Understanding User Cognition: From Spatial Ability to Code Writing and Review

Yu Huang
http://www-personal.umich.edu/~yhhuy
June 17, 2021
We want to **improve productivity** and **reduce cost** in software development and maintenance.
### The Human Aspect Matters

- Early study of industrial developers found order-of-magnitude individual variations

<table>
<thead>
<tr>
<th>Metric</th>
<th>Poorest</th>
<th>Best</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debugging Hours Algebra</td>
<td>170</td>
<td>6</td>
<td>28:1</td>
</tr>
<tr>
<td>Debugging Hours Maze</td>
<td>26</td>
<td>1</td>
<td>26:1</td>
</tr>
<tr>
<td>CPU Seconds Algebra</td>
<td>3075</td>
<td>370</td>
<td>8:1</td>
</tr>
<tr>
<td>CPU Seconds Maze</td>
<td>541</td>
<td>50</td>
<td>11:1</td>
</tr>
<tr>
<td>Code Writing Hours Algebra</td>
<td>111</td>
<td>7</td>
<td>16:1</td>
</tr>
<tr>
<td>Code Writing Hours Maze</td>
<td>50</td>
<td>2</td>
<td>25:1</td>
</tr>
<tr>
<td>Program Size Algebra</td>
<td>6137</td>
<td>1050</td>
<td>6:1</td>
</tr>
<tr>
<td>Program Size Maze</td>
<td>3287</td>
<td>651</td>
<td>5:1</td>
</tr>
<tr>
<td>Run Time Algebra</td>
<td>7.9</td>
<td>1.6</td>
<td>5:1</td>
</tr>
<tr>
<td>Run Time Maze</td>
<td>8.0</td>
<td>0.6</td>
<td>13:1</td>
</tr>
</tbody>
</table>

The Human Aspect Matters

- Early study of industrial developers found order-of-magnitude individual variations

<table>
<thead>
<tr>
<th>Metric</th>
<th>Poorest</th>
<th>Best</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debugging Hours Algebra</td>
<td>170</td>
<td>6</td>
<td>28:1</td>
</tr>
<tr>
<td>Debugging Hours Maze</td>
<td>26</td>
<td>1</td>
<td>26:1</td>
</tr>
<tr>
<td>CPU Seconds Algebra</td>
<td>3075</td>
<td>370</td>
<td>8:1</td>
</tr>
<tr>
<td>CPU Seconds Maze</td>
<td>541</td>
<td>50</td>
<td>11:1</td>
</tr>
<tr>
<td>Code Writing Hours Algebra</td>
<td>111</td>
<td>7</td>
<td>16:1</td>
</tr>
<tr>
<td>Code Writing Hours Maze</td>
<td>50</td>
<td>2</td>
<td>25:1</td>
</tr>
<tr>
<td>Program Size Algebra</td>
<td>6137</td>
<td>1050</td>
<td>6:1</td>
</tr>
<tr>
<td>Program Size Maze</td>
<td>3287</td>
<td>651</td>
<td>5:1</td>
</tr>
<tr>
<td>Run Time Algebra</td>
<td>7.9</td>
<td>1.6</td>
<td>5:1</td>
</tr>
<tr>
<td>Run Time Maze</td>
<td>8.0</td>
<td>0.6</td>
<td>13:1</td>
</tr>
</tbody>
</table>

The Human Aspect Matters

- How to *measure* cognitive processes?
  - Conduct behavioral experiments
  - “Stopwatch” and “Scoresheet”
    - *Time* and *accuracy*
The Human Aspect Matters

- How to *measure* cognitive processes?
  - Conduct behavioral experiments
  - “Stopwatch” and “Scoresheet”
    - Time and accuracy
The Human Aspect Matters

- How to **measure** cognitive processes?
  - Conduct behavioral experiments
  - "**Stopwatch**" and "**Scoresheet**"
    - **Time** and **accuracy**
    - **What** but not **why**
      - Generalization, recommendation, transformation
    - **Overlook** what is actually going on
      - Miss information
    - **Limited** research findings
The Human Aspect Matters

● How to *measure* cognitive processes?
  ● Conduct behavioral experiments
  ● “Stopwatch” and “Scoresheet”
  ● Self-reporting
The Human Aspect Matters

