processing for language, with implications for CS training. Finally, we find a neural relationship between coding and expertise, including a more grounded explanation than prior work.

CCS CONCEPTS

• **Software and its engineering** → **Software creation and management**; • **Social and professional topics** → *User characteristics*.

KEYWORDS
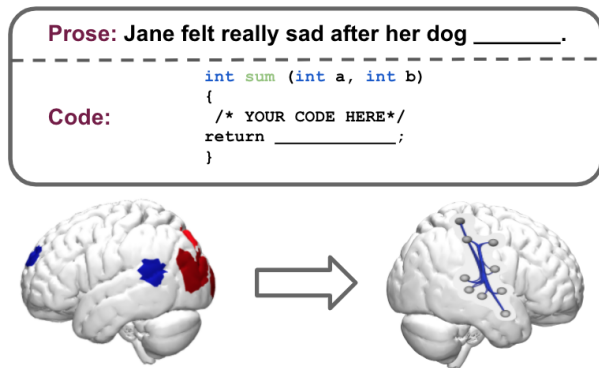
functional connectivity, fMRI, code writing, expertise

1 INTRODUCTION

Neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), provide researchers with a wealth of information related to brain activity, and are being applied more frequently to study the cognitive processes of programming [25, 35, 55, 63, 78, 86, 87]. These studies have examined code review [35], data structure manipulation [55], debugging [15], code reading [86], and code writing [63], reporting activity in regions associated with working memory, top-down control, and spatial reasoning, among others. Some studies have also compared coding with prose processing [63] or spatial reasoning [55]. These neurological findings have led to longitudinal studies exploring transfer training in CS undergraduate education [29]. Imaging studies unveil the neurological foundations of, and provide new approaches for, investigating and understanding programming activities, which paves the way for effectively improving CS pedagogy, technology transfer, and workforce retraining [35, 55, 63]. However, prior work in CS has only shown what specific brain regions are involved in programming. Brain areas do not function in isolation, but interact and are connected structurally and functionally [27, 52, 103].

Standard fMRI analyses are *modular*, meaning activity is modeled individually for each region in the brain, and these modules are assumed to be independent from one another [71]. Ignoring how disparate regions interact to accomplish complex coding tasks may result in overlooked connections and missed information. Suppose, informally, that debugging involves brain region *A* interacting with

Figure 1: High-level illustration. The proposed functional connectivity analysis reveals brain interactions mediating writing tasks, language, and expertise, which may be overlooked by standard modular analyses.

region *B*, while coding involves region *A* interacting with region *C*. Although how *A* conditionally interacts is critical, modular analyses will not see *A* as distinguishing or contrasting them (since it is active for both tasks); we desire an analysis that sees $A \rightarrow B$ and $A \rightarrow C$ edges. To gain a better understanding of programming on a neurological level, we present a *functional connectivity* analysis of an archival dataset on code writing. As shown conceptually in Figure 1, we investigate programming tasks using more than the standard localized analyses, present results that were not discovered before, and indicate the potential of functional connectivity analysis on cognitive processes of programming.

Instead of localizing brain activities to certain brain areas, functional connectivity analysis is based in graph theory, and provides information about interactions between brain regions by quantifying the statistical correlations of neurophysiological activity [26, 38]. Functional connectivity analysis can reveal (1) *which* regions of brain activations are significantly correlated and (2) *how much* they are correlated with each other in a statistical manner, which cannot be covered by standard modular analysis. Furthermore, there are no additional requirements on fMRI data collection for functional connectivity analysis, making it feasible and relatively easy for software engineering researchers to adapt it to archival data. Still new to computer science, functional connectivity has been widely used in cognitive and clinical psychology to understand complex neural processes. For instance, it has been effective in testing new medications for depression [2] and attention deficit hyperactivity disorder (ADHD) [62]. It has revealed new information about the neural substrate of schizophrenia [3] and Alzheimer's disease [22], which cannot be localized to any specific region in the brain. Furthermore, functional connectivity analysis has resolved conflicting evidence in cognitive psychology, such as for *apraxia of pantomime* (incapability of mimicking the use of a tool without holding it, e.g., brushing one's teeth without holding a toothbrush)) [90]. With functional connectivity analysis, researchers found this behavior was due to a breakdown in communication between multiple areas, rather than any particular localized region [90]. These studies in cognitive and clinical psychology provide models and motivations

for analyzing functional connectivity in software engineering tasks like code writing (which have potentially-conflicting evidence).

Since the pioneering work by Siegmund *et al.* on understanding cognitive processes in software engineering with fMRI in 2014 [86], there has been a growth of interest in the community [14, 24, 31, 35, 54–56, 73]. In recent years, fMRI has also been adapted to other computer science domains to help researchers understand and improve their design models [95]. Since fMRI is relatively higher-cost, researchers in software engineering have begun releasing their de-identified fMRI datasets [35, 54, 55]. We use this particular fMRI dataset because it is the most recent to compare code writing and prose writing, but the choice was otherwise arbitrary. Even though this dataset or others may not have been made with functional connectivity analysis in mind, there are no special requirements to run these analyses. It is therefore feasible and relatively easy for software engineering researchers to apply functional connectivity analyses to archival data.

To the best of our knowledge, only one previous paper in the psychology literature has considered connectivity and computing [15], and no such studies have been published in computing venues. We present a study that applies functional connectivity in software engineering. Our approach operates on an archival fMRI dataset and results in new findings that were overlooked before but are revealed in terms of of connectivity.

The contributions of this paper are as follows:

- The first study that uses functional connectivity analysis to investigate code writing. We present the value and feasibility of such analyses to software engineering researchers.
- The first study in software engineering that operates entirely on archival fMRI data. Such examples reduce the barrier to entry by demonstrating new insights without new, expensive experiments.
- Functional connectivity results that clarify and advance previous understandings. We find a significant link between the higher mathematics and language areas for coding. We identify an area with important functional connectivity for both prose writing and coding. We find a strong connectivity result between coding and language processing. Finally, we also find a relationship between coding and expertise.

2 BACKGROUND AND MOTIVATION

In this section, we introduce the basic idea of functional connectivity analysis and how it can reveal important findings in medical imaging. We also discuss the motivation of adapting functional connectivity analysis to software engineering. We finish by summarizing the neurological findings for code and prose writing from a previous fMRI study [63], which we re-analyze but with a different approach and new findings.

2.1 Functional Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a medical imaging technology that uses the magnetic properties of biological tissue to differentiate between structures [99]. *Functional* MRI (fMRI) applies this to oxygenated and de-oxygenated blood in the brain, which have slightly different magnetic properties, and can be used as a proxy to measure brain activation [23]. These scanners are incredibly precise,

and can measure brain activity in 1-millimeter cubes, called voxels ("volume" and "pixel"). When activity in a given region spikes, there is a short lag followed by a *blood-oxygen-level-dependent* (BOLD) response, which provides that region with oxygen and nutrients [17]. Researchers can measure these BOLD responses to make inferences about which brain regions are enlisted for a task. Scientists might design different experimental conditions, such as looking at houses or faces [93], and by *subtracting* the patterns of activation elicited by both types of stimuli, they can determine which patterns of activation are due to visual processing, and which are due to stimuli-specific processing. Such subtractions are called *contrasts*.

2.2 Functional Connectivity

Functional connectivity is defined as the temporal coincidence of spatially distant neurophysiological events [37]. In this type of analysis, two regions are considered to show functional connectivity when there is a statistical relationship between the time series of recorded brain activations [26]. To give an analogy, standard analysis looks at activity in the brain as if each region is a worker in a cubicle, while connectivity analysis models brain activity like a cooperative environment where hubs of specialized activity collaborate with other regions to work on complex tasks. Similar to standard analysis, functional connectivity analysis does not indicate causality or chemical mechanisms underlying brain activations. Instead, it applies to the voxel-level and provides information of time-course correlations between brain regions.

In modern psychology, researchers developed functional connectivity analysis to model the brain as a network system characterized by principles of segregation and integration [5, 6]. Segregation is similar to the idea of localization, where brain functions map onto specific regions, but segregation assumes regions and their processes are interactive [38]. The Human Connectome Project began in 2009 to map all the distinct regions of the brain, finding 180 areas per hemisphere that are separated by function, cortical architecture, or connectivity patterns [40]. There are temporal classifications as well, with some processes happening on the scale of milliseconds [61] and other processes, such as learning or memory, happening on the scale of weeks or years [42]. This modularity is integral for the brain's adaptability, where one component can change without affecting the rest of the system [5]. Informally phrased in terms of graph theory, these modules represent *nodes* [6].

Regions are integrated with one another, communicating information to accomplish complex tasks, and are connected with structural white matter [27, 70]. These links between regions can be considered *edges* between nodes in the brain [49]. Researchers have found that a region's intrinsic and extrinsic connections are entirely unique, going so far as to call each region's pattern of connections its 'fingerprint' [77]. It is important to note, however, that two regions do not have to be structurally connected to be functionally connected [53]. It is therefore difficult to understand a module's impact and role in a larger network without having measures of its structural and functional connections.

In terms of graph theory or distributed systems, to the best of our knowledge it is still unknown in the Psychology literature what the analogue of a *message* traveling between modules is, but it is clear that complex behaviors and properties emerge

through an interaction between regions or nodes. Viewed from a top-down perspective, these emergent behaviors and neural processes are non-reducible [75]. With vision for example, we recognize faces [39] and perceive motion [59], but even though vision is a cohesive experience, these processes rely on the coordination of different areas [100]. Participants have self-reported that prose writing feels similar to code writing [63], but the unreliability of self-reporting [55] warrants the use of neuroimaging and network-level measurements to unravel the nuanced processes of coding.

