

Neurological Divide: An fMRI Study of Prose and Code Writing

Ryan Krueger¹, **Yu Huang**¹, Xinyu Liu², Tyler Santander³,
Westley Weimer¹, Kevin Leach¹

¹University of Michigan

²Georgia Institute of Technology

³University of California, Santa Barbara

July 10, 2020



Thank You to the Collaborators!



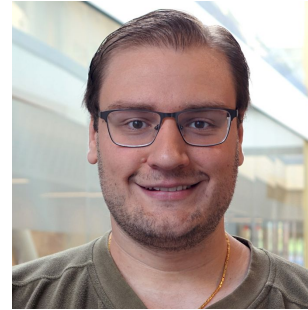
Ryan Krueger



Xinyu Liu



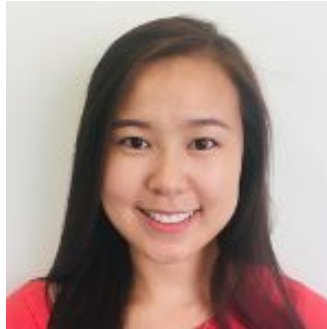
Dr. Tyler Santander



Dr. Kevin Leach



Dr. Westley Weimer



**Yu Huang is going on the
Job Market this year!**
yhhy@umich.edu

Motivation



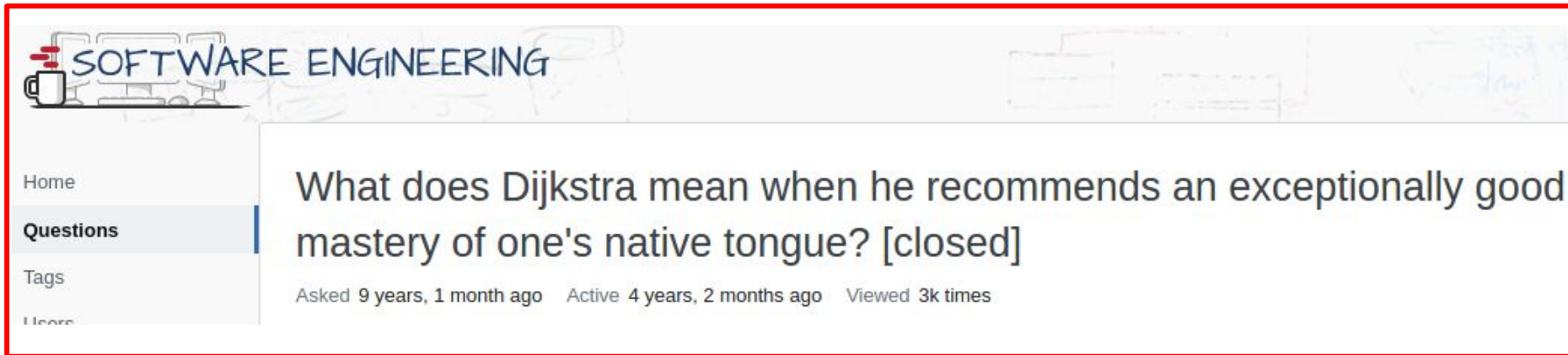
How do we tell truths that might hurt?

Besides a mathematical inclination, an exceptionally good mastery of one's native tongue is the most vital asset of a competent programmer.

Dijkstra might be right.

However, readers may take it in a different way and become really concerned...