- How to **measure** cognitive processes?
  - Conduct behavioral experiments
  - “Stopwatch” and “Scoresheet”
  - **Self-reporting**
    - Unreliable
The Human Aspect Matters

- How to **measure** cognitive processes?
  - Conduct behavioral experiments
  - “Stopwatch” and “Scoresheet”
  - Self-reporting
    - Unreliable

“Can we read your mind?”
The Human Aspect Matters

- **How to measure** cognitive processes?
  - Conduct behavioral experiments
  - “Stopwatch” and “Scoresheet”
  - Self-reporting
    - **Unreliable**

“Can we read your mind?”

Medical Imaging & Eye Tracking
Thesis Statement

It is possible to meaningfully and objectively measure user cognition to understand the role of spatial ability, fundamental processes and stereotypical associations in certain software engineering activities by combining medical imaging and eye tracking.
Challenges

- Many techniques: EEG, PET, fMRI, fNIRS, …
  - Non-invasive
  - fMRI and fNIRS
    - Sampling the brain rapidly with high spatial resolution
Challenges

- Many techniques: EEG, PET, fMRI, fNIRS, …
  - Non-invasive
  - fMRI and fNIRS
    - Sampling the brain *rapidly* with *high spatial resolution*
    - > 37,000 publications collected in PubMed in 2020 but relatively new to CS: less than two dozens in SE
Challenges

- Many techniques: EEG, PET, fMRI, fNIRS, …
  - Non-invasive
  - fMRI and fNIRS
    - Sampling the brain rapidly with high spatial resolution
    - > 1500 publications in Psychology but relatively new to CS: less than a dozen in SE
    - Streetlight effects?
Challenges

- Many techniques: EEG, PET, fMRI, fNIRS, …
  - Non-invasive
  - fMRI and fNIRS
    - Sampling the brain rapidly with high spatial resolution
    - > 1500 publications in Psychology but relatively new to CS: less than a dozen in SE
    - Streetlight effects?
      - Environment
Challenges

- Many techniques: EEG, PET, fMRI, fNIRS, ...
  - Non-invasive
  - fMRI and fNIRS
    - Sampling the brain rapidly with high spatial resolution
    - > 1500 publications in Psychology but relatively new to CS: less than a dozen in SE
    - Streetlight effects?
      - Environment
      - Contrast-based tasks
Challenges

● Many techniques: EEG, PET, fMRI, fNIRS, …
  ● Non-invasive
  ● fMRI and fNIRS
    ○ Sampling the brain *rapidly* with *high spatial resolution*
    ○ > 1500 publications in Psychology but relatively new to CS: less than a dozen in SE
    ○ Streetlight effects?
      ● Environment
      ● Contrast-based tasks
      ● Experimental design
Outline

● Introduction
● Investigating cognition in software engineering
  ● Understanding the neural representations of data structures
  ● Comparing prose writing and code writing
  ● Understanding bias in code reviews
● Career Plan
● Summary
Hypothesis Time!

Is balancing AVL trees like playing arcade claw machines?
Hypothesis Time!

Is balancing AVL trees like playing arcade claw machines?

Yes! Very similar!

No! Very Different!
How do human brains represent **data structures**? Is it more like **text** or more like **3D objects**?
How do human brains represent data structures? Is it more like text or more like 3D objects?
How do human brains represent data structures? Is it more like text or more like 3D objects?
Spatial Ability

● Mental rotations
  ● The determination of spatial relationships between objects and the mental manipulation of spatially presented information
  ● **Measured by mental rotation tasks**: 3D objects
  ● Related to success in STEM
About Medical Imaging

- **fMRI** and **fNIRS**
- **BOLD** signals
- Contrast design
- Rigorous data analysis: false positives
fMRI vs. fNIRS

Measure brain activities by calculating the blood-oxygen level dependent (BOLD) signal

- Functional Magnetic Resonance Imaging
- Functional Near-InfraRed Spectroscopy
fMRI vs. fNIRS

Measure brain activities by calculating the blood-oxygen level dependent (BOLD) signal

- **Functional Magnetic Resonance Imaging**
  - Magnets
  - Strong penetration power
  - Lying down in a magnetic tube

