Studying Programmers Without Programming: Investigating Expertise Using Resting State fMRI

Zachary Karas*, Benjamin Gold*, Violet Zhou†, Noah Reardon†, Thad Polk†, Catie Chang*, Yu Huang*
*Vanderbilt University
{z.karas, benjamin.p.gold, catie.chang, yu.huang}@vanderbilt.edu
†University of Michigan
{violetz, nreardon, tpolk}@umich.edu

***Abstract*—Expert programmers are more effective at coding activities, but the reasons for this remain elusive. Accordingly, recent research has used neuroimaging such as fMRI to analyze how expert programmers might *think* as they perform coding activities. Those experiments have all involved specific programming tasks (i.e., comprehension), but have been unable to detect systematic differences based on coding experience. By using tasks, however, those studies may limit the number and type of brain networks involved. In Cognitive Neuroscience, researchers commonly analyze resting-state data, in which participants' brain activity is recorded as they lay idle in the scanner. The brain's functional organization is plastic, and can change with experience. These changes can be measured at rest, making this a suitable data type for studying how programming activities affect neural organization over time. In this paper, we analyzed the resting state scans from 150 participants, 96 of whom were programmers. We found increased connectivity in programmers between brain regions involved in language, math, and the temporal attention. Non-programmers demonstrated more connectivity with regions involved in social and emotional cognition. We found that as years of programming experience increases, connectivity decreases between two regions associated with visual processing during reading and articulation, respectively.**

***Index Terms*—Programming Expertise, fMRI, Resting State**

I. INTRODUCTION

Coding has increasingly become a necessary skill in the workforce [1]. To meet this economic and societal need, researchers have investigated how to best train and educate new programmers [2]. One approach for this is to identify the traits of expert programmers that make them experts, and help novices attain these [2]. Research into the cognitive factors of programming has framed this into 'how can we help novices *think* like experts?' [3, 4]. There is a vibrant history of behavioral research in Software Engineering (SE) that has investigated the cognitive processes of programming activities [3, 4, 5, 6], with some current research employing techniques like neuroimaging [7, 8]. Neuroimaging provides an exclusive vantage point into the cognition of programmers, and has helped ground the perceptual experiences of programmers in measurable findings (i.e., the influence of code complexity on brain activity [9]). Despite the progress, this technology has struggled to identify systematic differences based on coding experience [8]. This is partially due to experimental designs of previous neuroimaging studies, which have scanned participants while they complete single tasks (i.e., debugging [10]). Those studies have helped to localize coding processes within the brain, but have only been able to expose expertise differences specific to those processes. Thus, the window through which researchers have studied the cognition of expertise is relatively narrow. As such, we still lack a foundational understanding of cognitive differences resulting from time spent on coding activities. A broader perspective on these differences could lead to improved practices for training new programmers, deeper insight into the cognition of coding, and further guidance for future research.

In Cognitive Neuroscience, researchers regularly analyze *resting state scans*, where participants are scanned in an fMRI machine as they lay idle [11]. This flexible experimental design can expose pluralistic differences between study populations because participants are not given a targeted task, and are not presented with external stimuli [11]. Using resting state scans, researchers study patterns in the innate, spontaneous activity of participants' brains [11]. This type of data offers many benefits in that it does not require the design of an experimental task, it can be more straightforward for comparisons without a task, and there exist many publicly available resting state datasets [12]. Resting state data has proved valuable because after extensive research on the brain "at rest," research has uncovered a consistent functional organization where certain brain regions are more likely to exchange information with one another [11]. These are collectively known as Resting-State Networks (RSNs). Some RSNs are actually more active during rest, with the Default Mode Network (DMN) being the most prominent. This distinctive constellation of brain regions is widely studied to measure the brain's intrinsic activity [13]. Researchers study these RSNs to measure the lasting changes in brain connectivity that result from lived experiences. For instance, researchers have found differences between mathematicians and non-mathematicians in the functional organization of several RSNs [14].

One may think that without a carefully designed task, participants' brain activity may be too variable for drawing any broader conclusions. However, RSNs are remarkably consistent between individuals because they reflect the underlying functional organization of the brain. In addition, these studies require fewer data points (5-10 minutes) for sufficient statistical power since they consist of continuous activity over a period of time, rather than event-related responses with a beginning, middle, and end [15]. Researchers can use nuances in these networks to infer differences between study populations. For instance, researchers have shown that brain activity of expert musicians was more correlated than that of non-musicians within a specific brain region [16]. These measurable differences are due to the brain's *plasticity*, where its structural and functional organization changes in response to environmental and genetic factors [17]. Researchers can measure how sustained patterns of co-activation during wakefulness can lead to lasting changes in the brain's functional organization at rest [18]. While increased time spent on coding activities may not make someone an expert, we hypothesize that sustained patterns of activity from time spent on programming activities (i.e., code reading, code writing, debugging, cod review, etc.) will lead to lasting changes in the functional organization of the brain, observable at rest.

We present an analysis of RSNs from 150 participants' resting state fMRI scans. This data originates from four fMRI datasets (collected on the same scanner at the University of Michigan from 2016 to 2021). Three of these datasets consist of programmers who were all graduate and undergraduate students, while one consists of non-programmers who were primarily STEM students. We first analyzed differences between programmers and non-programmers to explore the lasting impacts of coding activities on brain connectivity. We then analyzed differences between novice, intermediate, and expert programmers to study how these patterns of brain connectivity might progress. We found that programmers exhibit more connectivity from language areas to those involved in temporal attention and mathematical processing. By contrast, non-programmers express more connectivity with regions associated with social cognition and emotional processing. Furthermore, we found that a functional connection between two brain regions *decreases* with more years of coding experience. Our contributions include the following:

- The first resting state fMRI analysis in SE.
- The first systematic neurological analysis regarding programming expertise. Our study revealed programmers have increased connectivity between attention, language, and math regions, and decreased connectivity with regions associated with social and emotional cognition.
- The first neurological comparison between programmers and non-programmers.
- The largest combined fMRI dataset for studying the cognition of coding.
- Data available upon request and analysis code here [1].

II. BACKGROUND AND RELATED WORK

In this section, we describe fMRI in SE, background information of RSNs, and previous work in Cognitive Neuroscience exploring the effects of expertise in RSNs.

A. *Functional Magnetic Resonance Imaging in SE*

Essentially, fMRI gives insight into brain activity by measuring oxygenated blood flow in the brain using the Blood-Oxygen Level Dependent (BOLD) signal [11]. Researchers in SE have used fMRI to study various coding tasks, such as code comprehension [19], code writing [7, 20], debugging [10], code complexity [9], and even gender biases during code review [21], among others [8, 22]. Those task-based experiments have reported the significance of regions associated with language [8], salience detection [10], mathematics [7], and mental rotation [22]. Previous studies have also used classification to inform our understanding of the cognition of coding, finding that specific brain regions can be used to distinguish between code reading and prose reading [8]. Those findings have begun to paint an impressionist picture of the cognitive skills involved in programming, to which we contribute in this study by uncovering differences in the network activity of programmers at rest.

B. *Resting State Scans*

To infer network characteristics of the brain at rest, researchers can use a technique called *functional connectivity analysis*, which calculates the correlation of brain areas' activity over time. This type of analysis can effectively be applied to both task-based experiments and resting state scans [11, 23], and has been used in SE to identify the brain regions that coordinate in code writing tasks [7]. When humans are at rest, regions in their brains are not responding to any external stimulus, but are nonetheless still active and selectively correlated with one another [15]. This, coupled with evidence that sustained patterns of co-activation can lead to lasting changes in the brain, implies there are measurable correlations at rest that are indicative of neurological changes over time. Leveraging this insight, researchers have analyzed resting state data to find meaning in the brain regions whose activity is significantly correlated at rest [16]. In this manner, resting state scans are not constrained by a task specifically designed to elicit certain types of activity, and can give insight into the spontaneous yet patterned activity of all RSNs.

While analyzing resting state scans can offer many benefits, there are important challenges related to extraneous sources of noise, such as head movement or even respiration in the scanner [15]. Researchers typically perform motion correction during preprocessing, and correct physiological noise using a library called RETROICOR [24], or Independent Component Analysis (ICA). The latter separates the brain signal into spatially independent patterns of activity [23], and enables the isolation of components associated with signal, motion, and physiology, and removal of components associated with the latter two. There are numerous analysis techniques for functional connectivity analysis for resting state fMRI, with the most common being ICA (as mentioned above) and seed-based connectivity analysis [11]. In seed-based analyses, researchers choose regions *a priori* and calculate their correlation with other brain areas. These "other brain areas" can be individual voxels in a seed-to-voxel analysis, or other populations of voxels in a Region-of-Interest (ROI)-to-ROI analyses. These seed regions are chosen based on prior literature and researchers' hypotheses, and can be defined using parcellation maps of the brain or spheres around voxels [25]. Seed regions are made up of numerous voxels, so as a first step researchers calculate

---

[1] https://github.com/largehappygroup/resting_state_connectivity/tree/main/better_replication_package

an average timecourse from this population of voxels. This average timecourse then serves as a reference to which all other voxel or region timecourses are correlated [11]. Through this process, researchers can uncover voxels or brain areas whose activity is highly correlated with that of the seed region. Averaging timecourses in this manner can also help to improve the signal-to-noise ratio since individual voxel timecourses are susceptible to the artifacts mentioned above, but broader regions can cancel out these local artifacts [23].