Functional connectivity analysis has successfully been applied to both resting-state and task-based experiments [36, 91]. In a resting-state experiment design, participants are not presented any stimuli and researchers can observe the spontaneous connections between regions [12]. These analyses have been used to identify differences between populations due to training [8], expertise [79], or even creativity [68]. Similarly, task-based analyses can also recover the same basic networks [30, 98]. In this paper, we apply functional connectivity analysis to task-related designs to explore differences in functional connectivity between different writing tasks.

2.3 Motivation: Functional Connectivity for SE

Functional connectivity analysis has been employed with increasing regularity in the cognitive science community, with over 130,000 functional connectivity studies published in the last ten years alone [74]. Connectivity analyses have had a greater clinical impact than fMRI on its own [36]. For example, these analyses have provided insight for conditions such as depression [2], attention deficit hyperactivity disorder (ADHD) [62], schizophrenia [3], and Alzheimer's disease [22], which cannot be localized to any specific region in the brain. For depression, the effectiveness of medications can be measured by calculating changes in correlation between regions before and after treatment [2]. Researchers have been able to uncover atypical connections in hub regions for bipolar disorder and schizophrenia, which are conditions affecting large networks in the brain and are not traceable to specific modules [3]. Inspired by research on functional connectivity in psychology, we believe similar benefits for software engineering may be available.

We hypothesize that functional connectivity can clarify prevailing ambiguities about the neurological basis of coding. Current neuroimaging studies report activity in language areas as well as spatial and mathematical reasoning areas while coding, but these findings are not consistent across tasks [63, 86]. Researchers have reported that code comprehension tasks elicit activity in language and attention areas [86, 87], suggesting code comprehension is akin to reading. In direct contrasts of code writing with prose writing however, researchers have found comparatively more activity in mathematical and spatial reasoning areas for programming [35, 55]. It is possible code comprehension is closer to prose reading than code writing is to prose writing, but we do not know for sure. Array and tree manipulation have elicited activity in regions associated with spatial and mathematical reasoning [55], so it is also possible only select subsets of coding activities rely on such processes.

Taken together, these findings suggest code reading and writing are similar yet distinct from natural language processes. This distinction was broadly characterized by Floyd *et al.* who found the

patterns of activation for code reading and prose reading can be distinguished with high accuracy [35], but this has not led to a closer understanding of the network-level interactions that enable someone to write code. Looking to cognitive psychology, researchers used functional connectivity in a similar situation where the findings were different depending on researchers' perspectives [90]. Castelhano *et al.* used a functional connectivity analysis to identify a network involved in debugging, so we aim to build off their findings and apply functional connectivity analysis to help reconcile contradictory evidence in the field.

Functional connectivity is a theoretical shift in neuroimaging analysis more than a technological advancement [6, 67], so further data collection is not always necessary. Functional connectivity analysis can be applied to existing neuroimaging data [83, 90]. In the present study, we report on findings from a connectivity analysis applied to fMRI data from Krueger *et al.* [63], which compares the neural processes of programming and prose writing. The previous publication by Krueger *et al.* found activity in left inferior frontal, bilateral parietal, and bilateral temporo-occipital regions. In using data from a previous experiment, we aim to demonstrate the efficacy of reexamining archival data as our analysis techniques improve.

3 STUDY DESIGN AND METHODS

We claim no novelty in the experimental design, participant recruitment or data acquisition associated with the archival dataset we analyze. On the contrary, we view the applicability of functional connectivity analysis to previously-collected archival data acquired without knowledge of this technique as an advantage of our proposed approach. This particular dataset comparing code writing and prose writing was chosen because it is the most recent on the topic, but the choice was otherwise arbitrary. We present a brief summary of the study, tasks, and participants in the archival dataset to provide context for interpreting our novel results, but clarify that our contribution is in an analysis of, and discovery of new results from, this existing data, and the proposal that others do the same, rather than in the initial collection of this data.

3.1 Archival Data

Study Overview. In this paper, we reuse the data from a study conducted by Krueger *et al.* [63] involving 30 (valid data from 29 of the participants) right-handed English speaking students (9 female, 1 fluid gendered) between the ages of 19–25 (M = 21.0, SD = 1.73) at a university in the United States. All participants were either undergraduate or graduate computer sciences students and conducted both code and prose writing tasks under an fMRI scan. Demographic data, and cumulative GPA were also collected from all participants. An MRI-safe keyboard was developed and used for the purposes of data acquisition for that experiment [63, Sec. 3.4]. In total, that previous study provides a dataset of approximately 700 Gigabytes of raw fMRI scan data.

Writing Tasks. In the Krueger *et al.* study, participants completed blocks of 17 fill-in-the-blank (FITB) questions and 9 free-response (FR) questions (as shown in Figure 2, retrieved from [63]). In between each question or trial was a rest period during which a fixation cross was presented on the screen for a random duration of 2–10 seconds. There was a code writing condition and a prose

Table 1: Demographics of the participants from the Krueger *et al.* [63] archival data analyzed in this paper.

Demographic Variables	# Participants	
Sex	Male	16
	Female	6
Gender	Men	16
	Women	5
	Fluid	1

writing condition for each type of block, FITB or FR, meaning there were four blocks total (FITB Code, FITB Prose, FR Code, FR Prose). Each block took about 20 minutes, and their order was pseudo-randomized. For both sets of FITB questions, participants had 30 seconds to respond. FITB code questions were created from established exercises in which part of the solution was replaced with a blank line, and FITB prose questions were created from adapted SAT questions specifically using non-math terms associated with quantitative reasoning [63]. Participants had 60 seconds to respond for all FR questions. FR prose questions were chosen from English as a Second Language (ESL) prompts.

Participants. The Krueger *et al.* study produced raw fMRI scan data from 22 participants after filtering for truncated or missing scan files. The 22 participants (5 female, 1 fluid gendered) were between the ages of 19 and 25 (M = 21.2; SD = 1.93). The demographics of the 22 participants is shown in Table 1. On average, participants had a GPA of 3.4/4 (min = 2.7, max = 4, SD = 0.29).

3.2 Functional Connectivity Analysis

Functional connectivity analysis finds statistical relationships between the time series of recorded brain activations [26], establishing the temporal coincidence of spatially distant neurophysiological events [37]. Very informally, for a computing audience, this analysis approach treats the brain as a graph in which the nodes correspond to regions of the brain (e.g., separated by brain architecture, using well-established bases such as the Human Connectome Project [40]) and the potential edges correspond to white matter links between regions [49]. Medical imaging captures information over time: for an individual carrying out a task, this corresponds to a time series of snapshots of the same graph nodes but with different nodes active at different points in time. The goal is to determine which of the potential edges are actually used for coordination during a particular task and thus which connections are associated with complex or emergent behavior. This is done via statistical inference: if the pattern of activation in one node over time matches the pattern of activation in another node over time, those nodes are said to be *functionally connected*. See Section 2.2 for background information on this technique.

In a typical fMRI study, the amplitude of the BOLD response for each condition is estimated and then contrasted to determine where in the brain activity is greater for one condition than another [35, Sec. IV]. This type of analysis has the benefit of averaging over multiple trials of the same condition to form a robust estimate of the activity for that condition. However, this comes at the cost of reducing the entire time-series to a single value per condition,
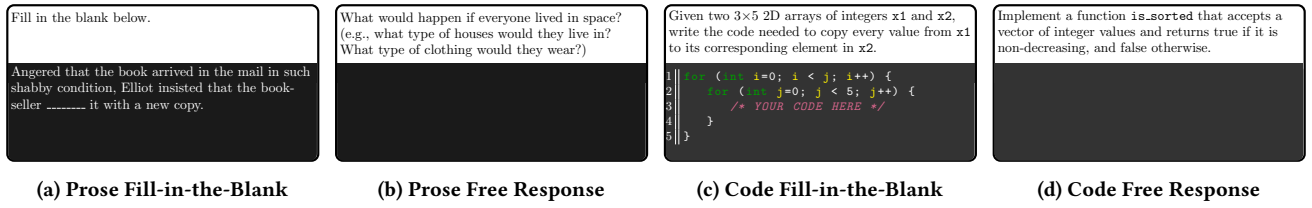
Figure 2: Illustrative examples of four categories of stimuli covering code and prose in fill-in-the-blank and free response scenarios. This figure is reproduced from Krueger *et al.* [63] to provide background information on the task types in the archival dataset used in this paper.

which essentially ignores most of the data collected during an fMRI scan and potentially loses valuable information.

Functional connectivity, on the other hand, uses the entire time-series to match the pattern of the BOLD response between different regions of the brain. Although the directionality of the functional connectivity cannot be determined — in other words, a significant correlation between regions does not provide information about whether one regionâĂŹs activity is causing the other regionâĂŹs activity [9] — it can nevertheless serve as an indicator of the organization of the brain. For example, functional connectivity shows reliable correlations between functionally similar areas such as the motor cortices [4], and functional connectivity between regions is strongly correlated with the underlying structural connectivity of the white matter tracts connecting those same regions [50].

This suggests that significant functional connectivity indicates not just co-activation, but the transfer of information between distant brain regions. Consequently, this method was best suited for our objective of determining whether and where there were differences in connectivity between conditions. This allows for more sophisticated inferences about the communication between brain regions under different cognitive conditions than is afforded by traditional univariate analyses.