- **Functional Near-InfraRed Spectroscopy**
  - Light
  - Weak penetration power
  - Wearing a specially-designed cap

Neural Representations of Data Structures (ICSE’19)
fMRI vs. fNIRS

Measure brain activities by calculating the blood-oxygen level dependent (BOLD) signal

- **Functional Magnetic Resonance Imaging**
  - **Magnets**
  - **Strong** penetration power
  - Lying down in a magnetic tube:
    - Cannot move

- **Functional Near-InfraRed Spectroscopy**
  - **Light**
  - **Weak** penetration power
  - Wearing a specially-designed cap:
    - More freedom of movement

Neural Representations of Data Structures (ICSE’19)
What is **BOLD** signal?

- **Blood-Oxygen Level Dependent (BOLD)** signal
- Blood flow and oxygen consumption as a **proxy** for brain activity
What is **BOLD** signal?

- Blood-Oxygen Level Dependent (**BOLD**) signal
- Blood flow and oxygen consumption as a **proxy** for brain activity
- Activation model: hemodynamic response function (**HRF**)
What is **BOLD** signal?

- **Blood-Oxygen Level Dependent (BOLD)** signal
- Blood flow and oxygen consumption as a **proxy** for brain activity
- Activation model: hemodynamic response function (HRF)
- Stimulus, HRF, design matrix, noise
  - Comprehensive quantitative model of BOLD signals
    - General Linear Model (GLM)
But it’s not so easy

- Brain activation does not work like this:
But it’s not so easy

- The brain signals are noisy
But it’s not so easy

- The brain signals are noisy
- Signal changes are small
Think in Terms of Contrasts!
Think in Terms of Contrasts!

- Controlled experimental design
  - Task A = “balancing trees + nervous + …”
  - Task B = “rotating 3D objects + nervous + …”
  - Contrast $A > B$: brain activations that vary between the tasks
Data Analysis

1. Preprocessing
2. First-level Analysis
3. Contrast & Group-level Analysis
Data Analysis

Preprocessing

- Slice-time Correction
- Voxel Displacement Maps (VDMs), Motion Correction, Co-registration & Normalization (MNI152)
- Spatial Smoothing
Data Analysis

Preprocessing

Slice-time Correction

Voxel Displacement Maps (VDMs), Motion Correction, Co-registration & Normalization (MNI152)

Spatial Smoothing
Data Analysis

Preprocessing

First-level Analysis

- **Within** individuals
- **General Linear Models**
  - Robust weighted least squares (rWLS)
  - For each experiment condition
Data Analysis

Preprocessing → First-level Analysis → Contrast & Group-level Analysis

- Pairwise contrast
  - Mean differences between conditions
  - Group-level random effects analyses
- Second-level GLM
  - Assess average activity across all subjects
Data Analysis

- Preprocessing

  First-level Analysis

  Contrast & Group-level Analysis

- Final results
  - A statistical parametric map of t-values (t-map) describing clusters of significant activity for a given task-related comparison
Data Analysis

- We need to be *careful*
  - 153,000 voxels or more
  - Spurious correlations due to multiple comparison: false positives
Data Analysis

- We need to be **careful**
  - 153,000 voxels or more
  - Spurious correlations due to multiple comparison: false positives
Data Analysis

- Preprocessing
- First-level Analysis
- Contrast & Group-level Analysis

- False discovery rate (FDR) correction (q<0.05)
Experiment Design

- Two types of tasks
  - Data structure manipulations
  - Mental rotations: 3D objects

**Experimental Design**

**Two types of tasks**

- Data structure manipulations
- Mental rotations: 3D objects

**Which object is the same as the original object, aside from its orientation?**

A. A
B. B

**What is the minimum number of swaps required to make the given array sorted?**

<table>
<thead>
<tr>
<th>Indices</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>nums</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

A. 1  B. 2

**Which of the candidate insertion sequences will produce the given BST?**

A. 5, 3, 8  B. 8, 3, 5
Experimental Design

- Two types of tasks
  - Data structure manipulations
  - Mental rotations: 3D objects
- fMRI and fNIRS: 1st time in SE
Experimental Design

- Two types of tasks
  - Data structure manipulations
  - Mental rotations: 3D objects
- fMRI and fNIRS: 1\textsuperscript{st} time in SE
- Largest in SE: 76 participants
Data Structures vs. Spatial Ability

- 70% of human participants believe there is no connection
- What is your answer?

