Statement of Purpose

Zachary Karas

June 6, 2023

If I did not draw, I would not be as good of a researcher. I make ink drawings of abstract shapes that push on one another, intersect, and stretch the white of the page (Fig. 1). The process of making a drawing can be very rewarding, like when I take a step back from a piece and see how the disparate parts come together, but it can also be overwhelming. I pressure myself to make the final piece perfect, but I have only a vague idea what it will look like. My experience with research has been similar, where each component of a study needs deliberate planning and proper execution, but it is still difficult to know the exact outcome. This uncertainty can be stressful, but making art for years has helped me become more persistent, adaptable, and self-assured as a researcher. In my two peer-reviewed publications, I have analyzed and interpreted neuroimaging data to explore the cognitive basis of programming and software engineering. I intend to draw upon these skills, pun intended, as a PhD student studying Computer Science at Vanderbilt University, to improve programmers' efficiency and make coding more accessible.

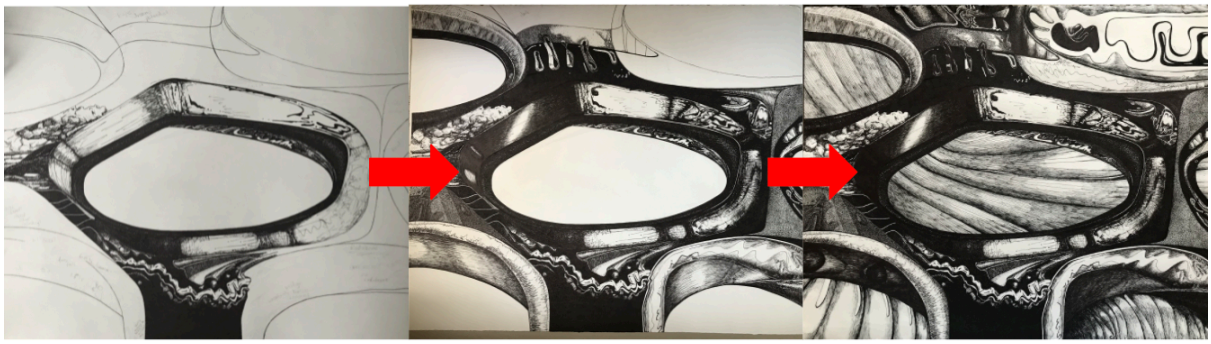

Figure 1: The progression of one my pieces *(title omitted)*

When I work on one particular area of a drawing, it usually veers away from my original intent, but persisting through these diversions gives a richness and character for which I could not have planned. In the moment, these intermediate steps can feel unsettling and chaotic, but now I see these stages as a sign of progress. Working through these complications directly applied to one of my research projects, but before discussing that, I will quickly summarize my research. I use neuroimaging to explore the cognitive basis of programming. Coding is a unique skill that uses language and math to manipulate data, but how do humans do it? Is coding more like writing, or more like a 3-D puzzle? I set out to answer these questions with my collaborators, and in my first project, we compared the neural activity of coding with that of mental rotation, and in my second, we compared code writing with prose writing. In my first project, I was responsible for neuroimaging data collection and analysis, but when I began analyzing the data I had collected, I saw the subsequent brain plots were blank. There was no significant

activity in any region. The deadline for a distinguished conference was fast approaching and my collaborators were relying on me, but I did not feel worried.

I considered the parts of our model where the issue could have originated, and hypothesized it was related to the timing of the stimuli. I had seen a similar issue in a dataset I had analyzed for a previous project, and believed we were making the same mistake by analyzing brain activity data from participants as they were idle, instead of engaged in a task. To fix the problem, I wrote preprocessing code to reannotate the parts each stimulus as either task or rest, and controlled for the latter which was now being ignored. Throughout this process, I was emailing my collaborators to keep them informed, all the while preparing for the possibility that it might not work. After analyzing the data again, however, I was happy to present brain plots showing significant activity in regions we named in our hypotheses, even after correcting for multiple comparisons. These figures formed the basis of the paper's results, which was published shortly thereafter in the *International Conference on Software Engineering* (20% acceptance rate).