Medical imaging data provides a high-resolution time series of snapshots of neural activity. Given that data, the key challenges are thus: (1) to abstract a large number of three-dimensional voxels into a small number of regions (nodes); and (2) to infer active edges based on statistical correlations between time series of node activations. We consider each challenge in turn. We used the CONN toolbox software for our numerical analysis, which is a widely used functional connectivity analysis package with well-maintained updates reflecting best practices [102]. The CONN toolbox is also compatible with other popular Matlab Packages used for fMRI data analysis (e.g., SPM).

3.2.1 Regions of the Brain. Multiple approaches have been proposed for partitioning the brain into regions or nodes, called *regions-of-interest* (ROIs) in the literature. Two common approaches include the Brodmann Areas (BAs) [69] and the the Harvard-Oxford Atlas [18], both of which represent well-established ways of dividing the brain based on behavior, anatomy and connectivity. In this paper we use the Harvard-Oxford Atlas, the default cortical atlas in the CONN toolbox.

Functional connectivity is typically computed one of three ways: between every pair of voxels (called a voxel-to-voxel analysis),

between a seed area and and every other voxel (called a seed-to-voxel analysis), or between pairs of regions (called an ROI-to-ROI analysis). Because the number of voxels is large (e.g., hundreds of thousands of voxel measurements may be taken in minutes [35, Sec. IV-A]) but the number of regions is typically much smaller (e.g., 48 cortical and 21 subcortical structural areas in the Harvard-Oxford Atlas), voxel-to-voxel analyses may complicate statistical analyses or otherwise suffer from scalability concerns [102, p. 129]. By contrast, the high level of abstraction in ROI-to-ROI analyses can lead to false positive concerns.

Since ROI-to-ROI and voxel-to-voxel approaches are more exploratory, we follow established practice in Psychology and use the seed-to-voxel analysis to uncover connectivity patterns given predetermined ROIs. We propose the use of seed-to-voxel analyses for the analysis of such software engineering data: they "show high reliability for well-characterized seed locations in resting state functional connectivity, both when using correlation- and regression-based measures to characterize functional connectivity" [102, p. 137]. This is appropriate in this context because previous work has identified and characterized seed regions relevant for various software engineering tasks. That is, the archival data that we propose to analyze, and other previously-published papers in software engineering (e.g., [63, 86]), already report regions of interest as a natural output of their non-connectivity analyses. Theoretically, functional connectivity analysis can be done independently (i.e., without any prior knowledge of pre-determined ROIs) with the ROIs selected by researchers. In practice, standard modular analysis (e.g., [63]) can first provide knowledge about ROIs to guide connectivity analysis while connectivity analysis can reveal truth that may be overlooked in modular analysis.

For our analysis, seed regions were chosen using results the published results of Krueger *et al.* [63]. In that paper, the authors reported localized brain activity based on Brodmann Areas [69], which were mapped to the Harvard-Oxford Atlas [18] automatically [102]. In particular, BA regions reported as having significant activity for the coding condition [63, Sec. 5.2–5.4] were taken as a potential ROIs. For this analysis, we chose a subset based on their overlap with the language network [33], and the peak regions of activity shown in Figures 5 and 7 of Krueger *et al.* [63]. Table 2 lists the seed regions used in this study as well as their associated BA regions.

3.2.2 Connectivity Inference. In this domain, *resting-state functional connectivity* is the correlation between two time-series from different regions of the brain. We employ a seed-to-voxel approach,

which first averages the time-series of all the voxels within a given region of interest. This averaged time-series is then correlated with the time-series of every other voxel in the brain to generate a whole-brain connectivity map.

The correlation coefficient r can be calculated by cross-correlating the static time-series of an ROI, represented by x, with the time-series of another voxel, represented by y:

$$r = (x^t \cdot x)^{1/2} \cdot b \cdot (y^t \cdot y)^{-1/2}$$

Where b represents the bivariate regression coefficient, calculated with the following formula:

$$b = (x^t \cdot x)^{-1} \cdot (x^t \cdot y)$$

Once r is calculated, it is then transformed to a z-value using Fisherâ˘A˘Zs inverse hyperbolic tangent function to make the values more closely approximate a normal distribution [102]. This is done to ensure that the assumption of normality is not violated for a second-level general linear model, as the untransformed correlation values tend to be skewed in the positive direction [102].

This z value is subsequently considered for significance during a two-step threshold to correct for false positives and multiple comparisons at the group level. In our analyses, an initial voxel-wise threshold of $p < 0.001$ creates clusters that are composed of individual voxels each passing a $p < 0.001$ threshold, and then a further cluster-wise alpha threshold of $p < 0.05$ only lets pass those clusters that, informally, we would see 5% of the time or less due to chance. The relatively conservative voxel-wise threshold we employ is to keep the nominal false positive rate below the 5% level, due to high autocorrelation in the fMRI signal [28].

3.2.3 Analysis Summary. The ultimate result of this analysis is a set of statistically-significant functional connections for a particular task (e.g., code writing). Informally, these are edges in the brain graph that denote temporal correspondences between spatially-distant neural regions. Because many complex cognitive activities involve coordination between multiple brain regions, functional connectivity has the potential to reveal insights and relationships not visible to previous, typical fMRI analyses.

4 RESULTS AND ANALYSIS

In this section, we investigate the following research questions, especially considering the relevant comparisons to the research questions explored in Krueger *et al.* [63]:

RQ1. How are brain regions involved in coding functionally connected to one another? Are there connections to brain regions involved in prose writing?

RQ2. How do fill-in-the-blank tasks compare to free response tasks in coding and prose writing?

RQ3. Do regions involved in the language network directly interact with regions involved in coding?

RQ4. What connections between brain regions are mediated by expertise in software engineering?

To answer the research questions, we analyze and discuss the results of functional connectivity analysis. All results are summarized in Table 2. For each seed region (column *Seed Regions*), we list the brain regions with which it has significant functional connectivity for three contrasted conditions: *Code > Prose, FR Code > FR Prose,*

and *FITB Code > FITB Prose.* Table 2 also lists the covariate analysis results of functional connectivity on expertise (measured by GPA). In this table, the *BA Regions* column lists the corresponding BA regions for every seed region to make it easier for readers to cross-reference with the original results [63], from which we obtained the archival dataset. Figures 3 and 4 accompany the table with visual plots of results from the connectivity analysis (voxel-wise $p < 0.001$, cluster corrected at $p < 0.05$).

4.1 RQ1 — High Level Code Writing vs. Prose Writing in Connectivity Analysis

Figure 3 summarizes our functional connectivity results for multiple seed regions; to explore general differences between the connectivity patterns of code writing and prose writing, we consider the *Code > Prose)* row. This row aggregates across both fill-in-the-blanks and free response tasks and corresponds to *(FITB Code + FR Code > FITB Prose + FR Prose)*. Values associated with the figure can be found in Table 2. For clarity, we separate the statistical results from the implications.

Statistical Results. Using the medial prefrontal cortex (MPFC, BA 6, 8) as a seed region revealed no significant differences in functional connectivity for code writing compared to prose writing. Using the left middle temporal gyrus (LMTG, BA 19, 37) as a seed region revealed significant functional connectivity for coding compared to prose in the left inferior temporal gyrus (LITG, BA 37; MNI: −54, −46, −8) with a cluster of 422 voxels and a peak t-statistic of 6.45 ($p < 0.001$). The right superior parietal lobule (RSPL, BA 5, 7) demonstrated significant functional connectivity with numerous regions, primarily for prose writing, including the right supramarginal gyrus (BA 40; MNI: 54, −20, 28; $k = 1333$; peak t-statistic = −8.24; $p < 0.001$), the central opercular cortex (BA 41; MNI: −40, 2, 12; $k = 178$; peak t-statistic = −4.9; $p = 0.03$), and the right insular cortex (MNI: 40, 2, −2; $k = 176$; peak t-statistic = −4.83; $p = 0.03$). The RSPL also demonstrated significant function connectivity with the left middle-frontal gyrus (LMFG, BA 46; MNI: −30, 10, 34) for coding ($k = 344$; peak t-statistic = 5.72; $p = 0.001$). Using the left inferior frontal gyrus (LIFG, BA 44) as a seed region revealed a significant functionally connected cluster in the right inferior temporal gyrus (RITG, BA 37; MNI: 50, −44, −8) for coding ($k = 428$; peak t-statistic = 5.89; $p < 0.001$). Using the right inferior frontal gyrus (RIFG, BA 44) as a seed region revealed significant functional connectivity in a cluster in the anterior cingulate cortex for writing (MNI: −6, −6, 30; $k = 210$; peak t-statistic = −6.90; $p = 0.015$). The *FR Code > FR Prose* contrast revealed significant functional connectivity with two seed regions: the LMTG with the left temporal pole (LTP, BA 38; MNI: −34, 18, −28; $k = 393$; peak t-statistic = −6.09; $p < 0.001$), and the RSPL with the right supramarginal gyrus (MNI: 56, −22, 30; $k = 226$; peak t-statistic = −5.55; $p = 0.01$). The *FITB Code > FITB Prose* contrast also revealed significant functional connectivity with two seed regions: the LMTG with the LITG (MNI: −60, −52, −12; $k = 324$; peak t-statistic = 6.6; $p < 0.01$) and the RSPL with the right perietal operculum cortex (RPOC, BA 41; MNI: 44, −40, 32; $k = 244$; peak t-statistic = −5.52; $p < 0.01$).

Table 2: Seed regions and regions revealed by the functional connectivity analysis (k = cluster size (voxels); t = peak t-statistic; p = voxel-wise p-value; "—" = no significant result).

