

Understanding User Cognition: from Everyday Behavior and Spatial Ability to Code Writing and Review

Yu Huang

University of Michigan
Dec 11, 2019



COLLEGE OF ENGINEERING
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

Break Down the Title

- A standard workday of a software developer

Break Down the Title

- A standard workday of a software developer



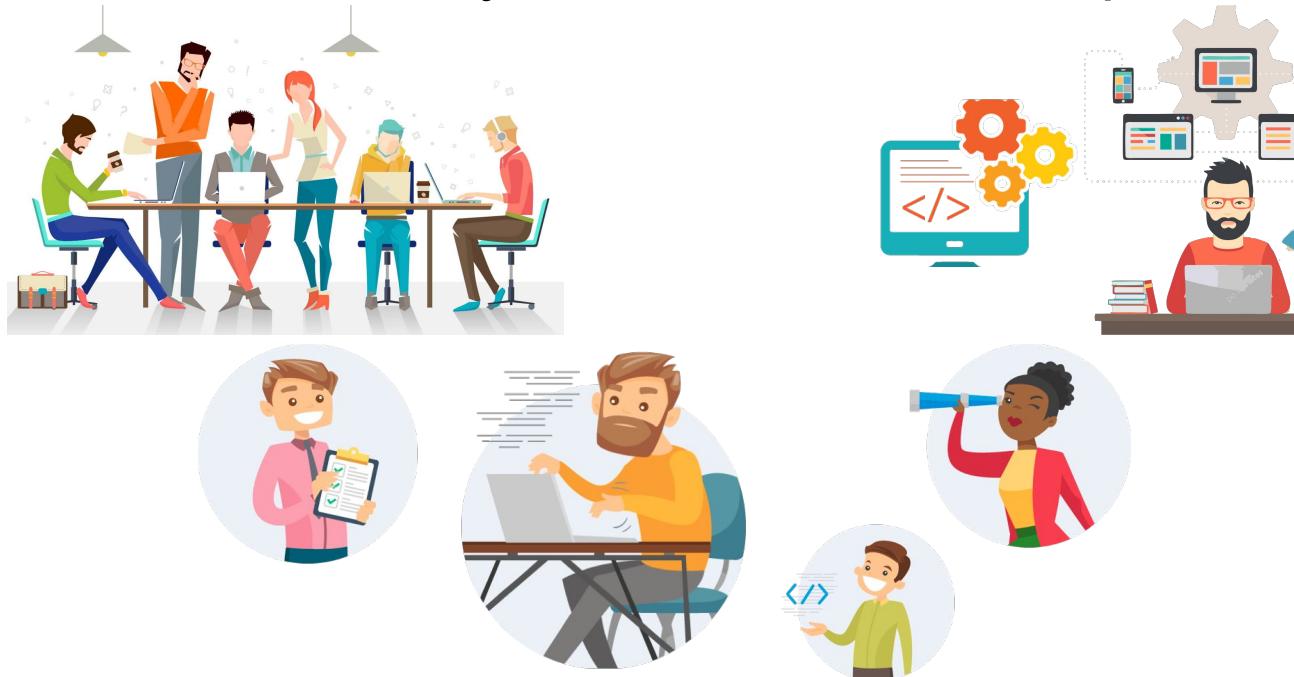
Break Down the Title

- A standard workday of a software developer



Break Down the Title

- A standard workday of a software developer



Break Down the Title

- A standard workday of a software developer

What could go wrong?

What is currently holding us back?

Break Down the Title

- A standard workday of a software developer



We need to talk about Silicon Valley's mental health problems

Ash Huang
4/09/15 2:49PM • Filed to



Depression & Anxiety in programming

Phil Walker Oct 9 '17 · 1 min read

#mentalhealth #discuss #career

I recently attended a developer retreat in beautiful Golden, Colorado and ran into something that surprised me. We were doing an ice breaker exercise and

WELL-BEING WISDOM WONDER PURPOSE SLEEP SPECIAL SECTIONS COMMUNITY

MENTAL HEALTH AT WORK // August 28, 2018

A Conversation On Workplace Mental Health in Silicon Valley

ew with Sunil Rajaraman
and Communications at Mind Share Partners



WELLNESS, STARTUPS, AND
MENTAL HEALTH: CONFRONTING
A GROWING CRISIS



On Dealing with Anxiety and Depression as a Developer

Michael Scott Hertzberg Sep 1 '18 Updated on Nov 21, 2018 · 1 min read

#discuss #career #health #life

Break Down the Title

- A standard workday of a software developer

The image is a collage of three screenshots from news articles. The top-left screenshot shows a headline "Silicon Valley Has a Health Crisis Too" with a sub-headline "We need to talk about Software developer problems". The author is Ash Huang, dated 4/09/15 2:49PM. The bottom-left screenshot shows a headline "Depression programs" by Phil Walker. The bottom-right screenshot shows a headline "Mental Health in UPS, AND CONFRONTING" with a sub-headline "Anxiety and Depression as a Developer". The central part of the collage has a large red text overlay that reads "Leads to impairment in academic functioning and relationship!".

- **62%** have mental complaints
- **31%** have mental ill-health
- **<1%** seeked for professional help

Leads to impairment in academic functioning and relationship!

I recently attended a developer retreat in beautiful Golden, Colorado and ran into something that surprised me. We were doing an ice breaker exercise and

Break Down the Title

- How can we be **more effective and efficient** in programming? What are the **cognitive processes** of programming? What affects our **decisions** in programming?

Break Down the Title

- How can we be **more effective and efficient** in programming? What are the **cognitive processes** of programming? What affects our **decisions** in programming?
 - Traditional research solutions: self-reporting

Break Down the Title

- How can we be **more effective and efficient** in programming? What are the **cognitive processes** of programming? What affects our **decisions** in programming?
 - Traditional research solutions: self-reporting
 - Unreliable**

Self-Reports in Organizational Problems and Prospects

Philip M. Podsakoff

Dennis W. Organ

Indiana University

Self-reports figure prominently in organizational and management research, but there are several problems associated with their use. This article identifies six categories of self-reports and discusses common method variance, the consistency of responses, social desirability, statistical and post hoc remedies and alternative methods for dealing with artifactual bias are presented. Recommendations for future research are also offered.

Faking It: Social Desirability Response Bias in Self-report Research

Australian Journal of Advanced Nursing
Volume 25 Issue 4 (June/Aug 2008)

van de Mortel, Thea F¹

Abstract: Objective: The tendency for questionnaires to elicit socially desirable responses is called social desirability bias. This bias can create false relationships or obscure true relationships. Methods: Several scales can be used to detect, minimize, and control social desirability bias in questionnaire-based research. The aim of this article is to review related studies that used questionnaire-based research methods to examine social desirability bias.