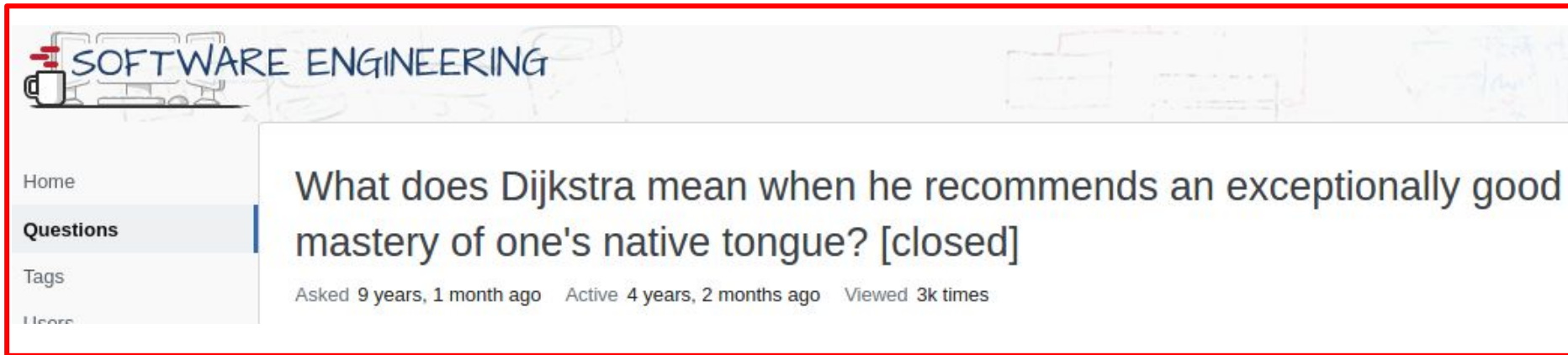
Motivation



The screenshot shows a Stack Overflow question page. The header includes a logo with a coffee cup and the text 'SOFTWARE ENGINEERING'. The question title is 'What does Dijkstra mean when he recommends an exceptionally good mastery of one's native tongue? [closed]'. Below the title, it says 'Asked 9 years, 1 month ago', 'Active 4 years, 2 months ago', and 'Viewed 3k times'. On the left, there is a navigation menu with 'Home', 'Questions', 'Tags', and 'Users'.

P.S. I have grown up in India. I speak [Bengali](#) at home; I speak [Marathi](#) in the community that I live in; [Hindi](#) is the national language and very widely spoken, so I know that, and in school and college I was taught with English as the first language. Of course, now I think in a multitude of languages and I must admit I don't have *mastery over any*. Is this really affecting my programming aptitude? If yes how ? and are there any *solutions*?

Motivation



The screenshot shows a Stack Overflow question page. The header includes the text 'SOFTWARE ENGINEERING' and a navigation menu with 'Home', 'Questions', 'Tags', and 'Users'. The question title is 'What does Dijkstra mean when he recommends an exceptionally good mastery of one's native tongue? [closed]'. Below the title, it says 'Asked 9 years, 1 month ago', 'Active 4 years, 2 months ago', and 'Viewed 3k times'.

Suggested Answer:

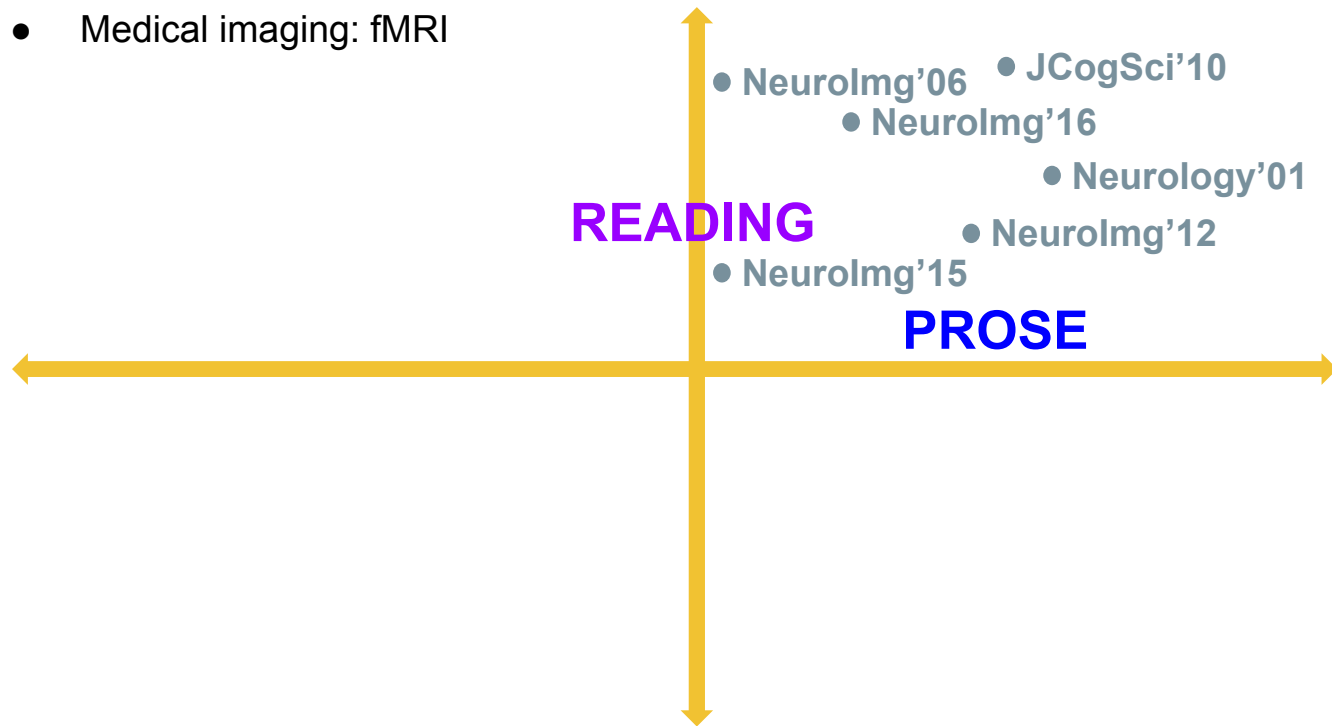
I believe this means that there is a direct correlation between a person's ability to learn a human language and a computer language. Both need the same set of human abilities and thinking capability. Take a look among your colleagues, and you will find that those with poor programming skills are also the ones who can't speak or write as clearly as others. Those who are good at picking human languages have the skills necessary to become good programmers too.

Motivation

- Objectively understanding the subjective cognitive process
 - Medical imaging: fMRI

Motivation

- Objectively understanding the subjective cognitive process
 - Medical imaging: fMRI