Seed Regions	BA Regions	Code > Prose	FR Code > FR Prose	FITB Code > FITB Prose	GPA
MPFC	(6, 8)	—	—	—	—
LMTG	(19, 37)	• **LITG** (−54, −46, −8) $k = 422; t = 6.45; p < 0.001$	• **LTP** (−34, 18, −28) $k = 393; t = -6.09; p < 0.001$	• **LITG** (−60, −52, −12) $k = 324; t = 6.6; p < 0.01$	—
RSPL	(5, 7)	• **RSG** (54, −20, 28) $k = 1333; t = -8.24; p < 0.001$ • **LMFG** (−30, 10, 34) $k = 344; t = 5.72; p < 0.001$ • **COC** (−40, 2, 12) $k = 178; t = -4.9; p < 0.05$ • **RIC** (40, 2, −2) $k = 176; t = -4.83; p < 0.05$	• **RSG** (56, −22, 30) $k = 226; t = -5.55; p < 0.05$	• **RPOC** (44, −40, 32) $k = 244; t = -5.52; p < 0.01$	*FR Code > FITB Code:* • **LLOC** (−34, −86, 28) $k = 218; t = -5.88; p < 0.05$ • **RAG** (64, −40, 40) ($k = 188; t = 4.79; p < 0.05$) *FR Code > FR Prose:* • **LFOC** (−26, 46, −10) $k = 523; t = -5.34; p < 0.001$
LIFG	(44)	• **RITG** (50, −44, −8) $k = 428; t = 5.89; p < 0.001$	—	—	—
RIFG	(44)	• **ACC** (−6, −6, 30) $k = 210; t = -6.9; p < 0.05$	—	—	—

Brain region acronyms: MPFC = medial prefrontal cortex; LMTG = left middle temporal gyrus; RSPL = right superior parietal lobule; LIFG = left inferior frontal gyrus; RIFG = right inferior frontal gyrus; LITG = left inferior temporal gyrus; RSG = right supramarginal gyrus; LMFG = left middle frontal gyrus; COC = central opercular cortex; RIC = right insular cortex; RITG = right inferior temporal gyrus; ACC = anterior cingulate cortex; LTP = left temporal pole; RPOC = right parietal operculum cortex; LLOC = left lateral occipital cortex; RAG = right angular gyrus; LFOC = left frontal orbital cortex.
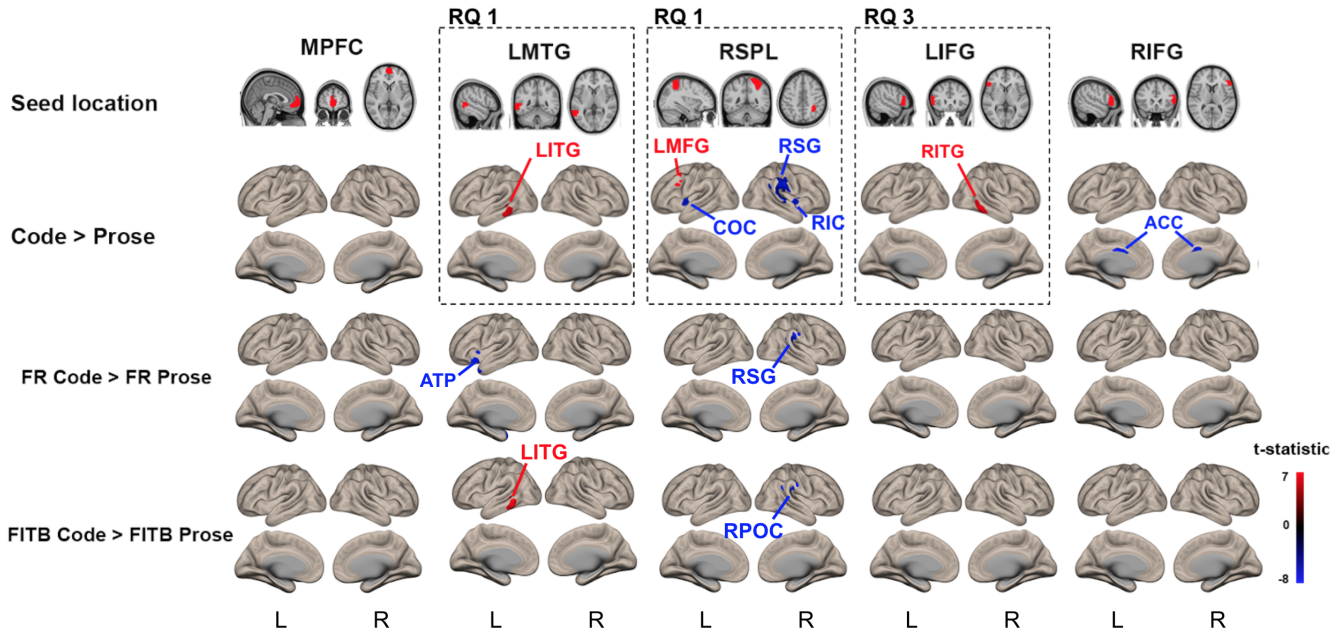
Figure 3: Function connectivity results using five seed regions (listed in the *Seed location* row): the medial prefrontal cortex (MPFC), left middle temporal gyrus (LMTG), right superior parietal lobule (RSPL), left inferior frontal gyrus (LIFG), and right inferior frontal gyrus (RIFG), ($p < 0.05$, uncorrected). Red corresponds to positive numbers in the contrast (left of the > symbol), and blue corresponds to negative numbers (right of the > symbol) in Table 2.

Implications. Krueger *et al.* reported higher brain activity for prose writing in the LMTG and higher brain activity for code writing in the LITG. Here we report significant functional connectivity between the LMTG and the LITG. These regions are adjacent to one another and implicated in semantic language processes [21].

Activity in both regions has been reported in previous code comprehension [86] and debugging studies [24], but we present the first evidence of significant functional connectivity between the two.

We also find the RSPL has more functional connectivity on the right hemisphere for prose writing compared to code writing, as

seen in Figure 3. Notably, this region displays significant functional connectivity with other regions for both code writing and prose writing, demonstrating that this area is involved in *both* processes. Previous findings reported higher activity for code writing in the RSPL [63], while we find stronger functional connectivity in this region for prose writing. We claim these findings are not necessarily contradictory since both analyses are measuring different features of brain functionality.

Previous findings report activity in the LIFG for both prose and code writing, but higher activity for prose writing in the *FR Code > FR Prose* contrast. This region was significantly functionally connected to the right inferior temporal gyrus (RITG). The LIFG is a component of the language network [33] and the RITG has been implicated in mathematical operations [51]. These results have larger implications for code writing and prose writing, which are discussed in more detail in Research Question 3. Activity was previously found for code writing in the anterior cingulate cortex [63], which is implicated in a broad array of higher cognitive functions [13]. We find this region demonstrates significant functional connectivity with the RIFG which has been implicated in inhibitory functions [47].

> We find an overlap of functional connectivity patterns for code writing and prose writing. There is a stronger connection for code writing among regions associated with semantic processing in natural language. Previous SE studies reported activity in both regions in different contexts, but we show a *functional link between the two*. The right superior parietal lobule, associated with visuospatial functions, shows significant connectivity with other regions for both code and prose writing.

4.2 RQ2 — Effects of Task and Language Types on Brain Function Connections

To explore the varying demands of the fill-in-the blank tasks and free response tasks, we analyzed *FR Code > FITB Code* and *FR Prose > FITB Prose* contrasts. These tasks did not yield any significant differences in functional connectivity, and are not shown in a table or figure. This might suggest there is a main effect of either coding or prose writing that outweighs any differences in functional connectivity between the two tasks. A slightly different interpretation might be there is little variation in the cognitive demands of both types of tasks. A task difference is uncovered for *FR Code > FITB Code* when participants' GPA is included as a covariate, which can be seen in the rightmost column of Table 2. These results related to expertise are discussed in more detail in Research Question 4.

> Significant functional connectivity differences among brain regions exist between language types (code vs. prose) but not between task types (fill in the blank vs. long response).

4.3 RQ3 — Language and Coding Connections

To investigate if the regions of the language network interact with regions involved in code writing, we consider the significant functional connectivity between the LIFG and the RITG in the *Code > Prose* contrast (MNI: 50, −44, −8; k = 428; peak t-statistic = 5.89; $p < 0.001$, uncorrected), which is shown in the *Code > Prose* row
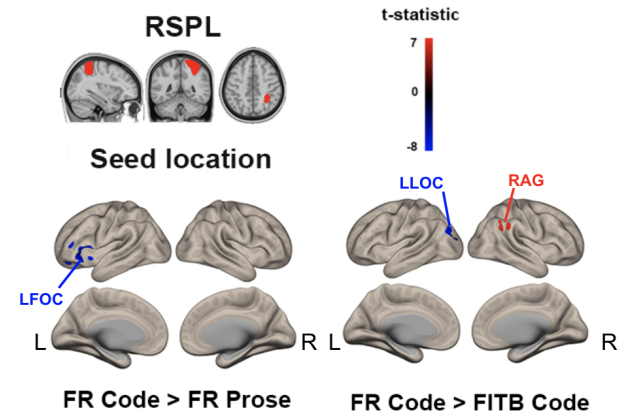
Figure 4: Functional connectivity results for expertise analysis using GPA. Only correlations that reached significance are plotted ($p < 0.05$). These were in the *FR Code > FR Prose* and *FR Code > FITB Code* contrasts with the right superior parietal lobule (RSPL) as a seed region. Red corresponds to *FR Code* in both images, and blue corresponds to *FR Prose* and *FITB Code*, from left to right.

and *LIFG* column of Figure 3. Table 2 shows the values associated with the figure.