Hypothesis Time!

Is balancing AVL trees like playing arcade claw machines?

- Yes! Very similar!
- No! Very Different!
Data Structures vs. Spatial Ability

- Data structure manipulations use the same parts of the brain as rotating 3D objects in the real world (spatial ability)
  - fMRI: more similarities than differences ($p<0.01$)
  - fNIRS: activation in the same brain regions ($p<0.01$)

Mental Rotation > Tree
The Role of Task Difficulty

- The brain works even **harder** for **more difficult** data structure tasks
  - Difficulty measurement
    - Mental rotations: angle of rotation
    - Data structure: size
  - fNIRS: no significant findings for the effect of task difficulty
Summary: Data Structures vs. Spatial Ability

- **Large human study**: 76 participants
- **fMRI vs. fNIRS**
- Data structure manipulations and mental rotations use the **same** brain regions
- **Task difficulty matters** for data structures
- Medical imaging can discover more than self-reporting
Summary: Data Structures vs. Spatial Ability

● This work may inform:
  ● Pedagogy and training
  ● Technology transfer
  ● Programming expertise
Summary: Data Structures vs. Spatial Ability

● This work may inform:
  ● Pedagogy and training
  ● Technology transfer
  ● Programming expertise

● The findings have been used to direct a longitudinal study
  ● Improve CS students’ performance on programming using spatial training
Summary: Data Structures vs. Spatial Ability

- This work may inform:
  - Pedagogy and training
  - Technology transfer
  - Programming expertise

Distilling Neural Representations of Data Structure Manipulation using fMRI and fNIRS

Yu Huang¹, Xinyu Liu¹, Ryan Krueger¹, Tyler Santander², Xiaosu Hu¹, Kevin Leach¹ and Westley Weimer¹

¹{yhyh, xinyuliu, ryankrue, xiaosuhu, kjleach, weimerw}@umich.edu, University of Michigan
²t.santander@psych.ucsb.edu, University of California, Santa Barbara
Outline

● Introduction
● Investigating cognition in software engineering
   - Understanding the neural representations of data structures
   - Comparing prose writing and code writing
   - Understanding bias in code reviews

● Career Plan
● Summary
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.
Hypothesis Time!

Is writing code like writing English?

Yes! Very similar!

No! Very Different!

Comparing prose writing and code writing (ICSE’20)
Are code writing and prose writing similar neural activities? Do I have to be good at English writing to become a good software developer?

```plaintext
I got in my ___ and I drove into town.

int sum(int a, int b) {
    /*YOUR CODE HERE*/
    return ___;
}
```
Challenge: Typing in fMRI

- fMRI-safe bespoke keyboard
- QWERTY keyboard
- Allow typing and editing
Challenge: Typing in fMRI

- fMRI-safe bespoke keyboard
- QWERTY keyboard
- Allow typing and editing
Experimental Design

- Two-by-two contrast task design: 30 participants

**CODE**

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 */
    }
}
```

**FITB**

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.
```

**PROSE**

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

```java
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?)

```java
1
```
Summary: Prose Writing vs. Coding Writing

- Code writing and prose writing are very distinct neural activities! ($2.4 < t < 6.2$)
  - Code writing: top-down control, memory, planning, spatial ability
  - Prose writing: language-related regions

Comparing prose writing and code writing *(ICSE’20)*
Comparing prose writing and code writing are very *distinct* neural activities! (2.4<t<6.2)

- **Code writing**: top-down control, memory, planning, spatial ability
- **Prose writing**: language-related regions

Comparing prose writing and code writing *(ICSE’20)*
Comparing prose writing and code writing (ICSE’20)

Summary: Prose Writing vs. Coding Writing

- Code writing and prose writing are very distinct neural activities! (2.4 < t < 6.2)
  - Code writing: top-down control, memory, planning, spatial ability
  - Prose writing: language-related regions

- Implications
  - Training and pedagogy
  - Broadening participation
  - Writing proofs?

Comparing prose writing and code writing (ICSE’20)
Outline

- Introduction
- Investigating cognition in software engineering
  - Understanding the neural representations of data structures
  - Comparing prose writing and code writing
  - Understanding bias in code reviews
- Career plan
- Summary
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

Delete the equal mark in case the array is like `{x,x,x...(n),y,y,y...(n+1)}`
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

Delete the equal mark in case the array is like \{x,x,x...\(n\),y,y,y,y...\(n+1\)}

Code changes
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