C. *Analyzing Expertise in Cognitive Neuroscience*

Though a resting state analysis of expertise in programming has not been implemented, neuroscientists have explored the effect of expertise in other domains using resting state data. A key principle in Neuroscience is that of plasticity, where the brain changes in response to lived experiences and genetic expression [17]. Plasticity is a broad term that can refer to neurological changes related to natural development and traumatic events alike [17]. Researchers have studied how experience with a skill can measurably alter the functional organization of one's brain at rest [18], causing an increase or decrease in functional connectivity. This has been reported in physical tasks such as gymnastics [26] and endurance running [27], along with cognitive tasks such as meditation [28] and math [14]. For example, in runners, researchers found an increase in functional connectivity between regions involved in motor control and executive function, reasoning that endurance running requires sustained focus [27]. Interestingly, greater experience in a domain can lead to *decreases* in functional connectivity as well. For instance, researchers found that mathematicians expressed lower levels of functional connectivity between the Left and Right Caudate Nuclei [14]. Changes in functional connectivity are not inherently good or bad, but the improvement in a skill can be reflected in the resultant cognitive changes. In this study analyzing resting state data of programmers, we also found that an increase of coding experience is associated with a decrease in functional connectivity between the Left Lingual Gyrus and the Left Precentral Gyrus. These regions are involved in attention during reading and articulation, respectively, and may reflect increased cognitive efficiency from time spent on programming tasks.

III. METHODS

In this section, we present details related to the resting state datasets, procedures used for data preprocessing, and the analysis of the resting state data at the subject and group levels.

A. *Datasets*

The resting state scans used in this study originate from four fMRI datasets, three of which are from computer science studies and publicly available, and one is from a Cognitive Neuroscience study. The original teams of researchers obtained IRB approval to conduct the studies, and used standard fMRI data collection procedures [29]. All studies used the same scanner for data collection. The three computer science datasets were collected using the same scanning parameters, while the parameters in the fourth dataset were slightly different, which we detail in this section.

Participants Our total dataset consists of resting state scans from four studies and 150 participants (64 females), ranging from 18–29 years old (mean = 22.248, SD = 2.556). We split this dataset into two samples according to the alignment of fMRI preprocessing procedures and the formatting of demographic data related to programming experience. We label these datasets A, B, C, and D, and detail their demographic information in Table I. All Computer Science students from datasets B, C, and D (n = 96) were graduates and undergraduates, recruited from the same upper-level Computer Science course in different years. Collected demographic information included age, gender, sex, and native language. Dataset C recorded coding experience in years [22], while dataset D recorded experience in semesters [20]. Participants in dataset B reported their coding experience either cumulatively, qualitatively, or by language [21]. In this study, we compared CS Students in datasets C and D based on their years of coding experience, and converted semesters in dataset D to years by counting two semesters as one year. We excluded dataset B from our expertise analysis because of inconsistencies in the formatting of participants' reported coding experience. In comparing programmers and non-programmers, we used studies A and B since their preprocessing was more aligned (which we detail in Section III-B). For the expertise analysis among programmers, we split datasets C and D into beginners, intermediates, and experts based on years of coding experience. This allowed us to analyze connectivity patterns along an approximate continuum, with the intermediate group also providing a sharper contrast between novices and experts. This information is detailed in Table II. Dataset A included information about the educational backgrounds of 38 participants, 31 of whom were STEM majors. Three participants who majored in CS were excluded from this dataset. Some participants for whom we did not have major information may have had CS experience, which we discuss in Section VI.

TABLE I: Demographics of our sample of participants ($n = 150$). In our analyses, we grouped dataset A with B, and dataset C with D. Datasets A and B were consistent in their preprocessing, while C and D were consistent in the information collected related to coding experience.

Label	Education	n	Avg. Age	Women	Men	Coding Experience
A	Non-CS	54	23.1	26	28	—
B	CS	37	21.9	16	21	mixed values
C	CS	33	22.1	14	19	in years
D	CS	26	21.1	9	17	in semesters

TABLE II: Group partitions of programmers from Datasets C and D, based on their years of coding experience.

Group	Years of Experience	n	Avg. Age	Women	Men
Novices	< 3 years	23	20.3	8	15
Intermediates	3 – 4 years	23	21.6	11	12
Experts	> 4 years	13	24	4	9

Scanning Parameters - Dataset A Scans were collected on an 3T MR750 GE Scanner with an 8-channel head coil. High resolution anatomical scans (T_1-weighted SPGR, 156 slices) were collected first with $TI = 500ms$, $FA = 15°$, and $FOV = 256mm^2$. Resting state scans (2D Gradient Echo spiral pulse sequence, 43 axial slices, $3mm$ thickness and no spacing) were collected in an interleaved, bottom-up sequence with TR = $2000ms$, TE = $30ms$, FA = $90°$, FOV = $220mm^2$. Eight minutes of resting state were collected from each participant at a slower TR, yielding 240 volumes for each.

Scanning Parameters - Dataset B, C and D Scans were collected on the same 3T MR750 GE Scanner, but with a 32-channel head coil. High-resolution anatomical scans (T_1-weighted spoiled gradient recall sequence, 208 slices, $1mm$ thickness) were collected first with TR = $2300.80ms$, TE = $24ms$, TI = $975ms$, and FA = $8°$. An estimate of the magnetic field homogeneity within the scanner was collected using a spin-echo fieldmap with a TR = $7400ms$, TE = $80ms$, and slice thickness of $2.4mm$. The resting state data (T_2^*-weighted multiband echo planar imaging sequence, 60 slices, $2.4mm$ thickness) was collected with a TR = $800ms$, TE = $30ms$, FA = $52°$, and an acceleration factor of 6. Including slice thickness, the isotropic voxel dimensions were [$2.4mm \times 2.4mm \times 2.4mm$]. Eight minutes of resting state were collected from each participant, yielding 600 volumes for each. We acknowledge the discrepancies in scanning parameters between datasets, and discuss our mitigation strategies in Sections III-C and VI (Limitations).

Atlases To standardize neuroimaging analysis and results across individuals, researchers use atlases such as the MNI152 template to align (register) participants' brain images [30]. This enables a shared coordinate space across participants and studies. Research also aims to map cognitive processes to specific regions and networks of the brain, with active research into how to best distinguish, or *parcellate*, brain areas. In this study, we used the Schaefer Atlas, which parcellates the cortex of the brain into 400 distinct regions [31]. We used this parcellation scheme in formalizing seed regions for functional connectivity analysis, with more details in Section III-C.

B. *Preprocessing*

To preprocess the resting state fMRI data for functional connectivity analysis, we implemented a standard pipeline using established libraries and methodologies [32]. The first 12 volumes were removed to eliminate effects of thermal noise and scanner drift. The scans were slice-timing corrected, then datasets C and D were unwarped [33]. Datasets A and B were not unwarped due to missing fieldmap data. Datasets A and B were corrected for physiological noise at this stage using RETROICOR [24]. The physiological data from datasets C and D were low quality, so we performed ICA correction at a later stage to remove physiological noise (detailed below). We then extracted the brain from the anatomical images of all datasets, and registered each to MNI152 space using affine then non-linear registration [34]. We then applied motion correction to the functional scans to align the volumes, and recorded motion parameters along 6 axes [35, 36]. Next, the functional scans were registered with the anatomical scans, and then registered to MNI152 space [34]. Next we applied gaussian smoothing ($3mm$ kernel) to the functional scans [35], which can help improve both the signal-to-noise ratio [32], and reliability for merging datasets [32].

We performed ICA on datasets C and D to correct for physiological noise, decomposing each participant's functional scan into 60 components [35]. For 16 participants we hand-classified each component as signal or noise, which remains the gold standard for identifying noise [37]. In classifying components as either signal or noise, the authors agreed on the criteria, considering: (1) the frequency of components, with higher frequencies suggesting cardiac signals; (2) the power spectra of components, supporting frequency information; (3) the location of components, noting that those in cerebrospinal fluid or arteries do not indicate neural activity; and (4) the location of components around the brain's outer rim, characteristic of motion artifacts. These hand-classifications were used as training data to create a study-specific classification [35], which we then applied to remove noisy components from all participants' data in these two datasets (threshold of 20). Finally, we regressed out the motion parameters recorded above, their first derivatives, and a linear trend, from each voxel's timecourse in all datasets for further motion correction.