We encountered other obstacles during the project, including the pandemic, but in each case I was able formulate a solution with my collaborators and calmly adjust. In a sense, I felt lucky we were able to resolve the challenges without any serious consequences, and I thank my experience with art for this adaptability and composure in the face of uncertainty. When I draw, even though I use ink, I hardly ever make a sketch with pencil beforehand. I cannot achieve the same fluid lines and organic shapes if I am too careful, so if something unexpected happens, I embrace it and find a way to make it work with the rest of the piece. In the previous research project I mentioned, I was experienced enough to fix the issues myself, but that was not always the case. In my next project, I was analyzing a different neuroimaging dataset from a different neuroimaging modality, and we were using a new analysis technique.

Standard analyses measure brain activity compared to a resting state, but we were interested in reanalyzing a particular dataset to measure the *connectivity* between brain regions. Most complex tasks are accomplished through the coordination of multiple regions instead of a single region, and standard analyses only provide information about individual areas. At the start of the project, I was responsible for applying functional connectivity analysis to uncover networks involved in code writing and prose writing. After a few weeks, however, I had not made much progress and another deadline was approaching. I wanted to analyze the data myself, but I did not want to jeopardize the project. The circumstances required a different approach, so I put aside my ego and recommended we ask Dr. Andrew Jahn, a neuroimaging specialist who was unknown to the others, for help. I had been emailing Dr. Jahn with questions up to that point, and my collaborators agreed asking him for help was the right decision at the right time.

My role on the project shifted once Dr. Jahn agreed to come onboard. I became the first author, conducting a literature review and directing the analyses. In our results we found a direct link between Broca's Area of the language network and a math region called the Number Form Area. We also used GPA from participants' computer science courses as a covariate, and found that participants with a higher GPA had stronger connectivity between areas of visual processing and areas for mathematical working memory. I contextualized these findings, that language and math regions are interconnected during coding tasks, for a computer science audience in our peer-reviewed paper, which was accepted at the conference *Foundations of Software Engineering* (19% acceptance rate). At the start of this project, I was responsible for analyzing the data, but the paper would not have been published in time if I had not suggested a change. Even with a plan in place, sometimes the circumstances require a change in course. Following impulses while I draw has helped me recognize these inflection points in research.

Even though research strives to be completely objective, there are myriad instances during a project when subjectivity is unavoidable. In tasks ranging from literature review to meeting

with collaborators, from finding relevant sources to discussing the direction of a project, I can feel self-assured by relying on the judgement I have honed with drawing. With every piece I have made, there have been countless times when I have needed to decide for myself how to fill a swath of whitespace. There is no formula and no measurable 'better' or 'worse' at these junctures, so it becomes a question of intuition and trusting what feels right. After drawing for years, I can rely on myself to make a good decision in nebulous circumstances when the outcome is uncertain.

One such research setting where this helped was in a meeting with my collaborators on the project involving functional connectivity analysis. We were discussing additional analyses to run, and one collaborator suggested we check whether there were gendered differences in patterns of connectivity. I believe this was well-intentioned, but I worried any differences between men's and women's brains could be taken out of context and misused in hiring, academic settings, or even conversations between peers. Finding gender differences could be especially sensitive in computer science, since there are already gender disparities, and it would be hard to prove whether these were due to genetic or experiential factors. At the same time, I was the junior member, and this collaborator had examined similar questions in the past for a good purpose. Still, I voiced my discomfort diplomatically, because no matter the outcome, I did not want a potential disagreement to distract us. As a group, we agreed not conduct the analyses. In this situation, I believe we were right not to pursue this question since we did not have data to interpret the results. Just as in drawing, one can say a lot with the material that is left off the paper.

These research experiences relate to my goals as a graduate student and afterwards: I would like to help students become better coders, and introduce more students to programming. From a cognitive perspective, coding is a unique skill in that it bridges the domains of math, language, and mental rotation. It is also one of the most lucrative in today's global economy, and yet our understanding of its neural basis is still limited. Results from this research could help inform IDE design, perhaps demonstrating whether tools such as Jupyter notebooks are helpful or distracting, and could also inform pedagogy in training future generations of coders. Ultimately, my goal is to help students, and I am interested in applying and developing more precise neuroimaging analyses to examine the neural correlates of software engineering. I believe my interests align well with those of Professors Yu Huang, Catie Chang, and Kevin Leach, and given my qualities as an artist and researcher, I feel I would be a good fit in the Computer Science department. I would be humbled for the opportunity to pursue a doctoral degree at Vanderbilt University.