```
Delete the equal mark in case the array is like {x,x,x...(n),y,y,y...(n+1)}
```

```
2 algorithms/cpp/majorityElement/majorityElement.cpp

```cpp
@@ -32,7 +32,7 @@ int majorityElement(vector<int> &num) {
    cnt++;
   }
- if (cnt >= num.size()/2) return majority;
+ if (cnt > num.size()/2) return majority;
```

```
35
35 }
36
37
38 return majority;
```
Code Review

- Code review is *critical* for software development
  - *Systematic* inspection, analysis, evaluation, and revision of code.
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

Beyond the Code Itself: How Programmers *Really* Look at Pull Requests

Denae Ford, Mahnaz Behroozi  
North Carolina State University  
Raleigh, NC, USA  
{dford3, mbehroozi}@ncsu.edu

Alexander Serebrenik  
Eindhoven University of Technology  
Eindhoven, The Netherlands  
a.serebrenik@tue.nl

Chris Parnin  
North Carolina State University  
Raleigh, NC, USA  
cjparnin@ncsu.edu

Understanding bias in code reviews (*FSE’20*)
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

Beyond the Code Itself: How Programmers *Really* Look at Pull Requests

Denae Ford, Mahnaz Behroozi
North Carolina State University
Raleigh, NC, USA
{dford3, mbehroozi}@ncsu.edu

Alexander Serebrenik
Eindhoven University of Technology
Eindhoven, The Netherlands
a.serebrenik@tue.nl

Chris Parmin
North Carolina State University
Raleigh, NC, USA
cjparmin@ncsu.edu
Code Review

- Code review is **critical** for software development
  - **Systematic** inspection, analysis, evaluation, and revision of code.

Beyond the Code Itself:
How Programmers *Really* Look at Pull Requests

Denae Ford, Mahnaz Behroozi  
North Carolina State University  
Raleigh, NC, USA  
{dford3, mbehroozi}@ncsu.edu

Alexander Serebrenik  
Eindhoven University of Technology  
Eindhoven, The Netherlands  
a.serebrenik@tue.nl

Chris Parmin  
North Carolina State University  
Raleigh, NC, USA  
cjparmin@ncsu.edu

Understanding bias in code reviews *(FSE’20)*
Hypothesis Time!

Do women and men review code in the same way?

Yes! Very similar!

No! Very Different!
Is there **bias** on **gender** and **identities** in code review? How do we characterize the bias?

- Systematically
- Objectively
- Rigorously
Experimental Design

● 60 C/C++ pull requests from GitHub
Experimental Design

- 60 C/C++ pull requests from GitHub
Experimental Design

- 60 C/C++ pull requests from GitHub
60 C/C++ pull requests from GitHub

Author images: Relabel the author information
- Chicago Face Database

Experimental Design

- 60 Pull Requests
- 20 Women
- 20 Men
- 20 Machine

Understanding bias in code reviews (FSE’20)
Experimental Design

- 60 C/C++ pull requests from GitHub
- Author images: **Relabel the author information**
- **Construction** of code review stimuli

60 Pull Requests

- 20 Men
- 20 Women
- 20 Machine

60 Stimuli: V1
60 Stimuli: V2

Understanding bias in code reviews (FSE’20)
Experimental Design

- 60 C/C++ pull requests from GitHub
- **Author images:** Relabel the author information
- **Construction:**
  - 60 Pull Requests
  - 20 Machine
  - 20 Men
  - 20 Women

Delete the equal mark in case the array is like \{x,x,x...(n),y,y,y,y...(n+1)\}

Owner:

Accept  Reject

60 Pull Requests

60 Stimuli: V1

60 Stimuli: V2

Understanding bias in code reviews (FSE’20)
Experimental Design

- Avoid social desirability bias
- 37 Participants
- Post-survey questions

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Men</td>
<td>21</td>
</tr>
<tr>
<td>Women</td>
<td>16</td>
</tr>
<tr>
<td>Undergraduate</td>
<td>26</td>
</tr>
<tr>
<td>Graduate</td>
<td>11</td>
</tr>
</tbody>
</table>
Summary: Biases in Code Review