Motivation

- Objectively understanding the subjective cognitive process

- Medical imaging: fMRI

- FSE'17 • BrainImg'18

- ICSE'14 • ICPC'18

- ISSRE'16 • ICSE'17

READING

- NeuroImg'06 • JCogSci'10

- NeuroImg'16

- Neurology'01

- NeuroImg'12

- NeuroImg'15

CODE

PROSE

Motivation

- Objectively understanding the subjective cognitive process

- Medical imaging: fMRI

• FSE'17 • BrainImg'18

• ICSE'14 • ICPC'18

• ISSRE'16 • ICSE'17

READING

PROSE

WRITING

• NeuroImg'06 • JCogSci'10

• NeuroImg'16

• Neurology'01

• NeuroImg'12

• NeuroImg'15

• LangLearn'1989

• JWR'2008

• TESOL'1992

• NeuroImg'2006

• WrComm'2000

• HBM'2013

• CogBrR'2001

CODE

Motivation

- Objectively understanding the subjective cognitive process

- Medical imaging: fMRI

• FSE'17 • BrainImg'18

• ICSE'14 • ICPC'18

• ISSRE'16 • ICSE'17

CODE

READING

PROSE

• NeuroImg'06 • JCogSci'10

• NeuroImg'16

• Neurology'01

• NeuroImg'12

• NeuroImg'15

• LangLearn'1989

• JWR'2008

• TESOL'1992

• NeuroImg'2006

• WrComm'2000

• HBM'2013

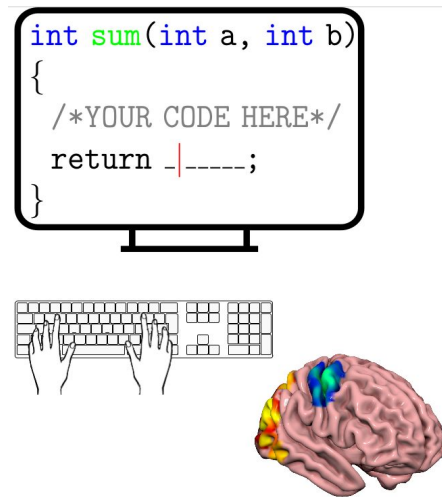
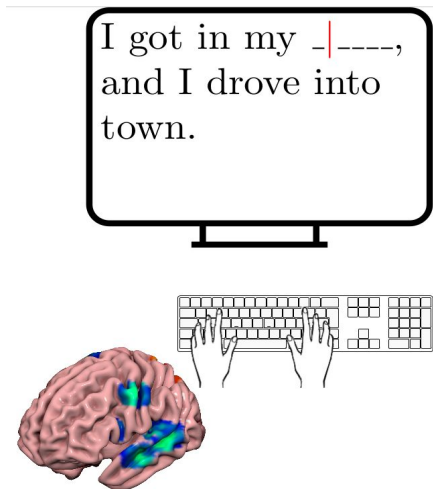
• CogBrR'2001



WRITING

High-level Question

- Are code writing and prose writing **similar** neural activities? Is being **good at writing** associated with being **a good software developer**?



Outline

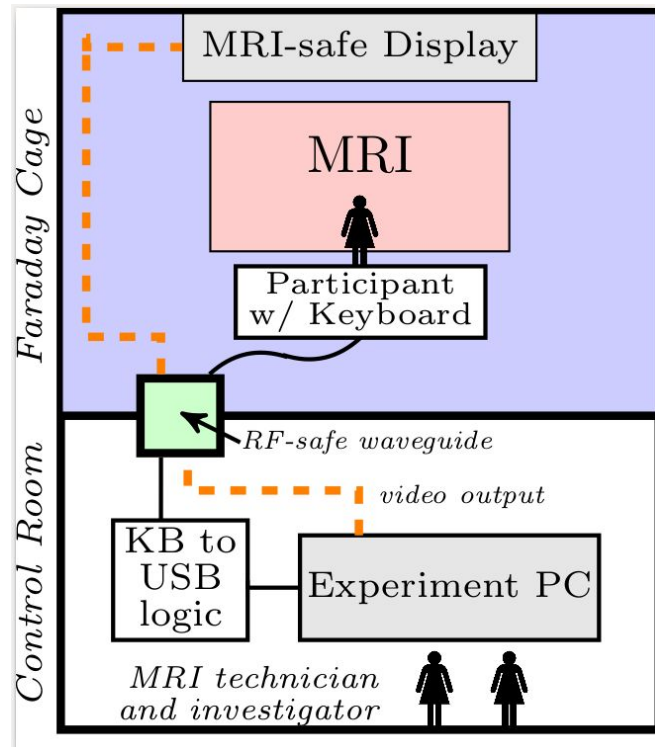
- Motivation
- High-level question
- **Challenges**
- **Experimental design**
- **Results**
- **Conclusions**