The LIFG is historically one of the primary regions associated with language production, called Broca's Area, and has been implicated in grammar as well as syntactic and semantic processing [45]. The RITG is known as the Number Form Area (NFA), and has been associated with visual processing of arabic numbers [105], nonsymbolic mathematics [51], geometry [1], and equation comprehension [1]. Activity in the RITG was reported previously by Krueger *et al.* in contrasts comparing code writing and prose writing. While activity in this region was then interpreted as evidence that code writing and prose writing are distinct from one another, here we show a strong link between language and mathematical processing. This suggests the answer to the question whether coding is the same as or different from prose writing is more nuanced. Both processes overlap with language regions, but coding involves at least one region (the NFA) unique to the complex cognitive task of programming.

> Unlike previous work that found code and prose writing to be distinct, we find a significant functional connection ($p < 0.001$) between Broca's area, a primary region of the language network involved in grammar and semantics, and a coding region involved in higher level mathematical operations, the Number Form Area.

4.4 RQ4 — Expertise in Code Writing

We explored the effects of expertise on changes in functional connectivity for code writing compared to prose writing. Figure 4 and the rightmost column of Table 2 show the functional connectivity between regions that is significantly impacted by participants' GPA.

As in previous questions, we see that the RSPL shows the highest amount of connectivity out of the seed regions chosen. For the *FR Code > FR Prose* contrast, activity in the RSPL becomes significantly more correlated with activity in the left frontal orbital cortex (MNI: $-26, 46, -10$; $k = 523$; peak t-statistic $= -5.34$; $p < 0.001$) as GPA increases. For the *FR Code > FITB Code* contrast, as GPA increases we find activity in the RSPL becomes significantly more correlated with activity in the right angular gyrus (RAG, BA 40; MNI: $64, -40, 40$) for free response as opposed to fill in the blank ($k = 188$; peak t-statistic $= 479$; $p = 0.03$) and the left lateral occipital cortex (LLOC, BA 19, 39; MNI: $-34, -86, 28$) for prose ($k = 218$; peak t-statistic $= -5.88$; $p = 0.01$).

Previous results cited activity in the fusiform gyrus as potential evidence of expertise [63], but cognitive psychologists are effectively using functional connectivity analysis to determine how the interaction between regions is mediated by proficiency [79]. Here we measure expertise in computer science by considering CS students' cumulative GPA. We find the RSPL, which has been implicated in visuospatial processing [16], exhibits different patterns of functional connectivity for both code and prose writing as GPA increases. These regions include the RAG for coding and the LLOC for prose. Increased functional connectivity with the RAG is particularly interesting within the coding tasks, as research finds this region is involved in math competence [44], mental arithmetic [43], and memory specifically related to mathematics [43], such as maintaining a mental representation of a number line [41].

> We find functional connectivity from the right superior parietal lobule as a seed region is significantly modulated by GPA in the *FR Code > FITB Code* contrast with a region involved in *mathematical calculation and memory*. The RSPL again shows the highest amount of connectivity in our present findings.

5 LIMITATIONS

The first potential limitation in the current study is the intrinsic susceptibility to noise in functional connectivity analysis, especially to participants' motion in the scanner [97]. Considering functional connectivity analysis seeks to correlate patterns of brain activity between regions, participants moving their heads can affect entire brain scans at once, potentially leading to false positive results. To mitigate these effects, participants were preemptively instructed to remain still during the experiment. The data used in the current experiment are also processed and denoised specifically for these types of noise artifacts by preprocessing software [102]. Scan data are considered usable for functional connectivity if less than 10% of volumes are flagged as problematic due to scanner issues or motion. No participants in the current analysis met that threshold.

Recently, there have been concerns related to the reliability of functional connectivity results. Researchers have identified important factors for increasing reliability, such as scan length, task design, and number of subjects. The current experiment follows guidelines presented by Noble *et al.* [76], even though the tasks were originally designed without considering reliability measures for functional connectivity. For instance, the average scan length analyzed by Noble *et al.* was 9.7 minutes ($SD = 6.56$). Participants in the data analyzed here completed 4 scans, each of which was roughly 20 minutes long, well above the average considered for good reliability. Noble *et al.* also report greatest reliability for experimental designs in which participants were awake and engaged in a task. Participants in the data analyzed here completed timed fill-in-the-blank and free response questions with breaks of 2–10 seconds in between trials, so there was little chance for subjects to sleep. Group level results also show functional connectivity in areas associated with attention and working memory, empirically suggesting subjects were actively engaged. Another potential limitation of the current study is the sample size. The samples analyzed here are comparable to other fMRI studies considered by Noble *et al.* ($M = 25.59$; $SD = 12.92$, corrected for outliers).

We acknowledge the use of cumulative GPA as a proxy for expertise may not effectively measure programming proficiency in computer science undergraduate and graduate students. Coding expertise has proved difficult to gauge [35, 64]. GPA has been found to correlate with learning as well as aptitude [89]. Previous work in the field has also used GPA [35], self-reporting measures [64], and status as either an undergraduate or graduate student [101]. We also note that functional connectivity analyses can be applied to this archival dataset with relative ease, even though the dataset was collected with modular analyses in mind.

The authors of the dataset we analyze considered threats to validity as well [63], which we will summarize here. Krueger *et al.* were primarily concerned with task design and whether chosen prompts genuinely tapped into the neurological processes of code writing and prose writing. This was accounted for in the experimental design by including both fill-in-the-blank questions and free response questions. Those authors also cited concerns that reading the prompts themselves affected brain activity measurements. This possible confound was accounted for in experimental contrasts, where overlapping activity between the two tasks was cancelled out. Lastly, the dataset authors considered the homogeneity of the subjects' experience with computer science. Participants had, on average, taken 5.2 semesters worth of computer science coursework, but no significant findings were claimed at the individual level, only at the group level.

6 RELATED WORK

In this section, we present our findings in context with previous research.

Neuroimaging in Computer Science. There is a growing interest in understanding the neural basis of software engineering [15, 19, 24, 31, 35, 54, 56, 57, 73, 78, 86], most of which have been conducted in the last five years. Recent studies have primarily used fMRI [54, 55, 63, 78, 86], fNIRS [31, 54, 56, 73], and occasionally EEG [19] to explore the neurological patterns in programming activities. These studies investigated code comprehension [86], expertise [64], code review [54, 96], debugging [15, 24], data structure manipulation [55], and code writing [63].

We consider the work of Krueger *et al.* to be closest to our current work because we use the same fMRI data [63]. Their study examined code writing against prose writing to compare the underlying cognitive bases of the two processes. Two types of questions, fill-in-the-blank and free response, were used to further hone in on the nuanced distinctions between the two.

Functional Connectivity. In the 1990s, researchers first measured the functional interactions between brain regions [37]. Friston defined functional connectivity in 1993 as "temporal correlations between remote neurophysiological events" [37], finding that verbal fluency relied on different networks in the brain. Functional connectivity analysis today has been used with increasing frequency to study a wide range of cognitive processes. Functional connectivity has been used to examine the effects of acupuncture [81], meditation [94], aging [32], and even baseball [85]. These versatile analyses have had a great clinical impact [36], such as settling contradictory neuroimaging results to advance researchers' understanding [90, 104], and elucidating neurological changes related to expertise [68, 79]. These findings helped form the basis for the current study. We present findings that illustrate neurological differences related to expertise on code writing, which has been overlooked using localization-based analysis, and help classify the neural substrates of programming. We are aware of one previous paper that applies functional connectivity analysis to neuroimaging data in software engineering [15]. This study identified a top-down network involved in decision making and error detection associated with debugging.

Expertise in Programming. Several studies examining the cognitive processes of coding have looked at programming expertise and whether there are generalized differences between experts and novices. In their 2014 paper, Siegmund *et al.* found both positive and negative correlations between activation strength of certain brain regions and programming experience. Floyd *et al.*'s work found the neural representations of natural and programming languages to be distinct, but that those differences are modulated by expertise [35]. However, Krueger *et al.*'s study, the closest study to this presented work, reported no significant evidence indicating the effects of expertise on the contrast between code and prose writing. Our primary objective with applying functional connectivity analysis was to clarify previous findings, and our results show the interaction between two distinct regions is mediated by expertise (see Section 4.4). Siegmund *et al.* used BA 40 in their analysis of expertise, which overlaps with the angular gyrus reported in our results. Interestingly, Siegmund *et al.* find a weak positive correlation between Java knowledge and this region on code comprehension. Instead of investigating the effect of expertise on brain activation patterns and associated functions, other work has tried to judge differences between novices and experts by measuring cognitive effort with neuroimaging. For exmaple, Lee *et al.* used EEG and found experts showed more high-frequency brain waves compared to novices in the left prefrontal and left premotor regions in code comprehension [64].

Code, Prose, and Writing. There is a long-standing interest in improving programming productivity in software engineering. Researchers have designed software tools and interfaces to increase developers' productivity [88], and have investigated different programming languages for different user groups [82] to develop effective CS education strategies [84]. Before the first medical imaging study examining the neurological patterns of code writing by Krueger *et al.* in 2020, researchers conducted behavioral studies to understand the psychology of code writing, including the cognitive load [11] and impact of expertise [106] on code writing.