International Journal for Quality in Health Care 1999; Volume 11, Number 3: pp. 187-192

Evidence of self-report bias in assessing adherence to guidelines

JAN LOMAS² AND DENNIS ROSS-DEGNAN¹

Journal of Business and Psychology, Vol. 17, No. 2, Winter 2002 (©2002)

UNDERSTANDING SELF-REPORT BIAS IN ORGANIZATIONAL BEHAVIOR RESEARCH

Stewart I. Donaldson

Claremont Graduate University

Elisa J. Grant-Vallone

California State University, San Marcos

Break Down the Title

- How can we be **more effective and efficient** in programming? What are the **cognitive processes** of programming? What affects our **decisions** in programming?
 - Traditional research solutions: **Unreliable** self-reporting
 - **Observed potential bias** of non-functional factors

✓ PEER-REVIEWED

Gender differences and bias in open source: pull request acceptance of women versus men

Research Software Engineering
Software Engineering
Josh Ter Emerson
May 1, 2019

2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)

Investigating the Effects of Gender Bias on GitHub

Nasif Imtiaz¹, Justin Middleton¹, Joymallya Chakraborty¹, Neill Robson¹, Gina Bai¹, and Emerson Murphy-Hill^{*2}

¹Department of Computer Science, North Carolina State University
²Google, LLC
 {simtiaz, jamiddl2, jchakra, nrobsn, rbai2}@ncsu.edu, emersonm@google.com

Geographical bias in GitHub : perceptions and reality

Rastogi, Ayushi; Nagappan, Nachiappan; Gousios, Georgios

URI: <https://repository.iitd.edu.in/ispui/handle/123456789/388>

Date: 2016-01-11

Trust in Automated Software Repair

The Effects of Repair Source, Transparency, and Programmer Experience on Perceived Trustworthiness and Trust

Authors
Tyler J. Ryan [✉](#), Gene M. Alarcon, Charles Walter, Rose Gamble, Sarah A. Jessup, August Capiola, Marc D. Pfahler

Authors and affiliations

Conference paper
First Online: 12 June 2019

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 11594)

Break Down the Title

- How can we be **more effective and efficient** in programming? What are the **cognitive processes** of programming? What affects our **decisions** in programming?
 - Traditional research solutions: **Unreliable** self-reporting
 - Observed potential bias of non-functional factors

Lack a **foundational understanding**

Desired properties for this proposal

- Bring all the concerns together:

Desired properties for this proposal

- Bring all the concerns together:
 - **Objective** measures
 - Not just self-reporting

Desired properties for this proposal

- Bring all the concerns together:
 - **Objective** measures
 - Not just self-reporting
 - **Foundational understanding** of software activities
 - What are the cognitive processes of programming?

Desired properties for this proposal

- Bring all the concerns together:
 - **Objective** measures
 - Not just self-reporting
 - **Foundational understanding** of software activities
 - What are the cognitive processes of programming?
 - **Higher-level** programming tasks
 - Data structures; code writing; code reviews

Desired properties for this proposal

- Bring all the concerns together:
 - **Objective** measures
 - Not just self-reporting
 - **Foundational understanding** of software activities
 - What are the cognitive processes of programming?
 - **Higher-level** programming tasks
 - Data structures; code writing; code reviews
 - **Generalizability** across different user groups
 - How is productivity mitigated by group difference

Insights

- It is now **possible** to conduct studies that acquire **objective data** to understand the underlying **cognitive processes** of certain tasks
 - Mobile crowdsensing (MCS); medical imaging; eye tracking

Insights

- It is now **possible** to conduct studies that acquire **objective data** to understand the underlying **cognitive processes** of certain tasks
 - Mobile crowdsensing (MCS); medical imaging; eye tracking
- We can adapt **scientific approaches and concepts** from **other domains** to assist our investigation and understanding of certain tasks
 - Social anxiety; spatial ability; creative writing

Insights

- It is now **possible** to conduct studies that acquire **objective data** to understand the underlying **cognitive processes** of certain tasks
 - Mobile crowdsensing (MCS); medical imaging; eye tracking
- We can adapt **scientific approaches and concepts** from **other domains** to assist our investigation and understanding of certain tasks
 - Social anxiety; spatial ability; creative writing
- It is now **possible** to study historically-subjective factors by designing rigorous **controlled experiments**
 - Contrast-based experiments

Thesis Statement

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

Proposal Overview: Four Components

Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing

Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representation of data structures

Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representation of data structures



Comparing prose writing and code writing

Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representations of data structures



Comparing prose writing and code writing



Understanding bias in code reviews

Monitoring Mental Health Using Mobile Crowdsensing (MCS)

- Can we monitor humans' mental health status **objectively** via their everyday behaviors in a **natural setting**?



Monitoring Mental Health Using Mobile Crowdensing

- **Sensus:** Cross-platform, general MCS mobile application for human-subject studies
- **A MCS-based framework:** understanding the relationship between human behaviors and mental health status



Sensus: Cross-Platform, General MCS



Sensus: Cross-Platform, General MCS

1. Target **heterogeneous** mobile infrastructures



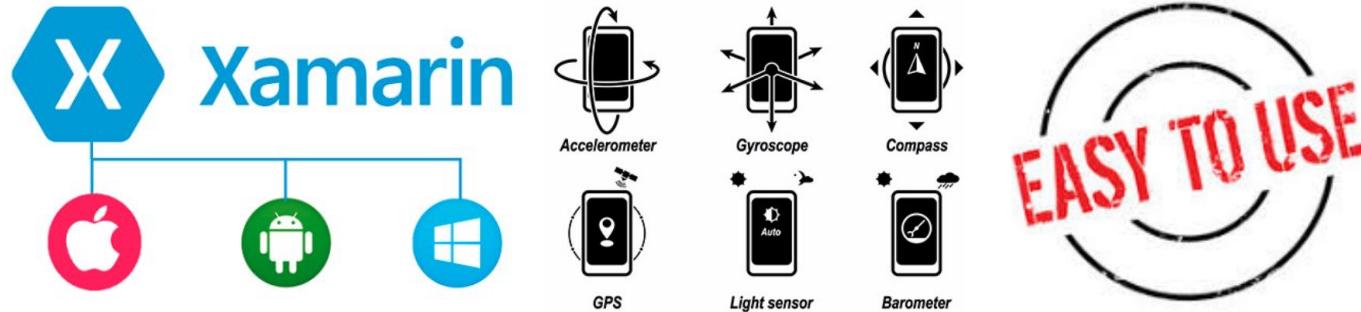
Sensus: Cross-Platform, General MCS

1. Target **heterogeneous** mobile infrastructures
2. Support a **wide range** of MCS-based human studies



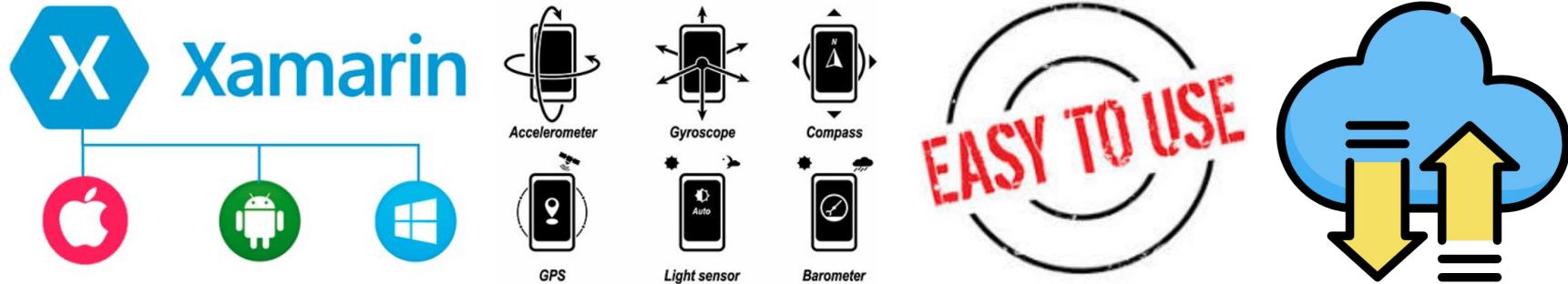
Sensus: Cross-Platform, General MCS

1. Target **heterogeneous** mobile infrastructures
2. Support a **wide range** of MCS-based human studies
3. **Eliminate** the need for **programming background**