Challenges

- Physics
 - Magnetic interference

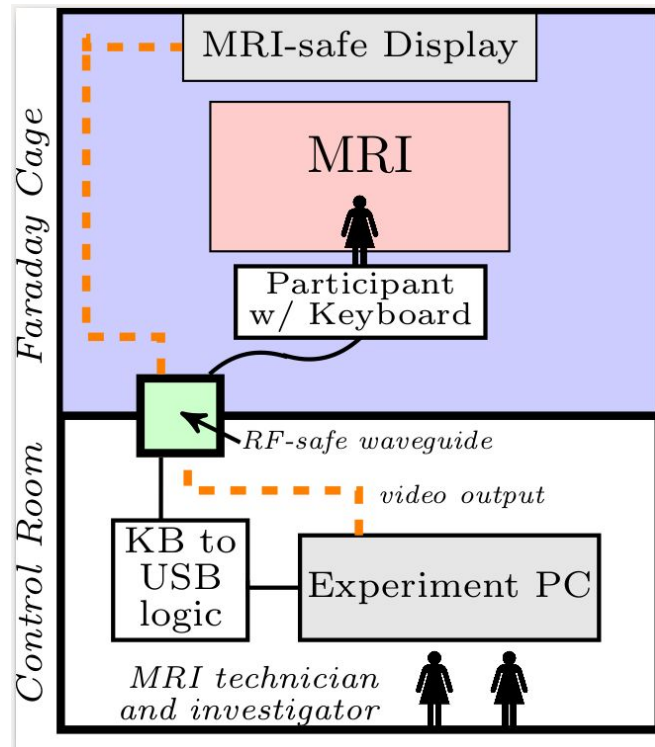
Challenges

- Physics
 - Magnetic interference
 - Solution
 - Employ an fMRI-safe bespoke keyboard



Challenges

- Physics
 - Magnetic interference
 - Solution
 - Employ an fMRI-safe bespoke keyboard



Challenges

- Physics
- Design
 - Contrast setup
 - Solution:
 - Two-by-two contrast task design

Experimental Design

- **Two-by-two contrast task design**
 - **Code writing** vs. **Prose writing**
 - **Fill in the blank (FITB)** vs. **Long response (LR)**

Experimental Design

- Two-by-two contrast task design

Given two 3x5 2D arrays of integer, x1 and x2, write the code needed to copy every value from x1 to its corresponding element in x2.

```
1 for(int i = 0; i < 3; i++){
2     for(int j = 0; j < 5; j++){
3         /* YOUR CODE HERE */
4     }
5 }
6 }
7
```

FITB

Fill in the blank below

```
1 Angered that the book arrived in the mail in
  such a shabby condition, Elliot insisted that
  the bookseller _____ it with a new copy.
```

CODE

PROSE

Implement a function is_sorted that accepts a vector of integer values and returns true if it is non-decreasing, and false otherwise

```
1 |
```

LR

What would happen if everyone lived in space? (e.g., What type of houses would they live in? What type of clothing would they wear?)

```
1 |
```

Experimental Design

- Two-by-two contrast task design

Given two 3x5 2D arrays of integer, x1 and x2, write the code needed to copy every value from x1 to its corresponding element in x2.

```
1 for(int i = 0; i < 3; i++){
2     for(int j = 0; j < 5; j++){
3         /* YOUR CODE HERE */
4     }
5 }
6 }
7 }
```

FITB

Fill in the blank below

```
1 Angered that the book arrived in the mail in
  such a shabby condition, Elliot insisted that
  the bookseller _____ it with a new copy.
```

Low-level

CODE

PROSE



Implement a function is_sorted that accepts a vector of integer values and returns true if it is non-decreasing, and false otherwise

```
1 |
```

LR

What would happen if everyone lived in space? (e.g., What type of houses would they live in? What type of clothing would they wear?)

```
1 |
```

Experimental Design

- Two-by-two contrast task design

Given two 3x5 2D arrays of integer, x1 and x2, write the code needed to copy every value from x1 to its corresponding element in x2.

```
1 for(int i = 0; i < 3; i++){
2     for(int j = 0; j < 5; j++){
3         /* YOUR CODE HERE */
4     }
5 }
6 }
7
```

FITB

Fill in the blank below

```
1 Angered that the book arrived in the mail in
  such a shabby condition, Elliot insisted that
  the bookseller _____ it with a new copy.
```

Low-level

CODE

PROSE

Implement a function is_sorted that accepts a vector of integer values and returns true if it is non-decreasing, and false otherwise

```
1 |
```

