Neurological Divide: An fMRI Study of Prose and Code Writing

Ryan Krueger\textsuperscript{1}, Yu Huang\textsuperscript{1}, Xinyu Liu\textsuperscript{2}, Tyler Santander\textsuperscript{3}, Westley Weimer\textsuperscript{1}, Kevin Leach\textsuperscript{1}

\textsuperscript{1}University of Michigan
\textsuperscript{2}Georgia Institute of Technology
\textsuperscript{3}University 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

What does Dijkstra mean when he recommends an exceptionally good mastery of one's native tongue? [closed]

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

Suggested Answer:

I believe this means that there is a direct corollation 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 neccessary 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

Yu Huang @ ICSE2020
Motivation

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

CODE
- FSE’17
- ICSE’14
- ISSRE’16
- ICSE’17

PROSE
- NeuroImg’06
- NeuroImg’15
- NeuroImg’16
- NeuroImg’12
- NeuroImg’15
- Neuroscience’01
- BrainImg’18
- ICPC’18
- ICSE’17
- JCogSci’10

Yu Huang @ ICSE2020
Motivation

- Objectively understanding the subjective cognitive process
  - Medical imaging: fMRI
    - FSE’17
    - BrainImg’18
    - ICSE’14
    - ICPC’18
    - ISSRE’16
    - ICSE’17

  - NeuroImg’06
  - NeuroImg’16
  - NeuroImg’15
  - Neurology’01
  - NeuroImg’12
  - NeuroImg’10

  - JCogSci’10
  - NeuroImg’16

  - NeuroImg’06
  - NeuroImg’12
  - NeuroImg’15
  - Neurology’01

  - WrComm’2000
  - TESOL’1992
  - HBM’2013
  - CogBrR’2001

- JWR’2008
- LangLearn’1989
- ICSE’14
- ISSRE’16
- ICSE’17

Yu Huang @ ICSE2020
Motivation

- Objectively understanding the subjective cognitive process
  - Medical imaging: fMRI
    - FSE’17
    - BrainImg’18
    - ICSE’14
    - ICPC’18
    - ISSRE’16
    - ICSE’17
    - NeuroImg’06
    - NeuroImg’16
    - Neurology’01
    - NeuroImg’12
    - NeuroImg’15
    - NeuroImg’16
    - NeuroImg’18
    - Neurology’01
    - NeuroImg’12
    - NeuroImg’15
    - NeuroImg’16
    - NeuroImg’18
  - Reading
    - JWR’2008
    - NeuroImg’06
    - TESOL’1992
    - WrComm’2000
    - HBM’2013
    - CogBrR’2001
  - Writing
    - LangLearn’1989
    - HBM’2013
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.

  ```java
  for(int i = 0; i < 3; i++){
    for(int j = 0; j < 5; j++){
      // YOUR CODE HERE
    }
  }
  ```

  Fill in the blank below:

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

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

  ```java
  ```

  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?)

  ```java
  ```
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.

```java
for(int i = 0; i < 3; i++){
    for(int j = 0; j < 5; j++){
        /* YOUR CODE HERE */
    }
}
```

Fill in the blank below

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

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?)
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.

```java
for(int i = 0; i < 3; i++) {
    for(int j = 0; j < 5; j++) {
        /* YOUR CODE HERE */
    }
}
```

Fill in the blank below

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

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

Low-level

CODE

FITB

High-level

PROSE

LR

Yu Huang @ ICSE2020
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.

```java
for(int i = 0; i < 3; i++) {
    for(int j = 0; j < 5; j++) {
        /* YOUR CODE HERE */
    }
}
```

Fill in the blank below

```java
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
- FITB

High-level

- PROSE
- LR

Yu Huang @ ICSE2020
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
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)
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
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

![Diagram](preprocessing→first-level-analysis)
Results

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

Preprocessing → First-level Analysis → Contrast & Group-level analysis
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 (Broadmann Areas)

![Brain images showing neural activity comparison between code and prose]

*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?**

![Brain images](image)

**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, planning, and spatial ability.

<table>
<thead>
<tr>
<th>Code FITB</th>
<th>Prose FITB</th>
<th>Code LR</th>
<th>Prose LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;</td>
<td></td>
<td>&gt;</td>
<td></td>
</tr>
</tbody>
</table>
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