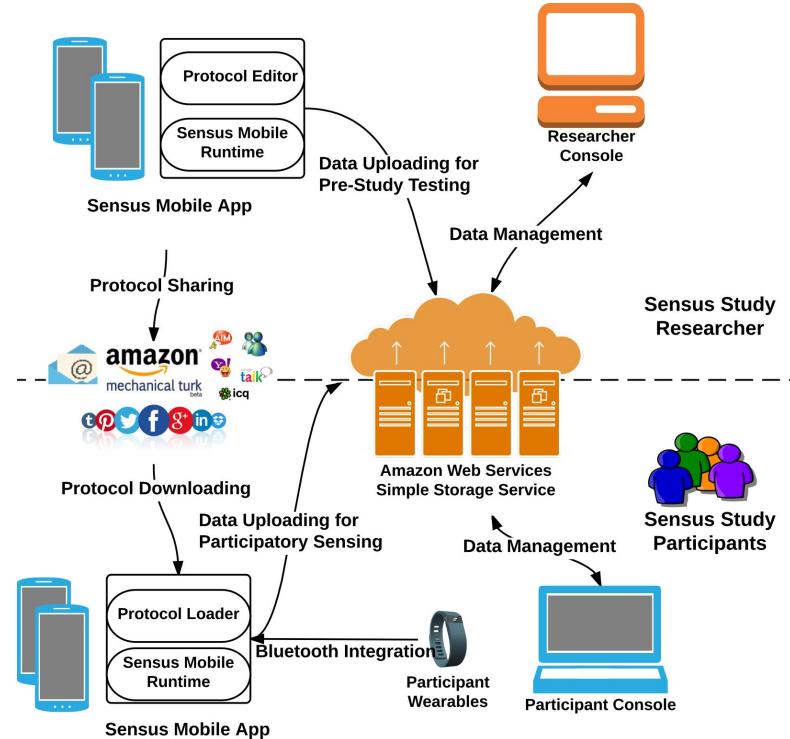
Sensus: Cross-Platform, General MCS

1. Target **heterogeneous** mobile infrastructures
2. Support a **wide range** of MCS-based human studies
3. **Eliminate** the need for **programming background**
4. Rely on **readily-available** mobile devices and cloud storage



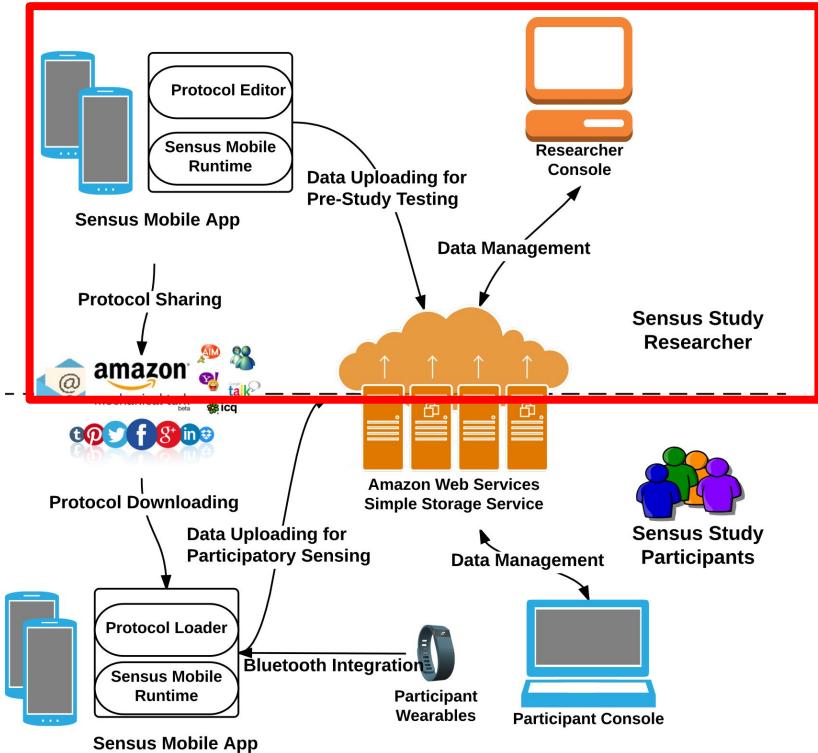
Architecture of Sensus: High-Level Design

- High-level design of **Sensus**
 - Cloud storage
 - Amazon AWS S3



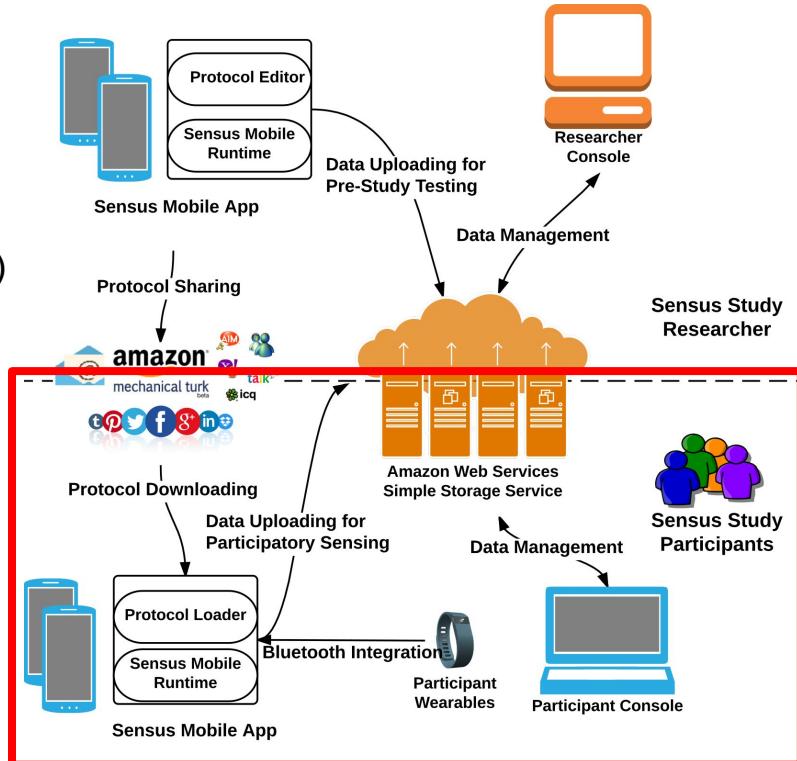
Architecture of Sensus: High-Level Design

- High-level design of **Sensus**
 - Cloud storage
 - Amazon AWS S3
 - Users
 - Researchers (study designers)
 - Participants



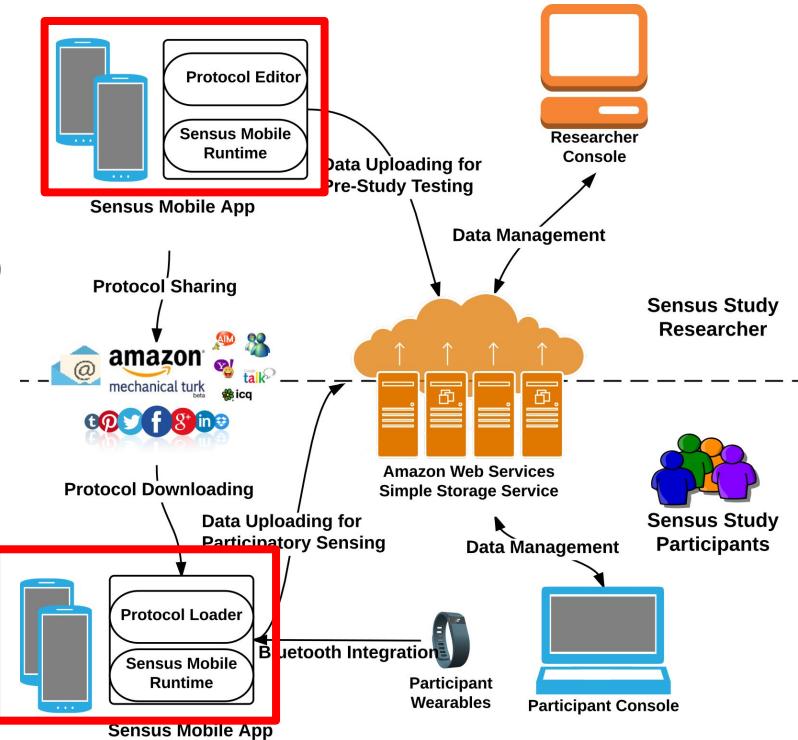
Architecture of Sensus: High-Level Design