LR

What would happen if everyone lived in space? (e.g., What type of houses would they live in? What type of clothing would they wear?)

```
1 |
```

High-level

Experimental Design

- Two-by-two contrast task design

Given two 3x5 2D arrays of integer, x1 and x2, write the code needed to copy every value from x1 to its corresponding element in x2.

```
1 for(int i = 0; i < 3; i++){
2     for(int j = 0; j < 5; j++){
3         /* YOUR CODE HERE */
4     }
5 }
6 }
7
```

FITB

Fill in the blank below

```
1 Angered that the book arrived in the mail in
  such a shabby condition, Elliot insisted that
  the bookseller _____ it with a new copy.
```

Low-level

CODE

PROSE

Implement a function is_sorted that accepts a vector of integer values and returns true if it is non-decreasing, and false otherwise

```
1 |
```

LR

What would happen if everyone lived in space? (e.g., What type of houses would they live in? What type of clothing would they wear?)

```
1 |
```

High-level

Experimental Design

- **Two-by-two contrast task design**
 - **Code writing** vs. **Prose writing**
 - **Fill in the blank (FITB)** vs. **Long response (LR)**
- **Source**
 - **Code**: Turing's Craft



The screenshot shows the Turing's Craft website. At the top left is the logo "turingscraft" with a green cube icon. To the right are links for "register" and "login". Below the logo is a navigation bar with links: "Our Innovations", "The Benefits", "CodeLabs", "About Us", and "Demo". The main content area features the heading "CodeLab™: A Powerful Tool for Programming Instruction". Below this heading is a paragraph of text describing CodeLab as a web-based interactive programming exercise system. To the right of the text are two callout boxes: "Try a Demo" with a link to register a free account for instructors, and "What Instructors say" with a quote from a professor at the University of Toronto praising the system.

turingscraft register | login

Our Innovations / The Benefits / CodeLabs / About Us / Demo

CodeLab™: A Powerful Tool for Programming Instruction

CodeLab is the web-based interactive programming exercise system for intro programming classes in Python, Java, C++, C, JavaScript, C#, VB and SQL. First offered in 2002 to reduce attrition and raise the overall level of the class, it is a seasoned system that has been used in over 400 institutions in 20 countries and analyzed over 135,000,000 (one hundred thirty-five million) exercise submissions from more than 300,000 students.

A CodeLab has 200-800 short exercises, each focused on a particular programming idea or language construct. The student types in code and the system immediately judges its correctness, offering hints when the submission is incorrect. Through this process, the student gains mastery over the semantics, syntax and common usage of the language elements.

Try a Demo
Click [HERE](#) to Register Your Free CodeLab Account. (Instructors Only)

What Instructors say
"Our students and instructors credit the labs and the CodeLab for improved marks." [Jeremy Sills, Professor, University of Toronto](#)

Experimental Design

- **Two-by-two contrast task design**
 - **Code writing** vs. **Prose writing**
 - **Fill in the blank (FITB)** vs. **Long response (LR)**
- **Source**
 - **Code**: Turing's Craft
 - **Prose**: Scholastic Assessment Test (SAT)



The screenshot shows the Turing's Craft website. At the top left is the logo "turingcraft" with a green leaf icon. To the right are links for "register" and "login". Below the logo is a navigation bar with links: "Our Innovations", "The Benefits", "CodeLabs", "About Us", and "Demo". The main content area features the heading "CodeLab™: A Powerful Tool for Programming Instruction". Below this is a paragraph describing CodeLab as a web-based interactive programming exercise system for intro programming classes in Python, Java, C++, C, JavaScript, C#, VB and SQL. It mentions that it was first offered in 2002 to reduce attrition and raise the overall level of the class, and that it has been used in over 400 institutions in 20 countries, analyzed over 135,000,000 (one hundred thirty-five million) exercise submissions from more than 300,000 students. Below the paragraph is a section titled "Try a Demo" with a link: "Click HERE to Register Your Free CodeLab Account. (Instructors Only)". To the right of the paragraph is another section titled "What Instructors say" with a quote: "Our students and instructors credit the labs and the CodeLab for improved marks." followed by "Jeremy Sills, Professor, University of Toronto".