In comparison, prose writing has been investigated on both behavioral and neurological levels for a much longer span of time. For example, psychologists have proposed a theory of the cognitive processes of writing in early 1980s [34], and have studied the effects of expertise and second-language proficiency to the second-language writing performance [20], effective methods for developing expertise in English as a Second Language students [65], and a social apprenticeship model for gaining writing expertise [7]. Furthermore, with medical imaging and the standard modular analyses, researchers have located brain activity for prose writing in the left hemisphere, particularly the left superior parietal lobe [72]. Other medical imaging studies have been conducted to investigate the brain's writing center during both left- and right-handed writing tasks [92] and to identify brain regions that are consistently involved in prose writing tasks [80]. Moreover, functional connectivity analysis has also been used in prose writing. For instance, in handwriting, researchers used traditional modular analyses and found regions-of-interest that distinguished motor processes from cognitive processes, and later only by employing functional connectivity analysis clarified those links [104]. The results presented in this paper follow exactly that pattern (e.g., Krueger *et al.* [63] and Ivanova *et al.* [57] identified regions of interest via modular analyses that served as the basis for our connectivity analysis), and we anticipate more such results to come from connectivity analysis applied to software engineering.

7 DISCUSSION

In this section we consider our results from a broader perspective and propose directions for future research.

Perhaps most intriguing is the finding of a functional link between Broca's area in the left inferior frontal gyrus and the Number Form Area in the right inferior temporal gyrus. Broca's area was identified by Paul Broca in 1865 as an integral region of the language network [10], which has been consistently supported by modern medical imaging studies examining language [33]. The number form area has been studied extensively as well, finding this area is involved in equation comprehension [1], and nonsymbolic mathematics [51]. To the best of our knowledge, no psychology research studying this area has examined coding, or reported a connection between the NFA and coding (cf. [1, 46, 105]).

Finding a significant, functionally connected cluster between the NFA and Broca's area implies the neural basis of coding is intrinsically tied to language processes, but intuitively, is set apart from language by enlisting nonsymbolic mathematical operations. This has implications for computer science education and training, especially considering findings reported by Endres *et al.* that spatial-like processing was less correlated with improved performance for novices learning to program [29]. The results were unexpected for Endres *et al.*, but are supported by our current findings.

This link between code writing and prose writing is further supported by our other results, including the significant functional connectivity between two left temporal regions involved in semantic processing. This evidence again suggests coding relies on similar processes as natural language. Future research may consider particular hypotheses to untangle the basis for this finding, such as

whether or not participants maintain values assigned to variables in working memory.

Our finding that a single region on the right hemisphere demonstrates significant functional connections with areas for both code writing and prose writing exposes avenues for future research as well. This region had the most connections compared to the other seed regions, and on both hemispheres, which might suggest this region acts as a hub. Many of the destination areas have been implicated in sensation [48], attention [58], and working memory [66]. For our analysis of expertise, we find the RSPL becomes more functionally connected to the RAG, which is associated with mental calculation [43] and mathematical memory [41]. Disentangling these various connections might provide a more fine-grained understanding of the differences between code writing and prose writing.

At a high level, the new findings revealed in the functional connectivity analysis about the connections between code and prose writing (as discussed above) may also bring in more thoughts on new programming paradigms as described in Knuth's early work on literate programming [60]. We hope our work can provide a starting methodology for future investigations on functional connectivity in software engineering tasks.

Measures of localized brain activity have provided the first look at the neural substrate of programming, but characterizing these complex processes in terms of brain activations has led to ambiguity and conflicting evidence. Uncovering these new results by applying functional connectivity analysis to archival fMRI data indicates the viability of this approach, and its potential to advance our understanding of the cognitive processes of coding.

8 CONCLUSION

Standard analyses of fMRI data assume brain regions are independent from one another, and may overlook or mischaracterize neurological findings. Brain regions are inherently interactive, connected to one another structurally and functionally. We propose that computer scientists use functional connectivity analyses, which consider temporal connections between spatially-distinct brain regions. We applied functional connectivity analysis to archival data to uncover new results related to the neural basis of code writing and coding expertise. We find the left inferior frontal gyrus, known as Broca's area, is functionally connected to the left inferior temporal gyrus, known as the Number Form Area for coding compared to prose writing ($p < 0.001$). Researchers in psychology are still studying this region to understand its role in mathematical processes, and to the best of our knowledge, have not reported findings related to coding in this area. Showing a functional link between Broca's Area and the Number Form Area uncovers a strong link between language and mathematics in characterizing the neural basis of coding.

We find the right superior parietal lobule (RSPL) exhibits significant functional connectivity with numerous regions across tasks, and may act as a hub for code writing and prose writing. The functional connectivity of the RSPL was modulated by expertise as well, which was measured using participants' GPAs. As participants' GPAs increased, the functional connectivity between the RSPL and the RAG increased for free-response coding questions compared to

fill-in-the-blank coding questions ($p < 0.05$). Processes related to mental calculation and mathematical memory have been attributed to the RAG, suggesting expertise might modulate the correlation between visuospatial processing and mental calculation. Previous studies have used different metrics to locate neurological differences attributable to expertise. We build off these findings by showing a significant link between two regions that varies with proficiency.

While one study in the psychology literature has considered connectivity for computing, to the best of our knowledge this is the first to perform a large-scale, seed-to-voxel functional connectivity analysis on archival data and explain the technique for a computing audience. We demonstrate this analysis technique can uncover new results in archival data, clarify ambiguities, and advance our understanding of coding in the brain.

REFERENCES

[1] Marie Amalric and Stanislas Dehaene. 2016. Origins of the brain networks for advanced mathematics in expert mathematicians. *PNAS* 113, 18 (2016).

[2] A Anand, Y Li, Y Wang, J Wu, S Gao, L Bukhari, VP Mathews, A Kalnin, and MJ Lowe. 2005. Antidepressant effect on connectivity of the mood-regulating circuit: an FMRI study. *Neuropsychopharmacology* 30, 7 (2005).

[3] M Argyelan, T Ikuta, P DeRosse, RJ Braga, KE Burdick, M John, PB Kingsley, AK Malhotra, and PR Szeszko. 2014. Resting-state fMRI connectivity impairment in schizophrenia and bipolar disorder. *Schizophrenia bulletin* 40, 1 (2014).

[4] B Biswal B, F Z Yetkin, V M Haughton, and J S Hyde. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine* 34, 4 (1995), 537–541.

[5] Danielle S Bassett and Michael S Gazzaniga. 2011. Understanding complexity in the human brain. *Trends in cognitive sciences* 15, 5 (2011), 200–209.

[6] Danielle S Bassett and Olaf Sporns. 2017. Network neuroscience. *Nature neuroscience* 20, 3 (2017), 353–364.

[7] Anne Beaufort. 2000. Learning the trade: A social apprenticeship model for gaining writing expertise. *Written communication* 17, 2 (2000), 185–223.

[8] Alexander Belden, Tima Zeng, Emily Przysinda, Sheeba Arnold Anteraper, Susan Whitfield-Gabrieli, and Psyche Loui. 2020. Improvising at rest: Differentiating jazz and classical music training with resting state functional connectivity. *NeuroImage* 207 (2020), 116384.

[9] Bharat B Biswal. 2012. Resting state fMRI: a personal history. *Neuroimage* 62, 2 (2012), 938–944.

[10] Paul Broca. 1861. Remarques sur le siège de la faculté du langage articulé, suivies dâĂŹune observation dâĂŹaphémie (perte de la parole). *Bulletin et Memoires de la Societe anatomique de Paris* 6 (1861), 330–357.

[11] J. S. Bruner, J. J. Goodnow, and G. A. Austin. 1956. *A study of thinking*. New York: John Wiley & Sons, Inc.

[12] Randy L Buckner and Justin L Vincent. 2007. Unrest at rest: default activity and spontaneous network correlations. *Neuroimage* 37, 4 (2007), 1091–1096.

[13] George Bush, Phan Luu, and Michael I Posner. 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends in cognitive sciences* 4, 6 (2000).

[14] João Castelhano, Isabel C. Duarte, Carlos Ferreira, João Duraes, Henrique Madeira, and Miguel Castelo-Branco. 2018. The Role of the Insula in Intuitive Expert Bug Detection in Computer Code: An fMRI Study. *Brain Imaging and Behavior* (May 2018).

[15] Joao Castelhano, Isabel C Duarte, Carlos Ferreira, Joao Duraes, Henrique Madeira, and Miguel Castelo-Branco. 2019. The role of the insula in intuitive expert bug detection in computer code: an fMRI study. *Brain imaging and behavior* 13, 3 (2019), 623–637.

[16] Maurizio Corbetta, Gordon L Shulman, Francis M Miezin, and Steven E Petersen. 1995. Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science* 270, 5237 (1995), 802–805.

[17] David D Cox and Robert L Savoy. 2003. Functional magnetic resonance imaging (fMRI)âĂIJbrain readingâĂİ: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19, 2 (2003), 261–270.

[18] R Cameron Craddock, G Andrew James, Paul E Holtzheimer III, Xiaoping P Hu, and Helen S Mayberg. 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human brain mapping* 33, 8 (2012), 1914–1928.

[19] Igor Crk, Timothy Kluthe, and Andreas Stefik. 2015. Understanding programming expertise: an empirical study of phasic brain wave changes. *ACM (TOCHI)* 23, 1 (2015).

[20] Alister Cumming. 1989. Writing expertise and second-language proficiency. *Language learning* 39, 1 (1989), 81–135.

[21] J Davey, HE Thompson, G Hallam, T Karapanagiotidis, C Murphy, I De Caso, K Krieger-Redwood, BC Bernhardt, J Smallwood, and E Jefferies. 2016. Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *Neuroimage* 137 (2016).