- High-level design of **Sensus**
 - Cloud storage
 - Amazon AWS S3
 - Users
 - Researchers (study designers)
 - Participants



Architecture of Sensus: High-Level Design

- High-level design of **Sensus**
 - Cloud storage
 - Amazon AWS S3
 - Users
 - Researchers (study designers)
 - Participants
 - Protocols
 - Sensing plans
 - Probes
 - Surveys
 - Customized scheduling
 - JSON file



Sensus: An Example Case

- A Sensus protocol example (iOS)

The figure consists of four screenshots of the Sensus mobile application interface, arranged horizontally. Each screenshot shows a different screen of the protocol creation and configuration process.

- Screenshot 1: Your Sensus Studies Protocol**
This screen shows basic protocol settings:
 - Name: SALMON
 - Shareable: On
 - Description: This is a protocol for the SALMON research study.
 - Participation Horizon (Days): 10
 - Contact Email: kcf3st@virginia.edu
 - Groupable: On
 - Force Reports to Remote: On
 - Reward Threshold: 0.5
- Screenshot 2: Protocol Probes**
This screen lists probe types categorized by color:
 - Red: Acceleration (Listening), Battery Level (Polling), Compass Heading (Listening), Facebook Profile (Polling), GPS Location (Polling), GPS Location (Listening), Phone Call Metadata (Polling), Points of Interest Proximity (Polling), Points of Interest Proximity (Listening), Scripted Interactions, Sound Level (Polling), Speed (Polling).
 - Green: None
- Screenshot 3: Inputs**
This screen configures an input for the RP1 probe:
 - Name: RP1
 - Label Text: How positive are you feeling?
 - Required: On
 - Tip Text: (empty)
 - Minimum: 1
 - Maximum: 100
 - Increment: 1
 - Left Label: (empty)
- Screenshot 4: Input Detail**
This screen shows the configuration for the RP1 probe:
 - RP1
 - Progress: 6%
 - Required fields are indicated with *
 - *1) How positive are you feeling?
A slider from "Not at all" to "Very positive" is shown, currently set around 50.
 - Previous and Next buttons are visible.



Sensus: Metrics

- **Sensus** can be used in real-world scalable human-subjects studies
 - Release **Sensus**
 - Conduct real-world studies using **Sensus**
- **Sensus** is easy for researchers without engineering background to use
 - Interview researchers who used **Sensus** but without engineering backgrounds



Sensus: Preliminary Results

- Apple App Store
- Google Play Store: **500+**
- > 200 subjects in research studies



SensusMobile

UVa Predictive Technology Laboratory



SensusMobile

UVA Apps, LLC



*Sensus development website: <https://predictive-technology-laboratory.github.io/sensus/index.html>

*For more design details, please refer to our paper: *Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber.*

[Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.](#) In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.



Sensus: Preliminary Results

- Apple App Store
- Google Play Store: **500+**
- **> 200** subjects in research studies
- Feedback from the Psychologists (2 studies)
 - **Easy to use, intuitive experience**



SensusMobile

UVa Predictive Technology Laboratory



SensusMobile

UVA Apps, LLC



*Sensus development website: <https://predictive-technology-laboratory.github.io/sensus/index.html>

*For more design details, please refer to our paper: *Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber.*

[Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.](#) In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.



Sensus: Preliminary Results

- Apple App Store
- Google Play Store: **500+**
- **> 200** subjects in research studies
- Feedback from the Psychologists (2 studies)
 - **Easy to use, experience is intuitive**
 - Does **not require extra** engineering knowledge as long as you know how to use a smartphone



SensusMobile
UVa Predictive Technology Laboratory



SensusMobile
UVA Apps, LLC

*Sensus development website: <https://predictive-technology-laboratory.github.io/sensus/index.html>

*For more design details, please refer to our paper: *Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber.*

[Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.](#) In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.



Sensus: Preliminary Results

- Apple App Store
- Google Play Store: **500+**
- **> 200** subjects in research studies
- Feedback from the Psychologists (2 studies)
 - **Easy to use, intuitive experience**
 - Does **not require extra** engineering knowledge as long as you know how to use a smartphone
 - **Able** to get the data they want and obtain **meaningful** results



SensusMobile
UVA Predictive Technology Laboratory



SensusMobile
UVA Apps, LLC

*Sensus development website: <https://predictive-technology-laboratory.github.io/sensus/index.html>

*For more design details, please refer to our paper: *Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber.*

[Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.](#) In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.



Sensus: Preliminary Results

- Apple App Store
- Google Play Store: **500+**
- **> 200** subjects in research studies
- Feedback from the Psychologists (2 studies)
 - **Easy to use, intuitive experience**
 - Does **not require extra** engineering knowledge as long as you know how to use a smartphone
 - **Able** to get the data they want and obtain **meaningful** results
 - A desktop or web-based protocol design tool would be useful



SensusMobile
UVa Predictive Technology Laboratory



SensusMobile
UVA Apps, LLC

*Sensus development website: <https://predictive-technology-laboratory.github.io/sensus/index.html>

*For more design details, please refer to our paper: *Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber.*

[Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.](#) In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.



Monitoring Mental Health Using Mobile Crowdensing

- Recall: Can we monitor humans' mental health status **objectively** via their everyday behaviors in a **natural setting**?
 - We already have an MCS mobile application: **Sensus**



Monitoring Mental Health Using Mobile Crowdensing

- **Sensus:** Cross-platform, general MCS mobile application for human-subjects studies
- **A MCS-based framework:** understanding the relationship between human behaviors and mental health status



A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Fine-grained human behaviors vs. Mental health status
 - Objective measures from **Sensus**
 - **GPS**: mobility patterns with semantics
 - **Accelerometer (3-axis)**: micro-level motions
 - **Smartphone metadata**: call and text logs



A MCS-based Framework: Understanding Behaviors and Mental Health Status

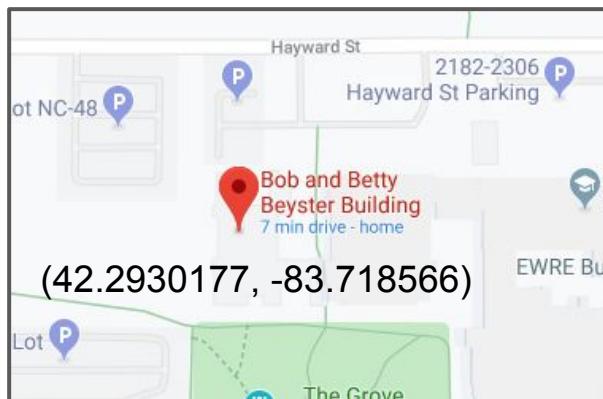
- Fine-grained human behaviors vs. Mental health status
 - Objective measures from **Sensus**
 - **GPS**: mobility patterns with semantics
 - **Accelerometer (3-axis)**: micro-level motions
 - **Smartphone metadata**: call and text logs
 - **Social anxiety levels**: SIAS score (0-80)