Results

- **Recruitment**
 - **30 participants**
 - 20 male vs. 10 female
 - 27 undergraduate vs. 3 graduate
- **Tasks**
 - Four randomized blocks
 - Code FITB: 17
 - Code LR: 9
 - Prose FITB: 17
 - Prose LR: 9

Results

- **Data analysis: we need to be careful**
 - Spurious correlation or false discovery from multiple comparisons



Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction

Craig M. Bennett¹, Abigail A. Baird², Michael B. Miller¹, and George L. Wolford³

¹ Psychology Department, University of California Santa Barbara, Santa Barbara, CA; ² Department of Psychology, Vassar College, Poughkeepsie, NY; ³ Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH

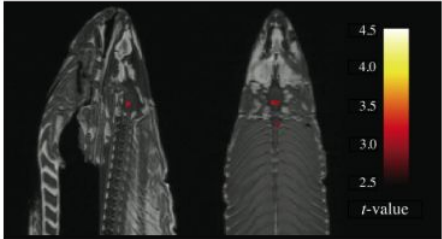
INTRODUCTION

With the extreme dimensionality of functional neuroimaging data comes extreme risk for false positives. Across the 130,000 voxels in a typical fMRI volume the probability of a false positive is almost certain. Correction for multiple comparisons should be completed with these datasets, but is often ignored by investigators. To illustrate the magnitude of the problem we carried out a real experiment that demonstrates the danger of not correcting for chance properly.

METHODS

Subject. One mature Atlantic Salmon (*Salmo salar*) participated in the fMRI study. The salmon was approximately 18 inches long, weighed 3.8 lbs, and was not alive at

GLM RESULTS



The figure shows two grayscale fMRI brain scans of a salmon. The left scan is a coronal view, and the right scan is an axial view. Both scans show red clusters indicating significant activation. A color scale on the right indicates t-values from 2.5 to 4.5.

Results

- **Data analysis: we need to be careful**
 - Spurious correlation or false discovery from multiple comparisons
 - Three steps

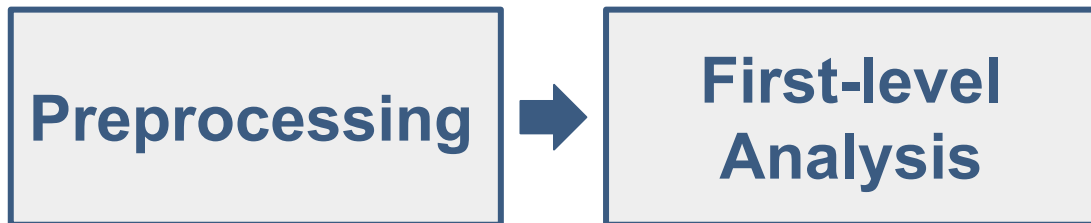
Results

- **Data analysis: we need to be careful**
 - Spurious correlation or false discovery from multiple comparisons
 - Three steps

Preprocessing

Results

- **Data analysis: we need to be careful**
 - Spurious correlation or false discovery from multiple comparisons
 - Three steps



Results

- **Data analysis: we need to be careful**
 - Spurious correlation or false discovery from multiple comparisons
 - Three steps



Results

- RQ1: Do **self reports** claim code writing is like prose writing?
- RQ2: Does the brain treat **code writing** like **prose writing**?
- RQ3: What **low-level** features explain code and prose writing?
- RQ4: What **high-level** features explain code and prose writing?