C. *Analysis*

To study functional changes in the brain from time spent on coding activities, we used seed-based functional connectivity to first compare programmers and non-programmers, then compare programmers amongst themselves based on their years of coding experience. Here we discuss our reasoning in selecting seed regions and our strategies for subject- and group-level analyses.

Seed Regions The results from task-based fMRI experiments in SE were integral in identifying key brain regions involved in coding activities. Specifically, previous studies have reported the involvement of the Left Inferior Frontal Gyrus (LIFG, Broca's Area) in code comprehension [9, 19] and code writing [20]. Significant functional connectivity was also reported during code writing between the LIFG and the Right Inferior Temporal Gyrus (RITG) [7, 38], known in Math Education research as the Number Form Area [38]. Previous research has shown that list operations elicit activity in both the Left Inferior Temporal Gyrus (LITG), also known as the Visual Word Form Area (VWFA), and the Right Superior Parietal Lobule (RSPL) [22]. The RSPL has been implicated in visuospatial processing, and was also found to exhibit patterns of functional connectivity that were affected by participants' GPA during code writing [7]. Due to the significance of the DMN, we also considered both hemispheres of one component in our analyses, the Right and Left Posterior Cingulate Cortices (RPCC/LPCC). These 6 areas define our regions of interest, which we localized to parcels of the Schaefer Atlas. In localizing these areas, we referred to MNI coordinates with peak connectivity reported in previous research [7], as well as figures of activation maps [7, 20, 22]. There was a degree of subjectivity involved in choosing seed regions, which we address in Section VI. The 6 seed regions we considered, their indices in the Schaefer Atlas, and the MNI coordinates within these parcels are detailed in Table III.

TABLE III: Seed regions for functional connectivity analyses, based on previous neuroimaging research in SE. MNI Coordinates refer to coordinates with peak activity and connectivity in prior research. Specific parcels were chosen by referring to MNI coordinates and activation maps in previous studies.

Region	MNI Coordinates	Schaefer Parcel
L Inferior Temporal Gyrus	[-59, -44, -10]	133
L Inferior Frontal Gyrus	[-50,30,-2]	172
L Posterior Cingulate Cortex	[-7, -49, 27]	192
R Inferior Temporal Gyrus	[61, -44, -9]	339
R Superior Parietal Lobule	[26, -42, 64]	284
R Posterior Cingulate Cortex	[8, -49, 27]	395

Subject and Group Level Analysis To measure the correlation of brain activity between brain regions, we performed an ROI-to-ROI functional connectivity analysis. We used parcels of the Schaefer Atlas to delineate these regions, and calculated the correlation from our seed regions to all other parcels in the Schaefer Atlas. For example, the LIFG corresponds to parcel 172 in the Schaefer atlas, which is comprised of 378 voxels. We averaged all the voxels' timecourses together within this parcel to obtain a mean timecourse, and did the same for the other 399 parcels. Using the mean timecourse from parcel 172, we then calculated its Pearson Correlation Coefficient with all 400 parcels' mean timecourses. We repeated this process for the other five seed regions to obtain measures of functional connectivity between brain regions. We also computed a correlation matrix by correlating every parcel's timecourse with one another. We first calculated correlation coefficients for each participant, then performed Fisher's z-transformation to calculate the coefficients' inverse hyperbolic tangents. This process is standard in fMRI analysis, and helps to normally distribute the data [39]. After computing these z-scores, we determined significant differences between groups' functional connectivity using two-sample t-tests on z-scores. We then corrected for multiple comparisons using the Benjamini-Hochberg method [40]. Lastly, our experts were significantly older than our novices ($p < 0.001$), so we performed linear regression to analyze whether our findings were attributable to years of coding experience or general effects of age. In Cognitive Neuroscience, t-tests are often used to assess group differences in resting-state date [28], while linear regression can be used to analyze relationships between brain measures and confounding variables like age, or factors like test scores [41].

Gaussian Process Classification We explored differences in our sample using hand-picked seed regions, but also sought to find important regions for delineating our sample using a data-driven approach, while still correcting for multiple comparisons. To this end, we used Gaussian Process Classification (GPC) to find regions informative for classifying between programmers based on their experience. This multivariate analysis has previously been used in neuroimaging studies for SE [8, 21], and considers whole-brain activity patterns in classifying participants' fMRI data. The study by Floyd *et al.* contains a more detailed explanation, but we give a brief description of relevant information for this study. This method classifies participants by first compressing fMRI data into a feature space using a linear kernel. Since the kernel is linear, the calculated feature space can be projected back into the original data space, enhancing the model's explainability. In other words, the decision function can be visualized as weights mapped to specific brain regions. Subsequently, we used the regions given high weights by the model as seed regions for an additional functional connectivity analysis. In training the model, we used an iterative leave-one-out cross-validation scheme. By classifying each participant, this approach yields a percent accuracy for classifying both groups, and an average Balance Accuracy (BAC) for the model as a whole.

IV. RESULTS

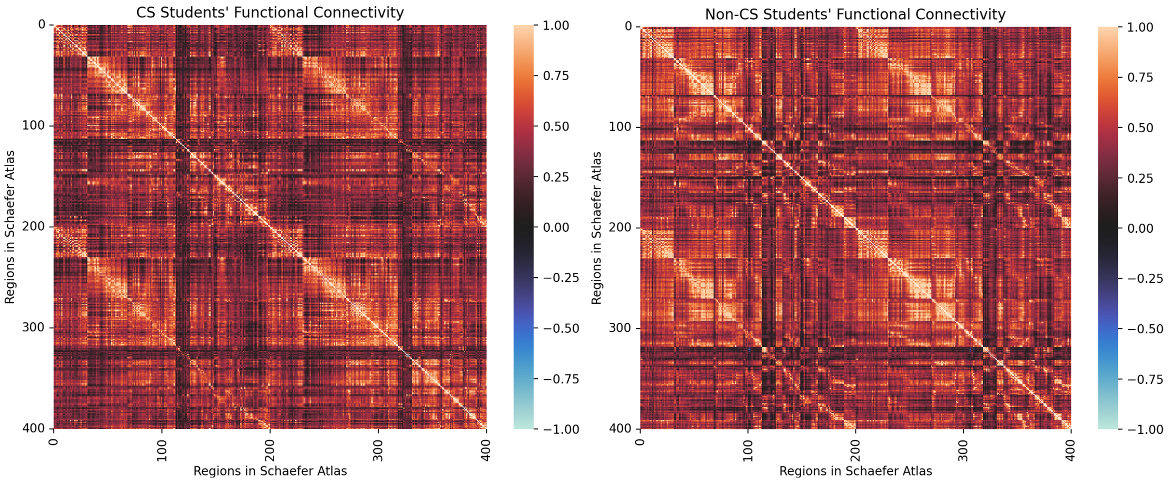
In this section, we discuss results from functional connectivity analyses between programmers and non-programmers, and between novice, intermediate, and expert programmers. We frame these into the following research questions:

- **RQ1** How do brain connectivity patterns of programmers compare to those of non-programmers?
- **RQ2** Among programmers, how are brain connectivity patterns mediated by experience?

A. *RQ1*

To understand how coding activities over time can influence the neural patterns of programmers, we first compared programmers and non-programmers. We conducted seed-based functional connectivity analyses on datasets A (n = 54) and B (n = 37). The datasets are consistent in their preprocessing, but differ in their composition: dataset A includes participants without a record of programming experience, while dataset B consists of graduate and undergraduate CS students. To investigate group-level patterns, we first examined the functional connectivity between all parcels (n = 400), yielding a 400×400 correlation matrix for each participant. Since the scanning parameters were slightly different between the two datasets, we mitigated nuances in data collection by performing Fisher's z-transformation separately for both studies. In this manner, the values in participants' correlation matrices represent scaled coefficients that follow a normal distribution for each group between -1 and 1. We then averaged these matrices across participants for both studies, yielding group-wide patterns of connectivity (Fig. 1). In this study, we primarily focus on seed-based analyses, which compare the statistical relationship of *one* seed region to the activity in every other cortical region. By contrast, the brain-wide measures of functional connectivity depicted in Figure 1 show how the activity of *every* region compares to that of every other region.