A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Semantics of locations

- (42.2930177, -83.718566) => School
- Point of Interest (POI) information obtained from *Foursquare*
- Clustering spatially and temporally
- Categories of location semantics



{
Education.
Bob and Betty Beyster Building.
Department of Computer Science
and Engineering.
University of Michigan.
}



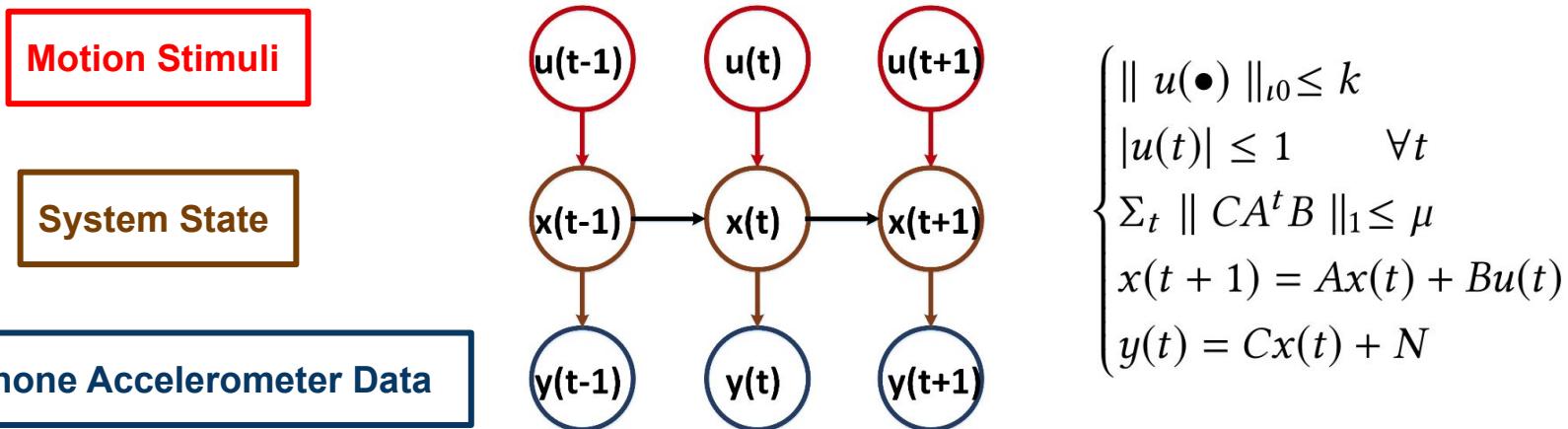
A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Semantics of locations
- Micro-level behaviors (behavioral dynamics)
 - Linear dynamic system (LDS)



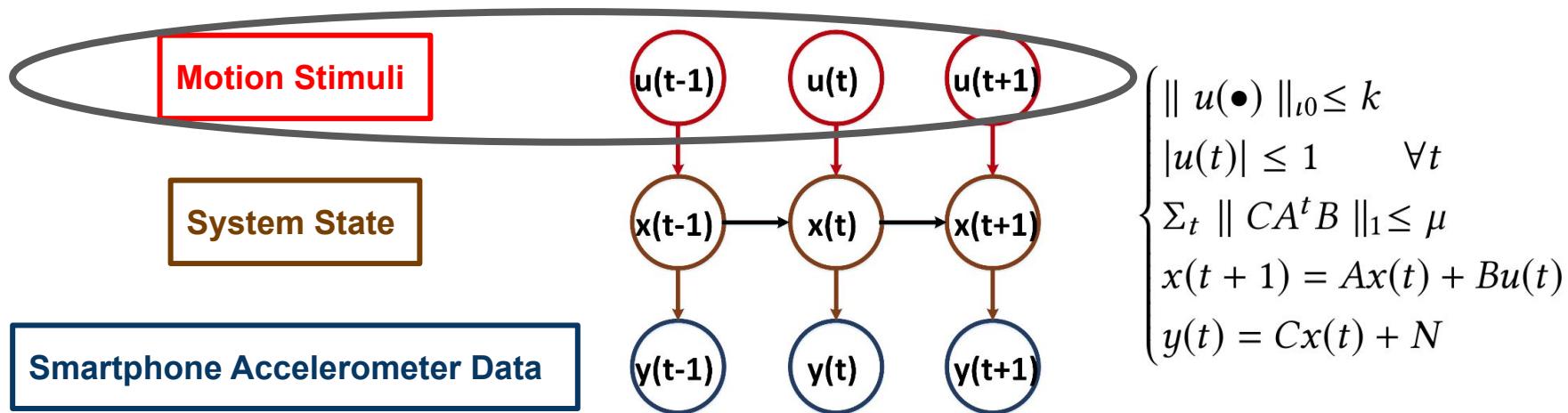
A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Semantics of locations
- Micro-level behaviors (behavioral dynamics)
 - Linear dynamic system (LDS)



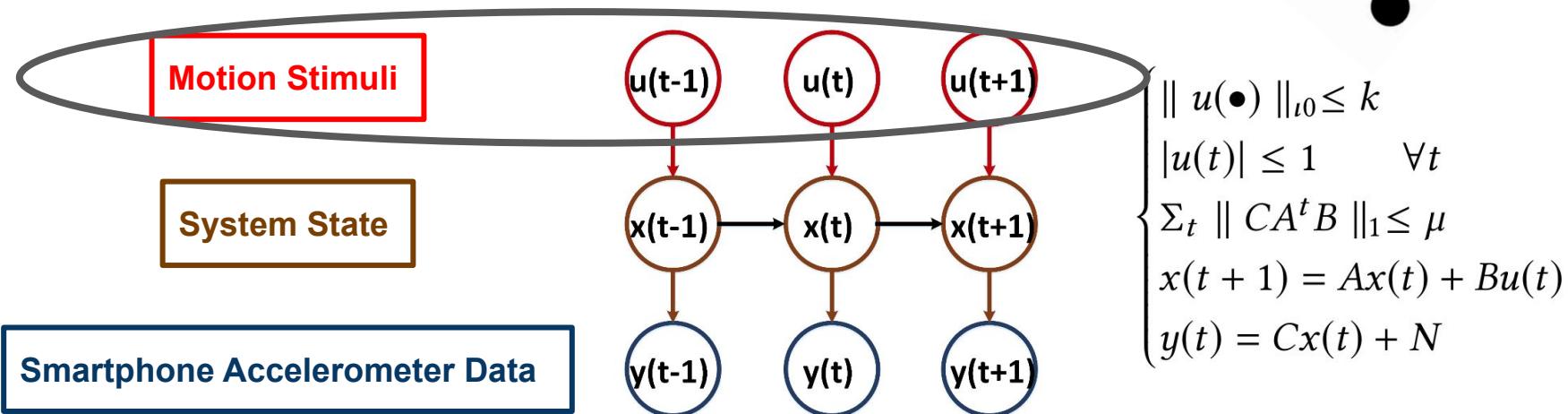
A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Semantics of locations
- Micro-level behaviors (behavioral dynamics)
 - Linear dynamic system (LDS)