- We find **universal biases** in how all participants treat code reviews as a function of the reviewers’ gender and apparent author:
Summary: Biases in Code Review

● We find **universal biases** in how all participants treat code reviews as a function of the **reviewers’ gender** and **apparent author**:

  ● **Behavioral difference**
    i. All participants spend **less time** evaluating the Pull Requests of **women** (p<0.01)
    ii. All participants are **less likely to accept** the Pull Requests of **machines** (p<0.05)
    iii. **Women reviewers** spent **less time** on all Pull Requests (p<0.0001)
Summary: Biases in Code Review

We find universal biases in how all participants treat code reviews as a function of the reviewers’ gender and apparent author:

- Behavioral difference
  i. All participants spend less time evaluating the Pull Requests of women (p<0.01)
  ii. All participants are less likely to accept the Pull Requests of machines (p<0.05)
  iii. Women reviewers spent less time on all Pull Requests (p<0.0001)

- Visual difference
  i. Men and women reviewers employ different high-level problem-solving strategies (p<0.001)
  ii. Men fixated more frequently (p<0.001), while women spent significantly more time analyzing Pull Requests messages and author pictures.
Summary: Biases in Code Review

- We find **universal biases** in how all participants treat code reviews as a function of the reviewers’ gender and apparent author:
  - **Neurological difference**
    - It is possible to **distinguish** women and men conducting code review
Summary: Biases in Code Review

- We find **universal biases** in how all participants treat code reviews as a function of the reviewers’ gender and apparent author:
  - **Neurological difference**
    1. It is possible to **distinguish** women and men conducting code review
- Participants’ **self-reported perception** in code review **do not align** with the objective observations.
  - Do not realize the existence of difference on gender
  - Bias against machines exists
Summary: Biases in Code Review

- We find universal biases in how all participants treat code reviews as a function of the reviewers’ gender and apparent author:
  - Neurological difference
    - It is possible to distinguish women and men conducting code review
- Participants’ self-reported perception in code review do not align with the objective observations.
  - Do not realize the existence of difference on gender
  - Bias against machines exists

Implications
- How should we design code review environment based on the differences?
- Should we avoid showing authors’ profiles?
- Is there any effective training to mitigate the biases?
Thesis Scope: Metrics for Human Factors in SE

Objective
Cognitive/Neuro Patterns
Visual Focus
Behaviors, Motions
Emotions
Feedback, Opinions

Subjective

Light-Weight

Interview
Survey

Smart Band
Body Sensor
Smartphone

Eye Tracker
fNIRS
fMRI
Thesis Scope: Metrics for Human Factors in SE

Objective
Cognitive/Neuro Patterns
Visual Focus
Behaviors, Motions
Emotions
Feedback, Opinions

Subjective

Light-Weight
Interview
Survey

Smart Band
Body Sensor

Smartphone

fNIRS
Eye Tracker
fMRI

Heavy-Weight
Publications: Supporting this Thesis

1. Distilling Neural Representations of Data Structure Manipulation using fMRI and fNIRS. *(SIGSOFT Distinguished Paper Award)*
   Yu Huang, Xinyu Liu, Ryan Krueger, Tyler Santander, Xiaosu Hu, Kevin Leach, Westley Weimer.
   41st ACM/IEEE International Conference on Software Engineering (ICSE 2019).

   Ryan Krueger, Yu Huang, Xinyu Liu, Tyler Santander, Westley Weimer, Kevin Leach.
   42nd ACM/IEEE International Conference on Software Engineering (ICSE 2020).

   Yu Huang, Kevin Leach, Zohreh Sharafi, Tyler Santander, Westley Weimer.
   ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2020)

4. Towards an Objective Measure of Developers’ Cognitive Activities.
   Zohreh Sharafi, Yu Huang, Kevin Leach, Westley Weimer.
   ACM Transactions on Software Engineering and Methodology, Volume 30, Issue 3 (TOSEM 2021)

2. **Applying Automated Program Repair to Dataflow Programming Languages.** **Yu Huang**, Hammad Ahmad, Stephanie Forrest, Westley Weimer. 10th International Workshop on Genetic Improvement (GI 2021 @ ICSE 2021). To Appear.