From these diagrams of functional connectivity, there are noticeable qualitative differences between programmers and non-programmers. Specifically, programmers exhibit lower levels of functional connectivity (darker bands) with regions in the medial frontal lobe (parcels 114-120, 319-324). By contrast, non-programmers appear to have more variety in their patterns of functional connectivity in these regions, perhaps influenced by heterogeneity in this group's educational background. Also, compared to non-programmers, programmers



(a) (b)

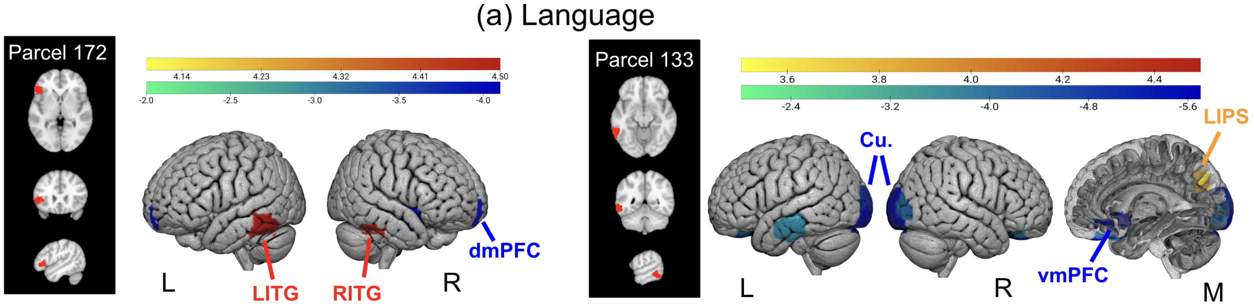
Fig. 1: Group-level functional connectivity measures for (a) programmers and (b) non-programmers. Values in these matrices represent the functional connectivity (i.e., Pearson correlation) from every cortical brain region to every other cortical brain region. Correlation coefficients in these figures were Fisher's z-transformed to normalize their distribution. These values now denote scaled correlation values following a normal distribution between -1 and 1.

demonstrate patterns of higher functional connectivity between regions around parcels 75 and 275. These parcels correspond to parietal areas on the left and right hemispheres of the brain, and may illuminate differences in functional connectivity from extended time spent on programming tasks. This can be informative for future research into additional areas that may specialize for programming-related tasks. To gain deeper insight into key regions of interest, we then conducted an ROI-to-ROI functional connectivity analysis using six seed regions based on previous neuroimaging studies in SE. Broadly, these regions can be grouped into categories related to Language (LIFG, VWFA), Math (RITG, RSPL), and the DMN (LPCC, RPCC). These regions, along with their Schaefer parcel numbers and MNI coordinates are summarized in Table III.

Again using an ROI-to-ROI functional connectivity analysis, we compared programmers and non-programmers based on their Fisher's z-transformed correlation values from the seed regions to all other parcels in the Schaefer atlas. To uncover group-level differences, we performed two-tailed t-tests between these values for programmers and non-programmers, and corrected for multiple comparisons using the Benjamini-Hochberg method ($q < 0.01$). This analysis uncovered numerous significant differences between programmers and non-programmers that survived correction for multiple comparisons. These differences in functional connectivity can be seen in Figure 2, with the five most significant results below the threshold for each seed region reported in Table IV. All t-statistics and q-values are included in Supplementary Material. To summarize the main findings here, programmers demonstrated significantly higher levels of functional connectivity between the LIFG and three bilateral parcels in the ITG (**parcel 330**: t = 4.573, $q < 0.01$; **parcel 151**: t = 4.545, $q < 0.01$; **parcel 69**: t = 4.053, $q < 0.01$). The Left ITG is well-studied as the Visual Word Form Area, where words are theoretically decoded, with the information then passing to Broca's Area in the LIFG [41]. The Right ITG is theorized to have a comparable role for decoding numbers [38], and also performing arithmetic [42]. We chose seed regions near these areas based on prior SE studies, and here find Broca's Area is more functionally connected in programmers to both.

For non-programmers, we found the same seed region expressed significantly more functional connectivity with the Right Anterior Insula (**parcel 302**: t = 4.436, $q < 0.01$) and with the Left dorsomedial Prefrontal Cortex (dmPFC, **parcel 168**: t = 4.081, $q < 0.01$). These regions have been implicated in social processing [43], as well as self evaluation [44], respectively. These results showing stronger connections between the LIFG and the bilateral ITG suggest that Broca's Area may be enlisted more for language or information processing in programmers, as opposed to emotional processing. Next, we found that programmers had significantly higher rates of functional connectivity between the VWFA and a parcel in the Left Intraparietal Sulcus (LIPS, **parcel 73**: t = 3.858, $q < 0.01$), a region often implicated in the temporal ordering of attention [45]. Non-programmers had significantly higher functional connectivity between the VWFA and 12 other parcels, with the most significant being bilaterally in the Cuneus (**parcel 221**: t = 6.414, $q < 0.0001$; **parcel 12**: t = 5.872, $q < 0.0001$; **parcel 213**: t = 5.651, $q < 0.0001$), and the bilateral vmPFC (**parcel 117**: t = 5.635, $q < 0.0001$; **parcel 323**: t = 5.605, $q < 0.0001$). The Cuneus has reportedly been involved in emotional perception and complex communicative tasks [46, 47], while the vmPFC has been implicated in emotional regulation [48] and memory consolidation [49]. A trend thus emerges where programmers express lower levels of functional connectivity between our seed regions and areas involved in emotional and social processing, which may suggest a reassignment of certain brain areas for information processing in programmers.

For regions related to math, we calculated the parcels that



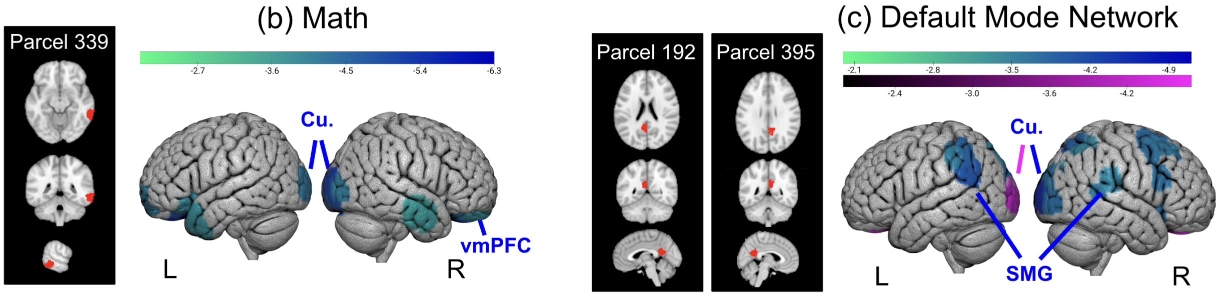


Fig. 2: Significant differences between programmers and non-programmers in patterns of functional connectivity from seed regions to other parcels ($q < 0.01$). Warm colors (red, yellow) represent t-values of regions where programmers had higher levels of functional connectivity. Cool colors (blue, teal, pink) show t-values where *non*-programmers had higher levels of functional connectivity. Seed regions are depicted to the left of the brain plots, grouped by function: (a) Language: parcel 172 is Broca's Area, a primary region of the language network involved in speech production and hierarchical processing. Parcel 133 is the Visual Word Form Area, theorized to decode letters and words from the visual system. (b) Math: parcel 339 is the Number Form Area, involved in arithmetic and the visual processing of Arabic numbers. (c) The Default Mode Network: parcels 192 (Blue) and 395 (Pink) are located in a hub of the DMN, which is more active at rest, involved in episodic memory processing and introspection. There were no significant differences for parcel 284, which was omitted from the figure. Brain acronyms are as follows: (L/R)ITG - Left/Right Inferior Temporal Gyrus, dmPFC - dorsomedial Prefrontal Cortex, Cu. - Cuneus, vmPFC - ventromedial Prefrontal Cortex, LIPS - Left Intraparietal Sulcus, SMG - Supramarginal Gyrus.

were functionally connected with the RSPL and the NFA. There were no significant differences between programmers and non-programmers in their patterns of functional connectivity with the RSPL. For the NFA, compared to non-programmers, programmers did not demonstrate significantly more functional connectivity with any other parcels. However, non-programmers showed significantly more functional connectivity between the NFA and 18 other parcels in the Schaefer Atlas. Again, this was especially pronounced bilaterally in the vmPFC (**parcel 117**: t = 6.606, $q < 0.0001$; **parcel 323**: t = 6.472, $q < 0.0001$) and bilatarally in the Cuneus (**parcel 221**: t = 5.35, $q < 0.0001$; **parcel 114**: t = 5.13, $q < 0.0001$). Even though the NFA has classically been associated with math and visual processing, here we found non-programmers have higher rates of functional connectivity from this region to areas associated with emotion and self reflection. A relation between this area and the regulation of negative emotions has been reported [50], but not investigated as thoroughly.