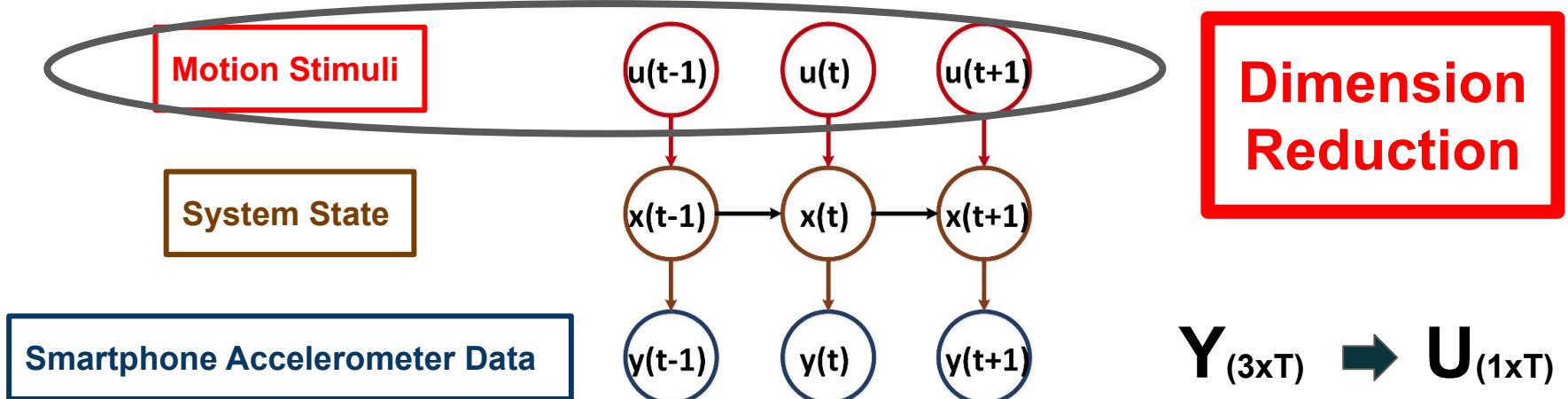
A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Semantics of locations
- Micro-level behaviors (behavioral dynamics)
 - Linear dynamic system (LDS)



A MCS-based Framework: Understanding Behaviors and Mental Health Status

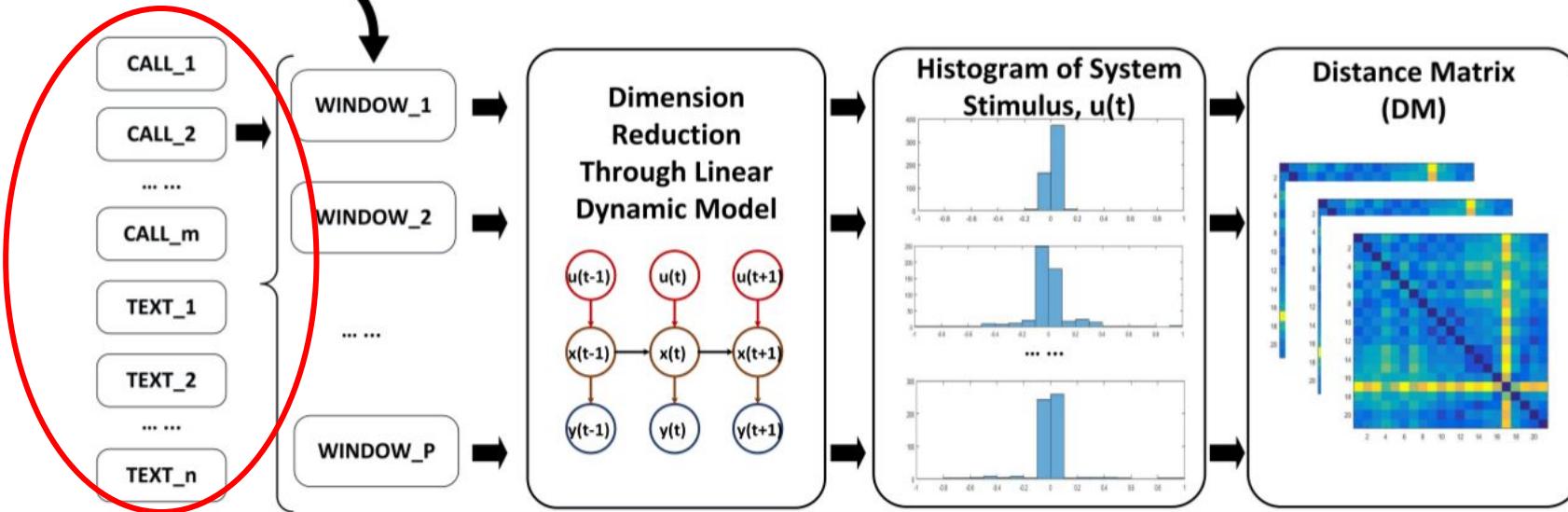
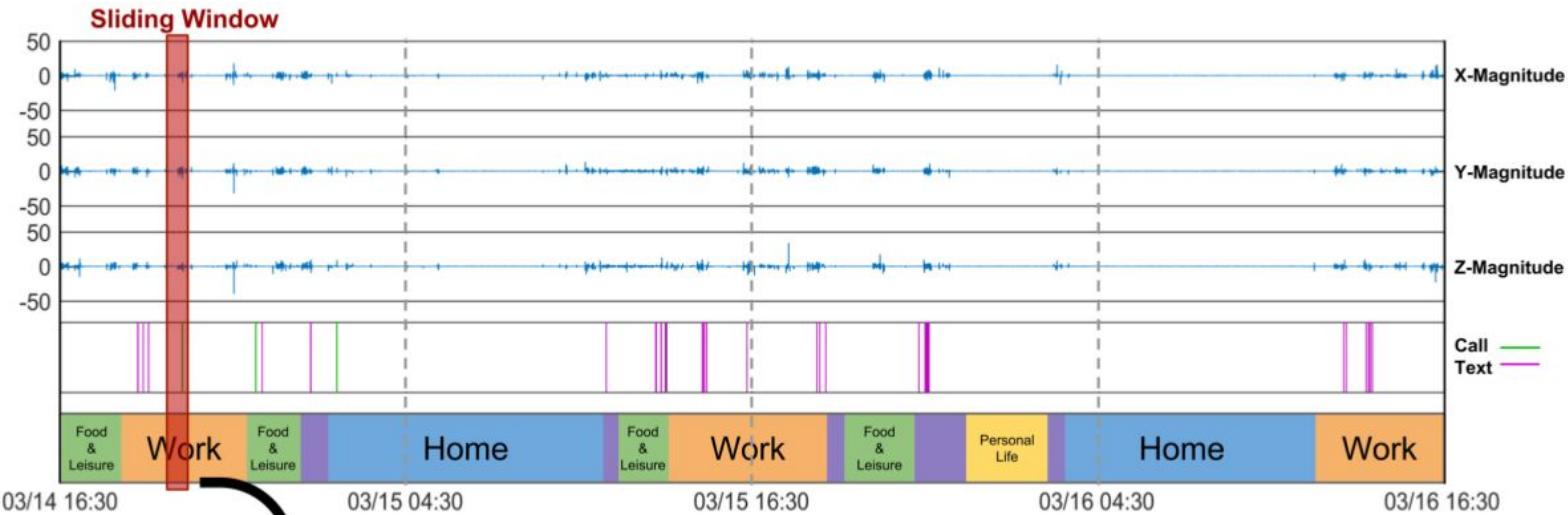
- Semantics of locations
- Micro-level behaviors (behavioral dynamics)
 - Linear dynamic system (LDS)