[22] Emily L Dennis and Paul M Thompson. 2014. Functional brain connectivity using fMRI in aging and Alzheimerâ£Žs disease. *Neuropsychology review* 24, 1 (2014), 49–62.

[23] Edgar A DeYoe, Peter Bandettini, Jay Neitz, David Miller, and Paula Winans. 1994. Functional magnetic resonance imaging (FMRI) of the human brain. *Journal of neuroscience methods* 54, 2 (1994), 171–187.

[24] J. Duraes, H. Madeira, J. Castelhano, C. Duarte, and M. C. Branco. 2016. WAP: Understanding the Brain at Software Debugging. In *International Symposium on Software Reliability Engineering.* 87–92.

[25] Sjoerd JH Ebisch, Dante Mantini, Roberta Romanelli, Marco Tommasi, Mauro G Perrucci, Gian Luca Romani, Roberto Colom, and Aristide Saggino. 2013. Long-range functional interactions of anterior insula and medial frontal cortex are differently modulated by visuospatial and inductive reasoning tasks. *Neuroimage* 78 (2013), 426–438.

[26] S.B. Eickhoff and V.I. MÃijller. 2015. Functional Connectivity. In *Brain Mapping*, Arthur W. Toga (Ed.). Academic Press, Waltham, 187 – 201.

[27] Simon B Eickhoff and Christian Grefkes. 2011. Approaches for the integrated analysis of structure, function and connectivity of the human brain. *Clinical EEG and neuroscience* 42, 2 (2011), 107–121.

[28] Anders Eklund, Thomas E. Nichols, and Hans Knutsson. 2016. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *PNAS* 113, 28 (2016), 7900–7905.

[29] Madeline Endres, Zachary Karas, Xiaosu Hu, Ioulia Kovelman, and Westley Weimer. 2021. Relating Reading, Visualization, and Coding for New Programmers: A Neuroimaging Study. In *ICSE.*

[30] Damien A Fair, Bradley L Schlaggar, Alexander L Cohen, Francis M Miezin, Nico UF Dosenbach, Kristin K Wenger, Michael D Fox, Abraham Z Snyder, Marcus E Raichle, and Steven E Petersen. 2007. A method for using blocked and event-related fMRI data to study âĂIJresting stateâĂİ functional connectivity. *Neuroimage* 35, 1 (2007), 396–405.

[31] Sarah Fakhoury, Yuzhan Ma, Venera Arnaoudova, and Olusola Adesope. 2018. The Effect of Poor Source Code Lexicon and Readability on Developers' Cognitive Load. In *International Conference on Program Comprehension.*

[32] Luiz Kobuti Ferreira and Geraldo F Busatto. 2013. Resting-state functional connectivity in normal brain aging. *Neuroscience & Biobehavioral Reviews* 37, 3 (2013), 384–400.

[33] Evelyn C Ferstl, Jane Neumann, Carsten Bogler, and D Yves Von Cramon. 2008. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping* 29, 5 (2008), 581–593.

[34] Linda Flower and John R Hayes. 1981. A cognitive process theory of writing. *College composition and communication* 32, 4 (1981), 365–387.

[35] Benjamin Floyd, Tyler Santander, and Westley Weimer. 2017. Decoding the representation of code in the brain: An fMRI study of code review and expertise. In *International Conference on Software Engineering (ICSE).* 175–186.

[36] Michael D Fox and Michael Greicius. 2010. Clinical applications of resting state functional connectivity. *Frontiers in systems neuroscience* 4 (2010), 19.

[37] KJ Friston, CD Frith, PF Liddle, and RSJ Frackowiak. 1993. Functional connectivity: the principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow & Metabolism* 13, 1 (1993), 5–14.

[38] Karl J Friston. 2011. Functional and effective connectivity: a review. *Brain connectivity* 1, 1 (2011), 13–36.

[39] Isabel Gauthier, Michael J Tarr, Jill Moylan, Pawel Skudlarski, John C Gore, and Adam W Anderson. 2000. The fusiform âĂIJface areaâĂİ is part of a network that processes faces at the individual level. *Journal of cognitive neuroscience* 12, 3 (2000), 495–504.

[40] Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. 2016. A multi-modal parcellation of human cerebral cortex. *Nature* 536, 7615 (2016), 171–178.

[41] Silke Gobel, Vincent Walsh, and Matthew FS Rushworth. 2001. The mental number line and the human angular gyrus. *Neuroimage* 14, 6 (2001), 1278–1289.

[42] Philip Goelet, Vincent F Castellucci, Samuel Schacher, and Eric R Kandel. 1986. The long and the short of long–term memoryâĂŤa molecular framework. *Nature* 322, 6078 (1986), 419–422.

[43] Roland H Grabner, Daniel Ansari, Karl Koschutnig, Gernot Reishofer, and Franz Ebner. 2013. The function of the left angular gyrus in mental arithmetic: evidence from the associative confusion effect. *Human brain mapping* 34, 5 (2013).

[44] Roland H Grabner, Anja Ischebeck, Gernot Reishofer, Karl Koschutnig, Margarete Delazer, Franz Ebner, and Christa Neuper. 2009. Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. *Human brain mapping* 30, 9 (2009), 2936–2952.

[45] Yosef Grodzinsky. 2000. The neurology of syntax: Language use without Broca's area. *Behavioral and brain sciences* 23, 1 (2000), 1–21.

[46] Mareike Grotheer, Brianna Jeska, and Kalanit Grill-Spector. 2018. A preference for mathematical processing outweighs the selectivity for Arabic numbers in the inferior temporal gyrus. *Neuroimage* 175 (2018), 188–200.

[47] Adam Hampshire, Samuel R Chamberlain, Martin M Monti, John Duncan, and Adrian M Owen. 2010. The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage* 50, 3 (2010), 1313–1319.

[48] Gesa Hartwigsen, Annette Baumgaertner, Cathy J Price, Maria Koehnke, Stephan Ulmer, and Hartwig R Siebner. 2010. Phonological decisions require both the left and right supramarginal gyri. *PNAS* 107, 38 (2010), 16494–16499.

[49] Yong He and Alan Evans. 2010. Graph theoretical modeling of brain connectivity. *Current opinion in neurology* 23, 4 (2010), 341–350.

[50] A M Hermundstad, D S Bassett, K S Brown, E M Aminoff, D Clewett, S Freeman, A Frithsen, A Johnson, C M Tipper, M B Miller, S T Grafton, and J M Carlson. 2013. Structural foundations of resting-state and task-based functional connectivity in the human brain. *PNAS* 110, 15 (2013), 6169–6174.

[51] Ian D Holloway, Gavin R Price, and Daniel Ansari. 2010. Common and segregated neural pathways for the processing of symbolic and nonsymbolic numerical magnitude: An fMRI study. *Neuroimage* 49, 1 (2010), 1006–1017.

[52] Christopher J Honey, Rolf Kötter, Michael Breakspear, and Olaf Sporns. 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *PNAS* 104, 24 (2007), 10240–10245.

[53] Christopher J Honey, Olaf Sporns, Leila Cammoun, Xavier Gigandet, Jean-Philippe Thiran, Reto Meuli, and Patric Hagmann. 2009. Predicting human resting-state functional connectivity from structural connectivity. *PNAS* 106, 6 (2009), 2035–2040.

[54] Yu Huang, Kevin Leach, Zohreh Sharafi, Nicholas McKay, Tyler Santander, and Westley Weimer. 2020. Biases and differences in code review using medical imaging and eye-tracking: genders, humans, and machines. In *Foundations of Software Engineering.* 456–468.

[55] Yu Huang, Xinyu Liu, Ryan Krueger, Tyler Santander, Xiaosu Hu, Kevin Leach, and Westley Weimer. 2019. Distilling neural representations of data structure manipulation using fMRI and fNIRS. In *(ICSE).*

[56] Yoshiharu Ikutani and Hidetake Uwano. 2014. Brain activity measurement during program comprehension with NIRS. In *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing.* IEEE, 1–6.

[57] Anna A Ivanova, Shashank Srikant, Yotaro Sueoka, Hope H Kean, Riva Dhamala, Una-May O'reilly, Marina U Bers, and Evelina Fedorenko. 2020. Comprehension of computer code relies primarily on domain-general executive brain regions. *Elife* 9 (2020), e58906.

[58] Shruti Japee, Kelsey Holiday, Maureen D Satyshur, Ikuko Mukai, and Leslie G Ungerleider. 2015. A role of right middle frontal gyrus in reorienting of attention: a case study. *Frontiers in systems neuroscience* 9 (2015), 23.

[59] Gunnar Johansson. 1975. Visual motion perception. *Scientific American* 232, 6 (1975), 76–89.

[60] Donald Ervin Knuth. 1984. Literate programming. *Comput. J.* 27, 2 (1984), 97–111.

[61] Thomas Koenig, Leslie Prichep, Dietrich Lehmann, Pedro Valdes Sosa, Elisabeth Braeker, Horst Kleinlogel, Robert Isenhart, and E Roy John. 2002. Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. *Neuroimage* 16, 1 (2002), 41–48.

[62] Kerstin Konrad and Simon B Eickhoff. 2010. Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. *Human brain mapping* 31, 6 (2010), 904–916.

[63] Ryan Krueger, Yu Huang, Xinyu Liu, Tyler Santander, Westley Weimer, and Kevin Leach. 2020. Neurological Divide: An FMRI Study of Prose and Code Writing. In *ICSE (Seoul, South Korea) (ICSE '20).* Association for Computing Machinery, New York, NY, USA, 678âĂŞ690.