Programmers and non-programmers deviated the most in their functional connectivity with respect to the LPCC and RPCC in the DMN. Compared to programmers, non-programmers had significantly more functional connectivity between the LPCC and 35 other parcels. This was again most significant bilaterally in the vmPFC (**parcel 323**: t = 5.269, $q < 0.001$; **parcel 117**: t = 5.058, $q < 0.001$), the Cuneus (**parcel 19**: t = 5.034, $q < 0.001$; **parcel 221**: t = 4.755, $q < 0.001$), and in the Right Anterior Cingulate Gyrus (RACG, **parcel 384**: t = 4.92, $q < 0.001$). The RPCC expressed a similar pattern of functional connectivity, where programmers did not exhibit significantly higher levels of functional connectivity than non-programmers did between this region and other parcels. However, non-programmers demonstrated significantly more functional connectivity between the RPCC and four other parcels, with the most prominent being bilaterally in the vmPFC (**parcel 323**: t = 4.769, $q < 0.01$; **parcel 117**: t = 4.646, $q < 0.01$), the RACG (**parcel 384**: t = 4.215, $q < 0.01$), and the Left Cuneus (**parcel 19**: t = 4.156, $q < 0.01$).

Programmers exhibited higher levels of functional connectivity between the Visual Word Form Area and the Left Intraparietal Sulcus ($q < 0.01$), regions involved in word processing and temporal attention, respectively. Non-programmers had higher levels of functional connectivity between seed regions and bilaterally in the Ventromedial Prefrontal Cortex and the Cuneus ($q < 0.01$), which are involved in emotional perception and regulation, respectively.

TABLE IV: Functional connectivity results between seed regions and parcels of the Schaefer Atlas, including t-statistics and q-values, corrected for Multiple Comparisons with the Benjamini-Hochberg method ($q < 0.01$). Schaefer parcel numbers corresponding to the brain areas are listed next to region names. We omitted the RSPL seed region from the table (parcel 284), because its functional connectivity patterns were not significantly different between programmers and non-programmers.

Seed Regions	Schaefer Parcel	Programmers > Non-Programmers	Non-Programmers > Programmers
Visual Word Form Area Decoding words	133	(73) **LSPL** (t = 3.858, $q < 0.01$)	(12) **L Cuneus** (t = 5.872, $q < 0.0001$) (117) **L vmPFC** (t = 5.635, $q < 0.0001$) (213) **R Cuneus** (t = 5.655, $q < 0.0001$) (221) **R Cuneus** (t = 6.414, $q < 0.0001$) (323) **R vmPFC** (t = 5.605, $q < 0.0001$)
Broca's Area Speech Production, Hierarchical Processing	172	(69) **LITG** (t = 4.053, $q < 0.01$) (151) **LITG** (t = 4.545, $q < 0.01$) (330) **RITG** (t = 4.573, $q < 0.01$)	(168) **L vmPFC** (t = 4.081, $q < 0.01$) (302) **R Insula** (t = 4.436, $q < 0.01$)
L Posterior Cingulate Cortex Default Mode Network	192	—	(19) **L Cuneus** (t = 5.034, $q < 0.001$) (117) **L vmPFC** (t = 5.058, $q < 0.001$) (221) **R Cuneus** (t = 4.755, $q < 0.001$) (323) **R vmPFC** (t = 5.269, $q < 0.001$) (384) **RACG** (t = 4.92, $q < 0.001$)
Number Form Area Number recognition, Mathematical Processing	339	—	(114) **L Cuneus** (t = 5.13, $q < 0.001$) (117) **L vmPFC** (t = 6.606, $q < 0.0001$) (213) **R Cuneus** (t = 4.958, $q < 0.001$) (221) **R Cuneus** (t = 5.35, $q < 0.0001$) (323) **R vmPFC** (t = 6.472, $q < 0.0001$)
R Posterior Cingulate Cortex Default Mode Network	395	—	(19) **L Cuneus** (t = 4.156, $q < 0.01$) (117) **L vmPFC** (t = 4.646, $q < 0.01$) (323) **R vmPFC** (t = 4.769, $q < 0.01$) (384) **RACG** (t = 4.215, $q < 0.01$)

Brain region acronyms: LSPL - Left Superior Parietal Lobule, vmPFC - ventromedial Prefrontal Cortex, (L/R)ITG - Left/Right Inferior Temporal Gyrus, RACG - Right Anterior Cingulate Cortex.

*B. **RQ2***

To understand how more time spent on coding activities can have a lasting impact on the functional organization of the brain, we next compared novice, intermediate, and expert programmers. Splitting the sample into three groups allowed us to analyze how neurological changes might progress. The intermediate group provides both a buffer between novices and experts for t-tests, and a continuum to analyze with linear regression, which we detail below. For this research question, we analyzed resting state scans from datasets C (n = 33) and D (n = 26), consisting entirely of programmers. To uncover trends of functional connectivity over time, we split our sample into three groups based on their years of coding experience: novices ($y.e. < 3$), intermediates ($3 <= y.e. <= 4$), and experts ($y.e. > 4$). We used the same six seed regions for an ROI-to-ROI functional connectivity analysis for the three groups, and performed two tailed t-tests between participants' z-scored correlation values. After correcting for multiple comparisons, no significant correlations survived. Factors contributing to this null result may include our relatively small sample size, the limited range of participants' years of coding experience, and the inherently low effect sizes in resting state data [51]. We discuss these factors further in Section VI.

None of our six hand-picked seed regions demonstrated functional connectivity differences that survived correction for multiple comparisons. However, there is a possibility that group differences relevant for coding activities were present in regions we did not consider as seeds. To test this, we implemented a data-driven analysis by first classifying the groups using GPC, then deriving seed regions based on the brain regions deemed important for distinguishing between groups. We again corrected for multiple comparisons to mitigate the false discovery rate. We first performed binary classifications between novices and intermediates, then between novices and experts. Our classification attained BAC accuracies above chance of 60.87% for novices and intermediates, and 63.38% for novices and experts. Using the brain map of feature weights, we isolated the clusters with the highest weights [35]. Specifically, we arbitrarily extracted the 5 largest clusters in the two classifications influencing the model to decide novice, intermediate, or expert. Many of these clusters were located in the cerebellum, which is not covered by the Schaefer Atlas. We therefore performed functional connectivity analysis using the Schaefer parcels associated with these cortical clusters as seeds, which are detailed in Table V.

TABLE V: Data-driven seed regions based on clusters that were ascribed high weights for classifications between groups. MNI Coordinates here represent peak weights in the clusters.

Group	Region	MNI Coordinates	Parcel
Experts	L Lingual Gyrus	[-24, 80, -18]	9
Experts	L Postcentral Gyrus	[-16, -30, 78]	67
Interm.	L Superior Frontal Gyrus	[-20, 70, 4]	176
Novices	L Superior Frontal Gyrus	[-7, -49, 27]	183
Interm.	R Superior Temporal Gyrus	[60, 8, -2]	231
Experts	R Postcentral Gyrus	[62, 10, 16]	242
Interm.	R Middle Frontal Gyrus	[22, 68, -6]	341
Interm.	R Middle Frontal Gyrus	[34, 62, -2]	343
Novices	R Middle Frontal Gyrus	[26, 60, 26]	348
Novices	R Frontal Pole	[2, 68, 0]	380
Interm.	R Superior Frontal Gyrus	[14, 52, 40]	386
Novices	R Superior Frontal Gyrus	[10, 34, 58]	389
Novices	R Superior Frontal Gyrus	[10, 20, 62]	391

We conducted another ROI-to-ROI functional connectivity analysis between the data-driven seed regions listed in Table V, and all other parcels in the Schaefer atlas. We similarly applied a Fisher's z-transformation, and conducted two-tailed t-tests between the z-scores of novices, intermediates, and experts. We again corrected for multiple comparisons using the Benjamini-Hochberg method. After this process, we found a significant difference between experts and novices in the level of functional connectivity between a seed region in the Left Lingual Gyrus (parcel 9) with a region in the Left Precentral Gyrus (PreCG, parcel 40). Specifically, as participants' years of coding experience increases, their functional connectivity between these two regions *decreases* (t = 4.469, $q < 0.05$). This difference may be innate among participants, so a longitudinal study would be required to establish causality. That being said, the Lingual Gyrus has been linked to visuospatial processing during reading and the perception of physical objects [52, 53], while the PreCG has been associated with articulation and oral movements [54]. Considering code as a 'language' that is primarily comprehended rather than spoken, these results suggest that more time spent on programming tasks may lead to a dissociation between these regions involved in visual processing during reading and oral movements.