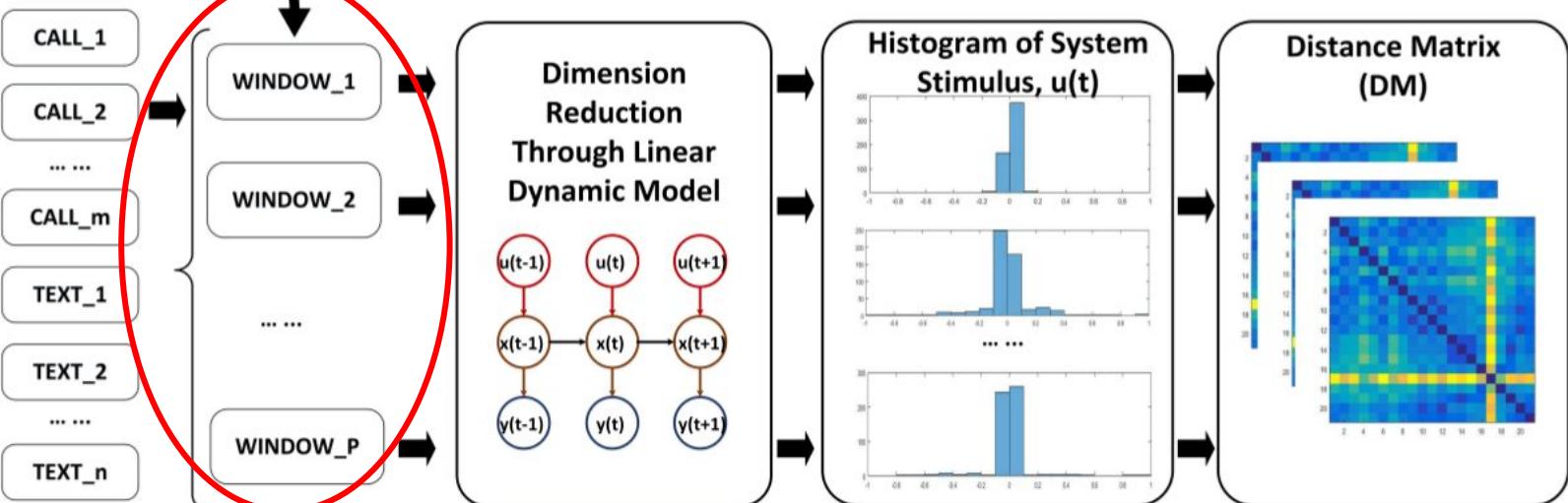


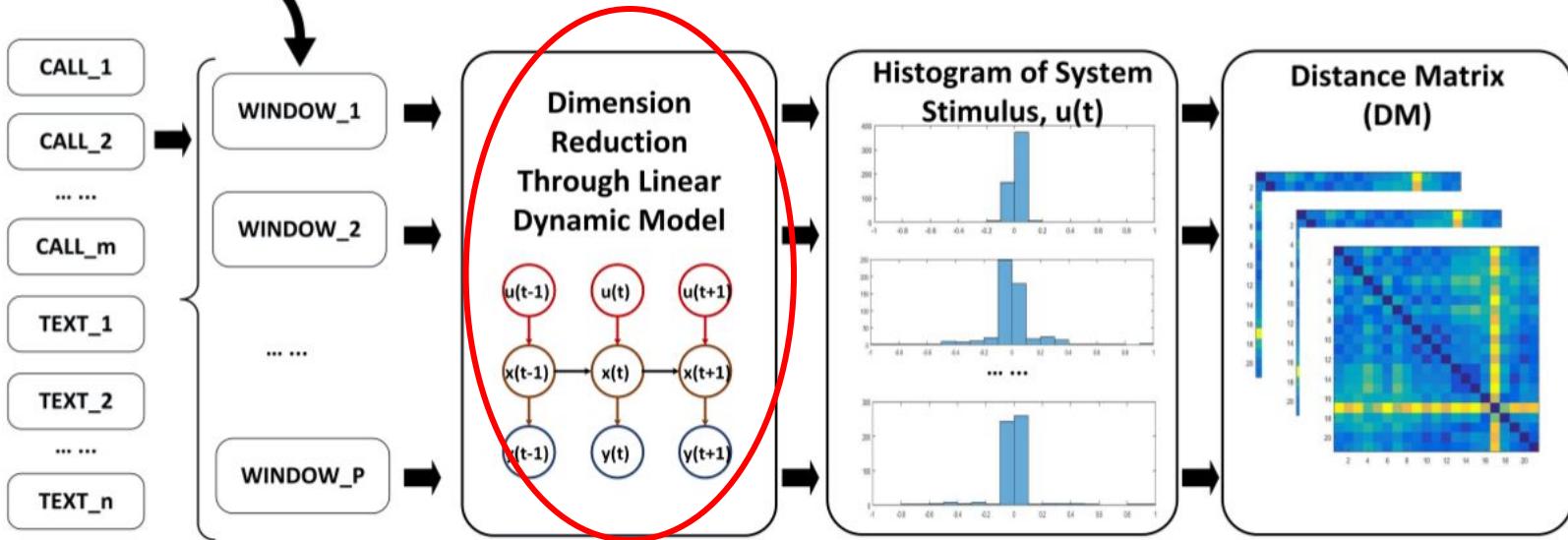
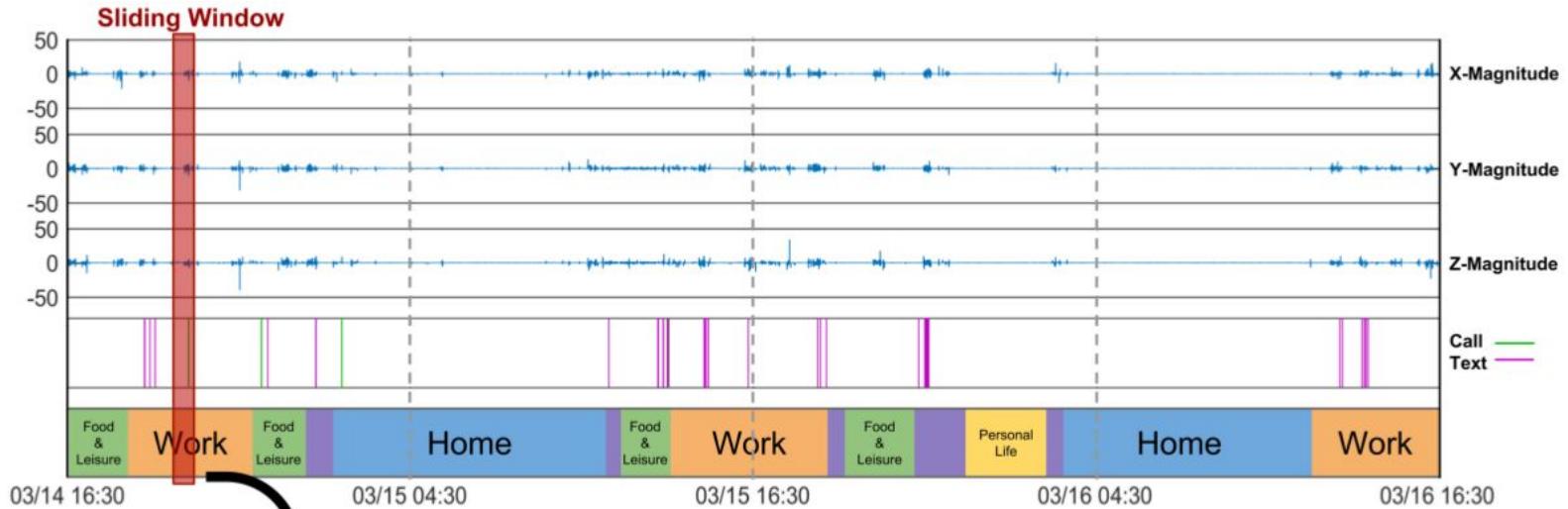
A MCS-based Framework: Understanding Behaviors and Mental Health Status