[64] S. Lee, A. Matteson, D. Hooshyar, S. Kim, J. Jung, G. Nam, and H. Lim. 2016. Comparing Programming Language Comprehension between Novice and Expert Programmers Using EEG Analysis. In *International Conference on Bioinformatics and Bioengineering.* 350–355.

[65] Ilona Leki. 1992. Building Expertise through Sequenced Writing Assignments. *Teachers of English to Speakers of Other Languages Journal* 1, 2 (1992), 19–23.

[66] H-C Leung, John C Gore, and Patricia S Goldman-Rakic. 2002. Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *Journal of cognitive neuroscience* 14, 4 (2002), 659–671.

[67] Kaiming Li, Lei Guo, Jingxin Nie, Gang Li, and Tianming Liu. 2009. Review of methods for functional brain connectivity detection using fMRI. *Computerized medical imaging and graphics* 33, 2 (2009), 131–139.

[68] Martin Lotze, Katharina Erhard, Nicola Neumann, Simon B Eickhoff, and Robert Langner. 2014. Neural correlates of verbal creativity: differences in resting-state functional connectivity associated with expertise in creative writing. *Frontiers in human neuroscience* 8 (2014), 516.

[69] Joseph A Maldjian, Paul J Laurienti, Robert A Kraft, and Jonathan H Burdette. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 3 (2003), 1233–1239.

[70] AR McIntosh and F Gonzalez-Lima. 1991. Structural modeling of functional neural pathways mapped with 2-deoxyglucose: effects of acoustic startle habituation on the auditory system. *Brain research* 547, 2 (1991), 295–302.

[71] Martin J McKeown and Terrence J Sejnowski. 1998. Independent component analysis of fMRI data: examining the assumptions. *Human brain mapping* 6, 5-6 (1998), 368–372.

[72] V Menon and JE Desmond. 2001. Left superior parietal cortex involvement in writing: integrating fMRI with lesion evidence. *Cognitive brain research* 12, 2 (2001), 337–340.

[73] Takao Nakagawa, Yasutaka Kamei, Hidetake Uwano, Akito Monden, Kenichi Matsumoto, and Daniel M German. 2014. Quantifying programmers' mental workload during program comprehension based on cerebral blood flow measurement: A controlled experiment. In *ICSE*.

[74] National Library of Medicine. [n.d.]. PubMed Search. Retrieved February, 2021 from https://pubmed.ncbi.nlm.nih.gov/?term=functional+connectivity.

[75] Denis Noble. 2008. *The music of life: biology beyond genes.* Oxford University Press.

[76] Stephanie Noble, Dustin Scheinost, and R Todd Constable. 2019. A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis. *Neuroimage* 203 (2019), 116157.

[77] Richard E Passingham, Klaas E Stephan, and Rolf Kötter. 2002. The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience* 3, 8 (2002), 606–616.

[78] Norman Peitek, Janet Siegmund, Sven Apel, Christian Kästner, Chris Parnin, Anja Bethmann, Thomas Leich, Gunter Saake, and André Brechmann. 2018. A look into programmersâĂŹ heads. *IEEE Transactions on Software Engineering* 46, 4 (2018), 442–462.

[79] Ana Luísa Pinho, Örjan de Manzano, Peter Fransson, Helene Eriksson, and Fredrik Ullén. 2014. Connecting to create: expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *Journal of Neuroscience* 34, 18 (2014), 6156–6163.

[80] Samuel Planton, Marieke Longcamp, Patrice Péran, Jean-François Demonet, and Mélanie Jucla. 2017. How specialized are writing-specific brain regions? An fMRI study of writing, drawing and oral spelling. *Cortex* 88 (2017), 66–80.

[81] Wei Qin, Jie Tian, Lijun Bai, Xiaohong Pan, Lin Yang, Peng Chen, Jianping Dai, Lin Ai, Baixiao Zhao, Qiyong Gong, et al. 2008. FMRI connectivity analysis of acupuncture effects on an amygdala-associated brain network. *Molecular Pain* 4 (2008), 1744–8069.

[82] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, and Yasmin Kafai. 2009. Scratch: Programming for All. *Commun. ACM* 52, 11 (Nov. 2009), 60–67.

[83] Jennifer L Robinson, Angela R Laird, David C Glahn, William R Lovallo, and Peter T Fox. 2010. Metaanalytic connectivity modeling: delineating the functional connectivity of the human amygdala. *Human brain mapping* 31, 2 (2010), 173–184.

[84] Jean Salac, Cathy Thomas, Chloe Butler, Ashley Sanchez, and Diana Franklin. 2020. TIPP&SEE: A Learning Strategy to Guide Students through Use-Modify Scratch Activities. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education.* 79–85.

[85] Jia-Hong Sie, Yin-Hua Chen, Chih-Yen Chang, Nai-Shing Yen, Woei-Chyn Chu, and Yuo-Hsien Shiau. 2019. Altered central autonomic network in baseball players: a resting-state fMRI study. *Scientific reports* 9, 1 (2019), 1–10.

[86] Janet Siegmund, Christian Kästner, Sven Apel, Chris Parnin, Anja Bethmann, Thomas Leich, Gunter Saake, and André Brechmann. 2014. Understanding

[87] source code with functional magnetic resonance imaging. In *International Conference on Software Engineering.* 378–389.

[87] Janet Siegmund, Norman Peitek, Chris Parnin, Sven Apel, Johannes Hofmeister, Christian Kästner, Andrew Begel, Anja Bethmann, and André Brechmann. 2017. Measuring neural efficiency of program comprehension. In *Foundations of Software Engineering.*

[88] David Canfield Smith. 1975. *Pygmalion: a creative programming environment.* Technical Report. STANFORD UNIV CA DEPT OF COMPUTER SCIENCE.

[89] Andrea Solimeno, Minou Ella Mebane, Manuela Tomai, and Donata Francescato. 2008. The influence of students and teachers characteristics on the efficacy of face-to-face and computer supported collaborative learning. *Computers & Education* 51, 1 (2008), 109–128.

[90] Christoph Sperber, Daniel Wiesen, Georg Goldenberg, and Hans-Otto Karnath. 2019. A network underlying human higher-order motor control: Insights from machine learning-based lesion-behaviour mapping in apraxia of pantomime. *Cortex* 121 (2019), 308–321.

[91] Michael C Stevens. 2016. The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation. *Neuroscience & Biobehavioral Reviews* 70 (2016), 13–32.

[92] Genichi Sugihara, Tatsuro Kaminaga, and Morihiro Sugishita. 2006. Interindividual uniformity and variety of the "Writing center": a functional MRI study. *Neuroimage* 32, 4 (2006), 1837–1849.

[93] James W Tanaka and Martha J Farah. 1993. Parts and wholes in face recognition. *The Quarterly journal of experimental psychology* 46, 2 (1993), 225–245.

[94] AA Taren, PJ Gianaros, CM Greco, EK Lindsay, A Fairgrieve, KW Brown, RK Rosen, JL Ferris, E Julson, AL Marsland, et al. 2017. Mindfulness meditation training and executive control network resting state functional connectivity: a randomized controlled trial. *Psychosomatic medicine* 79, 6 (2017), 674.

[95] Mariya Toneva and Leila Wehbe. 2019. Interpreting and improving natural-language processing (in machines) with natural language-processing (in the brain). *arXiv preprint arXiv:1905.11833* (2019).

[96] Hidetake Uwano, Masahide Nakamura, Akito Monden, and Ken-ichi Matsumoto. 2006. Analyzing Individual Performance of Source Code Review Using ReviewersâĂŹ Eye Movement. In *Eye Tracking Research Applications.*

[97] Koene RA Van Dijk, Mert R Sabuncu, and Randy L Buckner. 2012. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 1 (2012), 431–438.

[98] Eleanna Varangis, Christian G Habeck, and Yaakov Stern. 2021. Task-based functional connectivity in aging: How task and connectivity methodology affect discovery of age effects. *Brain and behavior* 11, 1 (2021), e01954.

[99] Marinus T Vlaardingerbroek and Jacques A Boer. 2013. *Magnetic resonance imaging: theory and practice.* Springer Science & Business Media.

[100] Brian Wandell and Stephen Thomas. 1997. Foundations of vision. *Psyccritiques* 42, 7 (1997).

[101] Mark Weiser and Joan Shertz. 1983. Programming problem representation in novice and expert programmers. *International Journal of Man-Machine Studies* 19, 4 (1983), 391–398.

[102] Susan Whitfield-Gabrieli and Alfonso Nieto-Castanon. 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity* 2, 3 (2012), 125–141.

[103] Jonathan Wirsich, Ben Ridley, Pierre Besson, Viktor Jirsa, Christian Bénar, Jean-Philippe Ranjeva, and Maxime Guye. 2017. Complementary contributions of concurrent EEG and fMRI connectivity for predicting structural connectivity. *NeuroImage* 161 (2017), 251–260.

[104] Yang Yang, Zhentao Zuo, Fred Tam, Simon J Graham, Ran Tao, Nizhuan Wang, and Hong-Yan Bi. 2019. Brain activation and functional connectivity during Chinese writing: An fMRI study. *Journal of Neurolinguistics* 51 (2019), 199–211.

[105] Darren J Yeo, Eric D Wilkey, and Gavin R Price. 2017. The search for the number form area: A functional neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews* 78 (2017), 145–160.

[106] Edward A Youngs. 1974. Human errors in programming. *International Journal of Man-Machine Studies* 6, 3 (1974), 361–376.