Age is a potential confounding variable, where programmers with more years of experience are likely to be older. Indeed, participants in the novice group are significantly younger than those in both the intermediate ($p < 0.05$) and the expert group ($p < 0.001$). As such, the significant difference in functional connectivity between parcels 9 and 40 may be due to age related effects rather than programming activities. To explore this possibility, we performed Ordinary Least Squares regression with age and years of experience as predictor variables, and functional connectivity measures between parcels 9 and 40 as the dependent variable. We analyzed novices, intermediates, and experts together (n = 59) to both improve statistical power and depict a continuum of coding experience. We found that years of coding experience is a strong and statistically significant predictor of functional connectivity between these two regions (β = -0.041, t = -2.788, $p < 0.01$), while age is not (β = 0.011, t = 0.610, p = 0.5). Thus, we can conclude with more certainty that this decrease in functional connectivity is attributable to years of coding experience, and not age (Fig. 4). Our results indicate that age is actually associated with a slight *increase* in functional connectivity between these two regions.

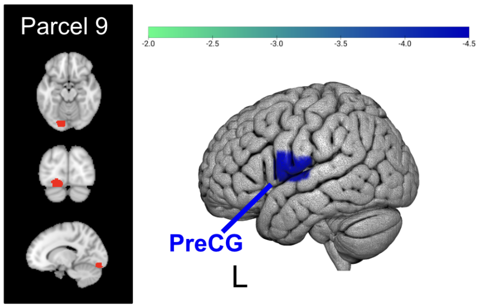


Fig. 3: As years of coding experience increases, there is a significant decrease in functional connectivity between a region in the Left Lingual Gyrus and the Left Precentral Gyrus ($q < 0.05$). PreCG - Precentral Gyrus.

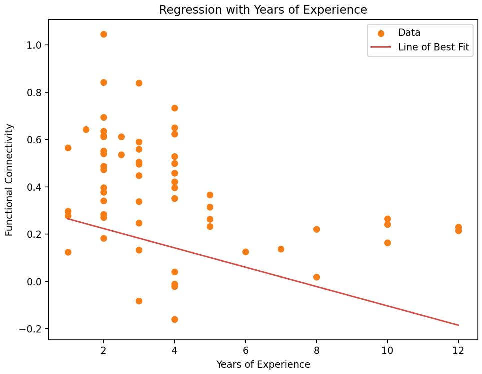


Fig. 4: Functional connectivity between the Left Lingual Gyrus (Schaefer parcel 9) seed region and a region in the Precentral Gyrus (Schaefer parcel 40) as it relates to years of coding experience. The line of best fit was calculated using the y-intercept (0.3056) and beta weight from the linear model detailed in Section IV-B.

A decrease in functional connectivity between regions that comes with more experience may be counter-intuitive, but a similar pattern has been reported in Cognitive Neuroscience. In elite gymnasts for instance, there is lower functional connectivity within the cerebellum, an area known for motor control, and within fronto-parietal and cingulo-opercular networks [26]. One may think that an elite gymnast would have higher functional connectivity within motor and attention regions. However, those authors interpreted the results to suggest that elite gymnasts were more automatic and efficient in their movements, and therefore benefited from less connectivity. In this study, this difference in our participants may be intrinsic, and it may also be task-dependent. Nonetheless, we can interpret our findings to suggest that weakened functional connectivity between these two regions may reflect enhanced neural efficiency for programming tasks.

> As programmers' years of experience increases, the functional connectivity decreases between a region in the Left Lingual Gyrus and a region in the Precentral Gyrus ($q < 0.05$), two regions known for visual processing during reading and articulation/oral movements, respectively. This pattern of functional connectivity is not explained by age, a possible confound ($p < 0.01$).

V. DISCUSSION

In this study, we uncovered meaningful changes in network activity of the brain due to programming experience. Here we present further interpretations of our findings and discuss their potential relevance for CS Education, practice, and research.

Interpretations We found consistent differences between programmers and non-programmers in their patterns of functional connectivity at rest. First and foremost, we found that in programmers, there is a stronger functional link between a low-level language processing region (VWFA) and another region involved in the temporal ordering of attention (LIPS). This connection has not previously been reported in SE, and perhaps reflects the coordination between programmers' visual and temporal attention mechanisms. As a potential explanation, there is an intrinsic order of execution in code, which is not necessarily present in natural language. Furthermore, codebases can have large and complicated structures, which may put additional strain on programmers to maintain a coherent understanding of the code. These factors may promote higher functional connectivity between these two regions for decoding low-level language and ordering temporal attention. We also found strengthened connectivity in programmers between the VWFA and another primary region of the language network, Broca's Area, which is thought to be involved in speech production and hierarchical processing [55]. The functional connection we found between two language regions may be unsurprising, but previous research found the VWFA is more strongly connected in adults to *attention* regions [41]. Those authors posited that the VWFA may in fact be responsible for integrating language and attention processes, so our results suggest that coding activities may *selectively enhance* the VWFA's functional connectivity with a hierarchical language region and a temporal attention region. These functional links present new opportunities for research and training, which we discuss below. Also noteworthy, we replicated previous findings of stronger functional connectivity for programming between Broca's Area and the visual Number Form Area [7], which is thought to decode numeric symbols [38].

Across the chosen seed regions, we found that non-programmers show significantly higher functional connectivity with regions involved in social cognition and emotional processing [46, 47, 48]. We are hesitant to draw conclusions about the social and emotional states of programmers without further carefully designed studies, but these results might imply certain longitudinal impacts from coding activities, and present intriguing avenues to explore in future work or interventions. For example, this may elucidate a cognitive distinction between natural language and code-related tasks, where code is a formal series of instructions, but natural language is inherently social and communicative. Previous study indicates that code and language processes become less distinguishable in coders with higher GPAs [8]. In our study, we considered years of experience, instead of GPA, and found evidence of possible dissociation between language and code; increased years of experience correlated with decreased connectivity between areas involved in reading and articulation. This may indicate nuances in using GPA versus years of experience, where the former may show effects of intensive study or innate differences, and the latter may show changes over time.

Implications Our findings highlighting differences in functional connectivity related to language, attention, and social cognition can be relevant for CS Education, practice, and further research. In training new programmers, these results present opportunities for tailoring educational materials to the cognitive profile of programming. For instance, many programming tools for new programmers are visual and block-based [56]. Our results, along with those of prior neuroimaging studies in SE, consistently underscore the importance of language regions for coding tasks [2, 7, 9]. This might imply an unmet need for early educational materials that target language components of programming, or how to combine language and programming education together. In this study, we found increased connectivity in programmers between language processing (VWFA, BA), temporal attention (LIPS), and mathematical processing regions (NFA). Researchers, educators, and companies could leverage this information to design novel training materials, such as word problems similar to those in math education that use narratives to train concepts like multiplication, but instead for concepts like loops or variables. Informally, if these patterns of connectivity arise naturally after years of programming, perhaps targeted training materials could accelerate this connectivity. Problems involving language and computational thinking may help promote brain connectivity that is beneficial for coding, but proper testing would be necessary. More broadly, our results suggest that targeting language skills, temporal attention, and mathematical processing may help novice programmers learning to code.

Previous SE research has asserted that novices and experts differ most notably in the knowledge they possess [3, 6]. Moreover, researchers have hypothesized that expert programmers efficiently encode this knowledge into structures of syntax and semantics, domain information, and control flow, among others [6]. This allows expert programmers to recognize design patterns, form more insightful hypotheses about unfamiliar code, and more effectively find root causes of bugs [3, 4, 5]. In this study, we compared programmers based on their years of coding experience. We found evidence that more experienced programmers become more efficient in their cognition, where the functional connectivity between two brain areas decreases with years of coding experience. This finding may be representative of our study population, which consists of graduate and undergraduate CS students, and warrants future longitudinal studies. Nonetheless, we did not find significant differences in brain regions associated with knowledge. In Cognitive Neuroscience, researchers have used fMRI analyses such as searchlight or Representation Similarity Analysis (RSA) to analyze brain activity in finer detail. Prior research has identified voxel representations of the same concepts elicited by either words or pictures [57], suggesting these analysis techniques could be applied in SE. More concretely, future research could scan participants completing coding tasks or use publicly available fMRI datasets [20, 21, 22], and apply searchlight analysis or RSA to study potential knowledge differences between experts and novices.