7. **Physiological Changes Over the Course of Cognitive Bias Modification for Social Anxiety.** Mehdi Boukhechba, Jiaqi Gong, Kamran Kowsari, Mawulolo K Ameko, Karl Fua, Philip I Chow, **Yu Huang**, Bethany A Teachman, and Laura E Barnes. Biomedical & Health Informatics (BHI), 2018 IEEE EMBS International Conference on, pages 422–425.
Publications: Others

8. I Did OK, But Did I Like It? Using Ecological Momentary Assessment to Examine Perceptions of Social Interactions Associated with Severity of Social Anxiety and Depression. Emily C Geyer, Karl C Fua, Katharine E Daniel, Philip I Chow, Wes Bonelli, Yu Huang, Laura E Barnes, and Bethany A Teachman. Behavior therapy, 49(6):866–880, 2018


Career Plan

● Next stop: Assistant Professor (tenure-track) in Vanderbilt
Career Plan

- Next stop: Assistant Professor (tenure-track) in Vanderbilt
- Research plan
Career Plan

● Next stop: Assistant Professor (tenure-track) in Vanderbilt

● Research plan
  ● Cognitive modeling for SE
    ○ Non-functional factors in software development process
    ○ Development of expertise in CS
    ○ Integration of cognitive models in automated program repair
Career Plan

● Next stop: Assistant Professor (tenure-track) in Vanderbilt
● Research plan
  ● Cognitive modeling for SE
    ○ Non-functional factors in software development process
    ○ Development of expertise in CS
    ○ Integration of cognitive models in automated program repair
  ● Sustaining participation in/via Open Source Software for Social Good
  ● Supporting HW/SW co-design
Acknowledgement
Acknowledgement

My advisor: Prof. Westley Weimer
Acknowledgement

My advisor: Prof. Westley Weimer
Acknowledgement

My committee members

Prof. Westley Weimer  Prof. Stephanie Forrest  Prof. Mark Guzdial  Prof. Ioulia Kovelman
Dr. Kevin Leach  
@UM

Dr. Zohreh Sharafi  
@UM

Dr. Tyler Santander  
@UC Santa Barbara

Dr. Xiaosu Hu  
@UM

Dr. Andrew Jahn  
@UM

Dr. Denae Ford  
@MSR

Dr. Tom Zimmermann  
@MSR

Hammad Ahmad  
@UM

Dr. Stephanie Forrest  
@ASU

Dr. Kevin Angstadt  
@St. Lawrence
Acknowledgement

WRG: Kevin Leach, Zohreh Sharafi, Kevin Angstadt, Madeline Endres, Colton Holoday, Yirui Liu, Hammad Ahmad, Annie Li, Fee Christoph, Yiannos Demetriou…

And all my friends!
Acknowledgement

WRG: Kevin Leach, Zohreh Sharafi, Kevin Angstadt, Madeline Endres, Colton Holoday, Yirui Liu, Hammad Ahmad, Annie Li, Fee Christoph, Yiannos Demetriou… And all my friends!
Acknowledgement

• My parents: Zhonglin Huang and Qian Xu
• My husband: Kevin Leach
• My parents-in-law: Richard and Linda Leach
• My brother-in-law: Eric Leach
Summary

- Measure cognitive processes in software activities

De-identified data is public: [http://www-personal.umich.edu/~yhhy/](http://www-personal.umich.edu/~yhhy/)
Summary

- Measure cognitive processes in software activities
- Novel concept, problems, and approaches

De-identified data is public: [http://www-personal.umich.edu/~yhhy/](http://www-personal.umich.edu/~yhhy/)
Summary

- Measure cognitive processes in software activities
- Novel concept, problems, and approaches
- Rigorous controlled experiments

De-identified data is public: http://www-personal.umich.edu/~yhhy/
Summary

- Measure cognitive processes in software activities
- Novel concepts, problems, and approaches
- Rigorous controlled experiments
- Potentials for broad impact

De-identified data is public: http://www-personal.umich.edu/~yhhy/