Results

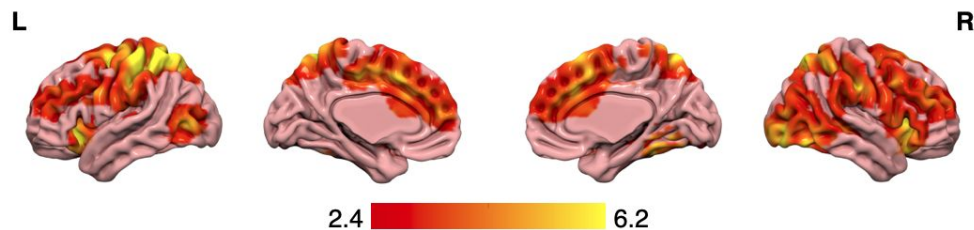
- RQ1: Do **self reports** claim code writing is like prose writing?
 - **38.5%** reported similarity between prose and code writing

Results

- RQ2: Does the brain treat **code writing** like **prose writing**?

Results

- RQ2: Does the brain treat **code writing** like **prose writing**?
 - **Significant** and **widely-distributed** difference in neural activity
 - More than 10 brain regions (Brodmann Areas)



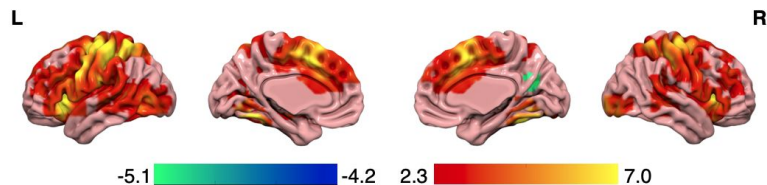
Code > **Prose**

Results

- RQ3: What **low-level** features explain code and prose writing?
- RQ4: What **high-level** features explain code and prose writing?

Results

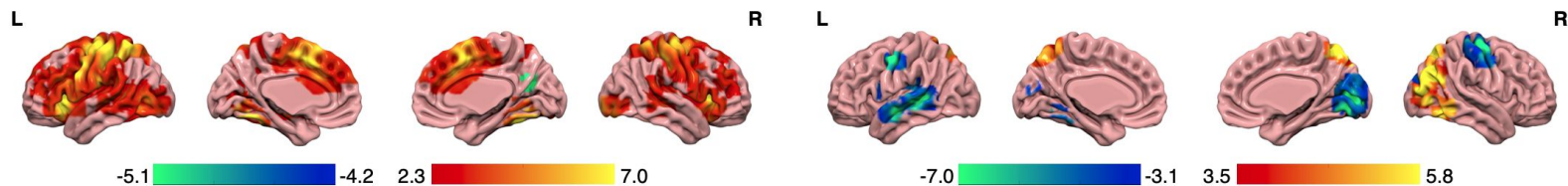
- RQ3: What **low-level** features explain code and prose writing?
 - **Low-level: code writing** requires more in parts of the brain associated with top-down control, planning, and categorization
- RQ4: What **high-level** features explain code and prose writing?



Code FITB > Prose FITB

Results

- RQ3: What **low-level** features explain code and prose writing?
 - **Low-level: code writing** requires more in parts of the brain associated with top-down control, planning, and categorization
- RQ4: What **high-level** features explain code and prose writing?
 - **High-level: prose writing** requires more in parts of the brain associated with language; **code writing** involves more in attention, memory, planing, and spatial ability.



Code FITB > Prose FITB

Code LR > Prose LR

Summary

- First fMRI study of code writing
 - Bespoke fMRI-safe QWERTY keyboard
 - Controlled, contrast-based experiment
- **Main result:** All analysis of all **code writing** tasks against **prose writing** tasks showed **distinct neural mechanisms**
- At a more granular level:
 - **Code FITB** > **Prose FITB**: top-down control, planning, categorization
 - **Code LR** > **Prose LR**: code involves more of the right hemisphere (spatial ability, planning)
prose involves more canonical left hemisphere (language production)
- Discussion
 - Pedagogy; Workforce retraining; Encouraging more diverse participation in computer science