In this study, we have shown that resting state scans are a powerful and convenient data source for researchers in SE to study the cognition of programming. We should note, however, that the effect sizes may be weaker in resting state data compared to those of task-based studies [23]. This can be seen in our null result in which there were no significant differences between novices, intermediates, and experts for the seed regions we chose. In conducting resting state studies, researchers can lower this hurdle by studying a large effect (i.e., perhaps programmers with a wider range of experience), increasing their statistical power with more participants, or approaching the data with more targeted hypotheses to reduce the impact of correcting for multiple comparisons. In this study, the differences between programmers and non-programmers were demonstrably large enough to expose meaningful changes in network activity as a result of coding. In this manner, we hope to provide guidance for researchers who wish to analyze this type of data and discover new knowledge about human and programming in the future.

VI. LIMITATIONS

In this section, we consider limitations of this study that may threaten the validity of our findings. These can be grouped into factors that may affect generalizability, and analysis decisions that may have influenced the results.

Generalizability First, brain activity was measured in an fMRI scanner, which is a highly restrictive experimental environment. As such, participants' brain activity might be unnaturally heightened or altered. Though all participants experienced the same experimental setup, which may control these effects, we acknowledge that anxious or claustrophobic participants may still exhibit heightened brain activity. Second, our participant sample may be relatively homogeneous, where most participants were college students. To mitigate this risk, our sample of non-programmers is diverse in its educational background, and may therefore offer a more robust comparison with programmers. Next, the programmers in our dataset may not represent a wide range of programming experience since most have between one and four years of coding experience, and some participants for whom we did not have educational information may have had programming experience. Furthermore, our expert group may not be representative of professional developers. Despite these limitations, we uncovered meaningful and statistically significant changes using conservative thresholds resulting from even this short time spent coding. We still acknowledge there may be stronger long-term effects of coding activities with a more experienced sample, and present our results as an initial, exploratory analysis of programming experience using resting state fMRI scans. Lastly, we attempted to isolate programming activities as an independent variable, but differences between populations in our study may be due to other variables or components of the CS degree. We mitigated these risks by correcting for multiple comparisons to reduce the false discovery rate, and performing linear regression in Section IV-B to explicitly test whether our findings were due to age as a confounding variable.

Construct Validity Our final results may have been influenced by our design decisions, as well as the nature of fMRI data itself. First, we compared two datasets whose scanning parameters were different. Combining fMRI datasets can be fraught with challenges, and the test-retest reliability can be low [32]. We acknowledge the risks and strove to mitigate them by first ensuring the data was all collected from the same scanner, then performing the same preprocessing steps, which are both recommended as best practice in neuroscience for combining datasets [29, 58]. Furthermore, we spatially smoothed the fMRI data and used an ROI-to-ROI functional connectivity analysis, both of which can improve the signal to noise ratio [23]. Next, we compared the datasets only based on their downstream statistics related to functional connectivity [32, 59]. Regardless, we encourage and welcome others to replicate the analysis we have done in this study.

The chosen seed regions might affect the functional connectivity analysis [15]. We attempted to mitigate this risk by precisely choosing seed regions based on MNI coordinates and figures in previous studies [7, 19, 22]. Furthermore, for replicability we included both Schaefer parcel numbers and MNI coordinates with respect to our seed regions and results. Next, we corrected our results for multiple comparisons by following established methods in Cognitive Neuroscience [40] and using a more conservative threshold of $q < 0.01$ (instead of $q < 0.05$) in our analyses. Lastly, we hand-classified components as signal or noise, as described in Section III, but did not formally obtain rater agreement. Some components may have been misclassified, which we attempted to mitigate by agreeing upon and following established criteria in Cognitive Neuroscience. These criteria rely on objective characteristics about component frequency, power spectra, and location.

VII. CONCLUSION

We present the first analysis of resting state fMRI data in SE. Analyzing resting state data enabled the first neurological comparison between programmers and non-programmers, and the concatenation of multiple datasets. We leveraged this data to study the influence of programming activities on the brain's functional organization. To this end, we compared non-programmers with programmers, then looked within programmers to explore how these neurological changes are mediated by years of experience. We found that programmers expressed stronger functional links compared to non-programmers in regions associated with language processing, temporal attention, and mathematical processing. Meanwhile, non-programmers expressed higher levels of connectivity with regions associated with social cognition and emotional processing. We interpret these findings to suggest that programming activities may selectively enhance the connectivity from language processing regions to temporal attention and hierarchical processing areas. The fMRI scanning parameters were different between these two datasets, which we attempted to mitigate, but we welcome any efforts to replicate our findings. Within our sample of programmers, we found a significant decrease in functional connectivity mediated by expertise between two regions associated with visual processing during reading and articulation, respectively. We interpret these findings to suggest increased neural efficiency in more experienced programmers.

VIII. ACKNOWLEDGMENTS

This research was supported in part by NSF grant CCF-2211429 and NIH grant R01 AG050523. We thank the participants who contributed their time, and our research groups at Vanderbilt University and the University of Michigan.

References

[1] M. Sigelman, S. Bittle, W. Markow, and B. Francis, “The hybrid job economy: How new skills are rewriting the dna of the job market,” *Burning Glass Technologies*, 2019.

[2] M. Endres, M. Fansher, P. Shah, and W. Weimer, “To read or to rotate? comparing the effects of technical reading training and spatial skills training on novice programming ability,” in *ESEC/FSE*, 2021.

[3] L. Gugerty and G. Olson, “Debugging by skilled and novice programmers,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, 1986.

[4] T. D. LaToza, D. Garlan, J. D. Herbsleb, and B. A. Myers, “Program comprehension as fact finding,” in *Proceedings of the the 6th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering*, 2007.

[5] A. J. Ko, B. A. Myers, M. J. Coblenz, and H. H. Aung, “An exploratory study of how developers seek, relate, and collect relevant information during software maintenance tasks,” *IEEE Transactions on software engineering*, vol. 32, no. 12, pp. 971–987, 2006.

[6] C. Parnin, J. Siegmund, and N. Peitek, “On the nature of programmer expertise.” in *Ppig*, 2017, p. 16.

[7] Z. Karas, A. Jahn, W. Weimer, and Y. Huang, “Connecting the dots: rethinking the relationship between code and prose writing with functional connectivity,” in *ESEC/FSE*, 2021.

[8] B. Floyd, T. Santander, and W. Weimer, “Decoding the representation of code in the brain: An fmri study of code review and expertise,” in *ICSE*. IEEE, 2017.

[9] N. Peitek, S. Apel, C. Parnin, A. Brechmann, and J. Siegmund, “Program comprehension and code complexity metrics: An fmri study,” in *ICSE*. IEEE, 2021.

[10] J. Castelhano, I. C. Duarte, C. Ferreira, J. Duraes, H. Madeira, and M. Castelo-Branco, “The role of the insula in intuitive expert bug detection in computer code: an fmri study,” *Brain imaging and behavior*, 2019.

[11] M. H. Lee, C. D. Smyser, and J. S. Shimony, “Resting-state fmri: a review of methods and clinical applications,” *American Journal of neuroradiology*, no. 10, 2013.

[12] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium *et al.*, “The wu-minn human connectome project: an overview,” *Neuroimage*, vol. 80, pp. 62–79, 2013.

[13] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, “A default mode of brain function,” *Proceedings of the national academy of sciences*, vol. 98, no. 2, pp. 676–682, 2001.

[14] M. Shim, H.-J. Hwang, U. Kuhl, and H.-A. Jeon, “Resting-state functional connectivity in mathematical expertise,” *Brain Sciences*, vol. 11, no. 4, p. 430, 2021.

[15] D. M. Cole, S. M. Smith, and C. F. Beckmann, “Advances and pitfalls in the analysis and interpretation of resting-state fmri data,” *Frontiers in systems neuroscience*, 2010.

[16] A. M. Zamorano, I. Cifre, P. Montoya, I. Riquelme, and B. Kleber, “Insula-based networks in professional musicians: evidence for increased functional connectivity during resting state fmri,” *Human brain mapping*, 2017.

[17] M. Costandi, *Neuroplasticity*. MIt Press, 2016.

[18] P. Cantou, H. Platel, B. Desgranges, and M. Groussard, “How motor, cognitive and musical expertise shapes the brain: Focus on fmri and eeg resting-state functional connectivity,” *Journal of Chemical Neuroanatomy*, 2018.

[19] J. Siegmund, C. Kästner, S. Apel, C. Parnin, A. Bethmann, T. Leich, G. Saake, and A. Brechmann, “Understanding understanding source code with functional magnetic resonance imaging,” in *ICSE*, 2014.

[20] R. Krueger, Y. Huang, X. Liu, T. Santander, W. Weimer, and K. Leach, “Neurological divide: an fmri study of prose and code writing,” in *ICSE*, 2020.

[21] Y. Huang, K. Leach, Z. Sharafi, N. McKay, T. Santander, and W. Weimer, “Biases and differences in code review using medical imaging and eye-tracking: genders, humans, and machines,” in *ESEC/FSE*, 2020.

[22] Y. Huang, X. Liu, R. Krueger, T. Santander, X. Hu, K. Leach, and W. Weimer, “Distilling neural representations of data structure manipulation using fmri and fnirs,” in *ICSE*. IEEE, 2019.

[23] Z. Chen and V. Calhoun, “Effect of spatial smoothing on task fmri ica and functional connectivity,” *Frontiers in neuroscience*, vol. 12, p. 327739, 2018.

[24] G. H. Glover, T.-Q. Li, and D. Ress, “Image-based method for retrospective correction of physiological motion effects in fmri: Retroicor,” *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 2000.

[25] L. Bonzano, M. Bortoletto, A. Zazio, C. Iester, A. Stango, R. Gasparotti, and M. Bove, “The hand motor hotspot for seed-based functional connectivity of hand motor networks at rest,” *Frontiers in Neuroscience*, 2022.

[26] J. Wang, M. Lu, Y. Fan, X. Wen, R. Zhang, B. Wang, Q. Ma, Z. Song, Y. He, J. Wang *et al.*, “Exploring brain functional plasticity in world class gymnasts: a network analysis,” *Brain Structure and Function*, 2016.

[27] D. A. Raichlen, P. K. Bharadwaj, M. C. Fitzhugh, K. A. Haws, G.-A. Torre, T. P. Trouard, and G. E. Alexander, “Differences in resting state functional connectivity between young adult endurance athletes and healthy controls,” *Frontiers in human neuroscience*, vol. 10, 2016.

[28] J. H. Jang, W. H. Jung, D.-H. Kang, M. S. Byun, S. J. Kwon, C.-H. Choi, and J. S. Kwon, “Increased default mode network connectivity associated with meditation,” *Neuroscience letters*, vol. 487, no. 3, pp. 358–362, 2011.

[29] N. K. Aurich, J. O. Alves Filho, A. M. Marques da Silva, and A. R. Franco, “Evaluating the reliability of different preprocessing steps to estimate graph theoretical measures in resting state fmri data,” *Frontiers in neuroscience*, vol. 9, p. 103902, 2015.

[30] P. K. Mandal, R. Mahajan, and I. D. Dinov, “Structural

brain atlases: design, rationale, and applications in normal and pathological cohorts," *Journal of Alzheimer's Disease*, vol. 31, no. s3, pp. S169–S188, 2012.

[31] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. Yeo, "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral cortex*, 2018.

[32] W. R. Shirer, H. Jiang, C. M. Price, B. Ng, and M. D. Greicius, "Optimization of rs-fmri pre-processing for enhanced signal-noise separation, test-retest reliability, and group discrimination," *Neuroimage*, 2015.

[33] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011.

[34] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ants similarity metric performance in brain image registration," *Neuroimage*, vol. 54, no. 3, 2011.

[35] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "Fsl," *Neuroimage*, 2012.

[36] O. Tange, *GNU parallel 2018*. Lulu. com, 2018.

[37] L. Griffanti, G. Douaud, J. Bijsterbosch, S. Evangelisti, F. Alfaro-Almagro, M. F. Glasser, E. P. Duff, S. Fitzgibbon, R. Westphal, D. Carone *et al.*, "Hand classification of fmri ica noise components," *Neuroimage*, 2017.

[38] J. Shum, D. Hermes, B. L. Foster, M. Dastjerdi, V. Rangarajan, J. Winawer, K. J. Miller, and J. Parvizi, "A brain area for visual numerals," *Journal of Neuroscience*, vol. 33, no. 16, pp. 6709–6715, 2013.

[39] C. M. Bennett and M. B. Miller, "fmri reliability: Influences of task and experimental design," *Cognitive, Affective, & Behavioral Neuroscience*, 2013.

[40] J. Durnez, S. P. Roels, and B. Moerkerke, "Multiple testing in fmri: an empirical case study on the balance between sensitivity, specificity, and stability," *Biometrical Journal*, vol. 56, no. 4, pp. 649–661, 2014.

[41] L. Chen, D. Wassermann, D. A. Abrams, J. Kochalka, G. Gallardo-Diez, and V. Menon, "The visual word form area (vwfa) is part of both language and attention circuitry," *Nature communications*, vol. 10, no. 1, p. 5601, 2019.

[42] D. Cheng, M. Li, J. Cui, L. Wang, N. Wang, L. Ouyang, X. Wang, X. Bai, and X. Zhou, "Algebra dissociates from arithmetic in the brain semantic network," *Behavioral and Brain Functions*, vol. 18, no. 1, p. 1, 2022.

[43] S. B. Eickhoff, A. R. Laird, P. T. Fox, D. Bzdok, and L. Hensel, "Functional segregation of the human dorsomedial prefrontal cortex," *Cerebral cortex*, 2016.

[44] K. Yaoi, N. Osaka, and M. Osaka, "Is the self special in the dorsomedial prefrontal cortex? an fmri study," *Social neuroscience*, vol. 4, no. 5, pp. 455–463, 2009.

[45] M. Capizzi, M. Martín-Signes, J. T. Coull, A. B. Chica, and P. Charras, "A transcranial magnetic stimulation study on the role of the left intraparietal sulcus in temporal orienting of attention," *Neuropsychologia*, 2023.

[46] T. Costa, F. Cauda, M. Crini, M.-K. Tatu, A. Celeghin, B. de Gelder, and M. Tamietto, "Temporal and spatial neural dynamics in the perception of basic emotions from complex scenes," *Social cognitive and affective neuroscience*, vol. 9, no. 11, pp. 1690–1703, 2014.

[47] I. Dumontheil, O. Küster, I. A. Apperly, and S.-J. Blakemore, "Taking perspective into account in a communicative task," *NeuroImage*, vol. 52, no. 4, 2010.

[48] J. C. Motzkin, C. L. Philippi, R. C. Wolf, M. K. Baskaya, and M. Koenigs, "Ventromedial prefrontal cortex is critical for the regulation of amygdala activity in humans," *Biological psychiatry*, vol. 77, no. 3, pp. 276–284, 2015.

[49] I. L. Nieuwenhuis and A. Takashima, "The role of the ventromedial prefrontal cortex in memory consolidation," *Behavioural brain research*, 2011.

[50] H. W. Koenigsberg, J. Fan, K. N. Ochsner, X. Liu, K. Guise, S. Pizzarello, C. Dorantes, L. Tecuta, S. Guerreri, M. Goodman *et al.*, "Neural correlates of using distancing to regulate emotional responses to social situations," *Neuropsychologia*, 2010.

[51] X.-Z. Jia, N. Zhao, B. Barton, R. Burciu, N. Carrière, A. Cerasa, B.-Y. Chen, J. Chen, S. Coombes, L. Defebvre *et al.*, "Small effect size leads to reproducibility failure in resting-state fmri studies," *BioRxiv*, p. 285171, 2018.

[52] A. Mechelli, G. W. Humphreys, K. Mayall, A. Olson, and C. J. Price, "Differential effects of word length and visual contrast in the fusiform and lingual gyri during," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 2000.

[53] L. Gates and M. G. Yoon, "Distinct and shared cortical regions of the human brain activated by pictorial depictions versus verbal descriptions: an fmri study," *Neuroimage*, vol. 24, no. 2, pp. 473–486, 2005.

[54] J. V. Baldo, D. P. Wilkins, J. Ogar, S. Willock, and N. F. Dronkers, "Role of the precentral gyrus of the insula in complex articulation," *Cortex*, 2011.

[55] M. Tettamanti and D. Weniger, "Broca's area: a supramodal hierarchical processor?" *Cortex*, 2006.

[56] J. Maloney, M. Resnick, N. Rusk, B. Silverman, and E. Eastmond, "The scratch programming language and environment," *ACM Transactions on Computing Education (TOCE)*, vol. 10, no. 4, pp. 1–15, 2010.

[57] B. J. Devereux, A. Clarke, A. Marouchos, and L. K. Tyler, "Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects," *Journal of Neuroscience*, vol. 33, no. 48, pp. 18 906–18 916, 2013.

[58] X. Li, N. B. Esper, L. Ai, S. Giavasis, H. Jin, E. Feczko, T. Xu, J. Clucas, A. Franco, A. S. Heinsfeld *et al.*, "Moving beyond processing and analysis-related variation in neuroscience," *BioRxiv*, pp. 2021–12, 2021.

[59] A. Weissenbacher, C. Kasess, F. Gerstl, R. Lanzenberger, E. Moser, and C. Windischberger, "Correlations and anticorrelations in resting-state functional connectivity mri: a quantitative comparison of preprocessing strategies," *Neuroimage*, vol. 47, no. 4, pp. 1408–1416, 2009.