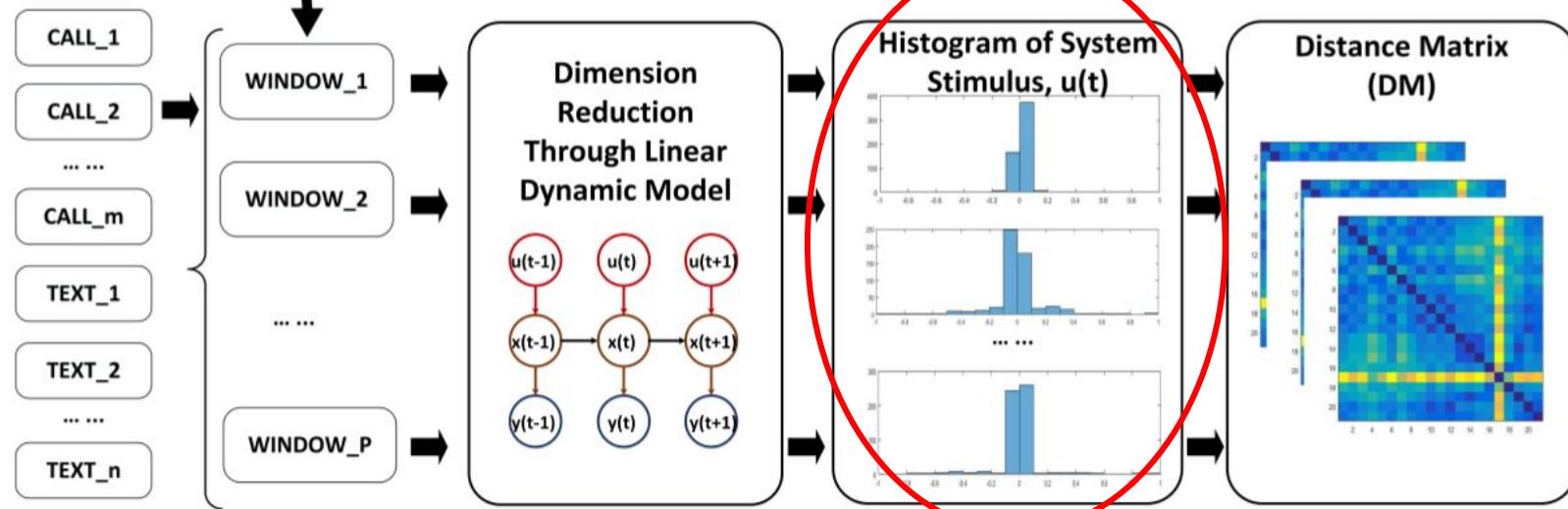
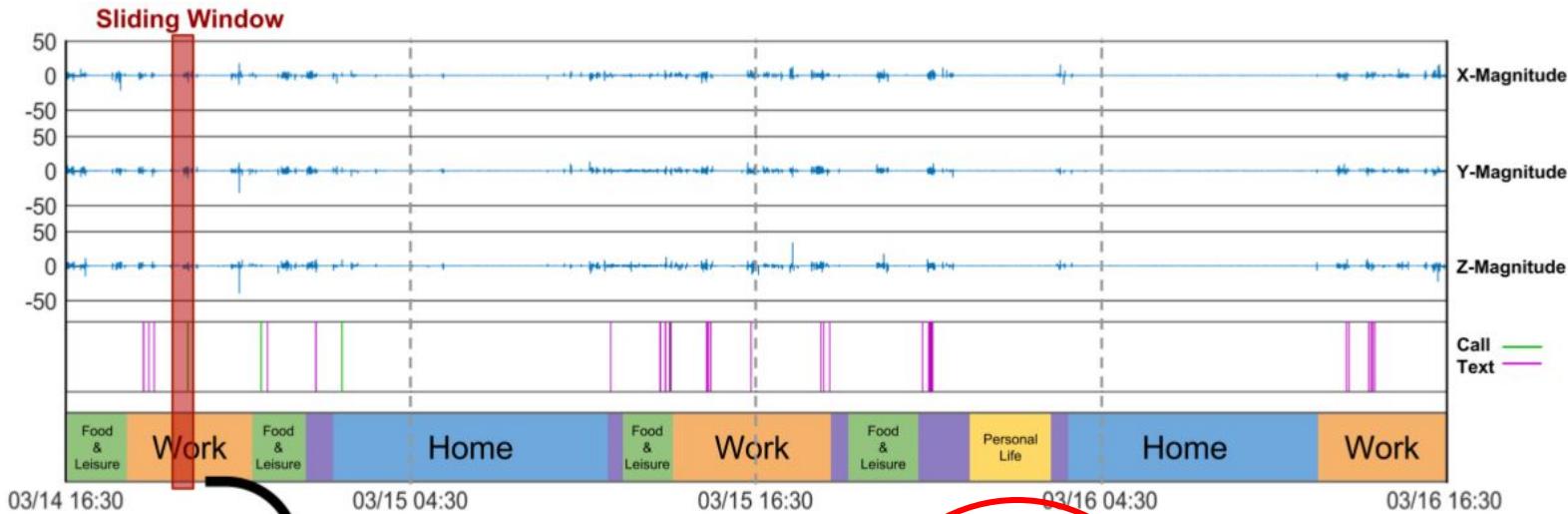
- The architecture of the MCS-based framework

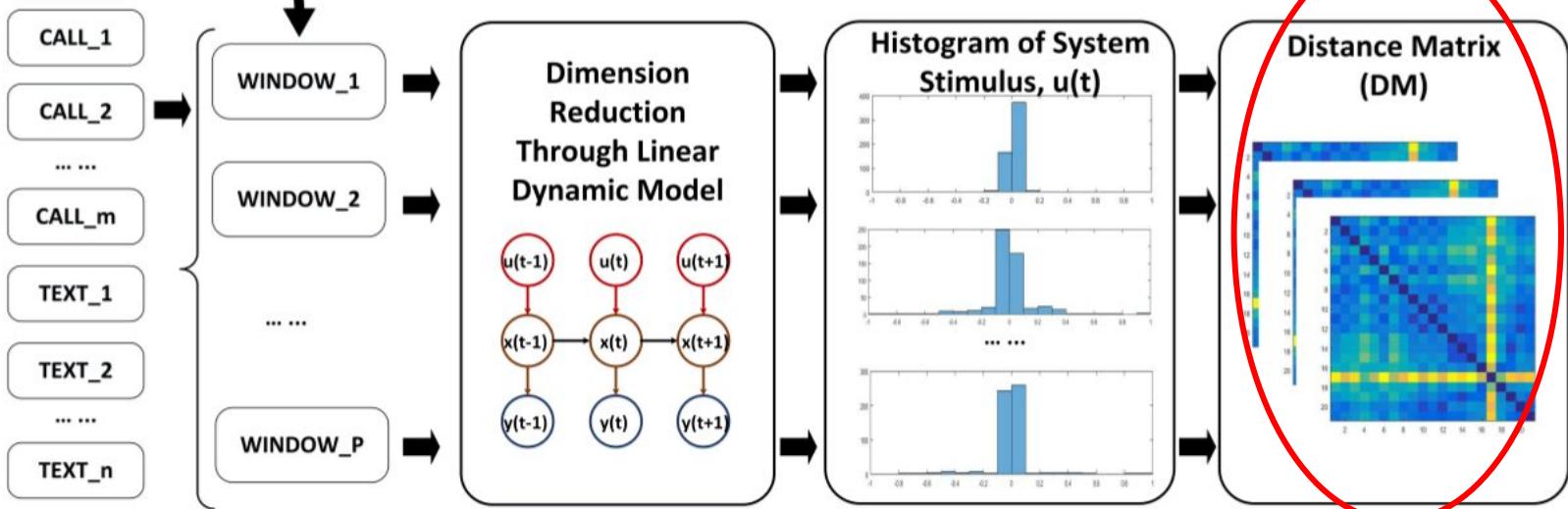
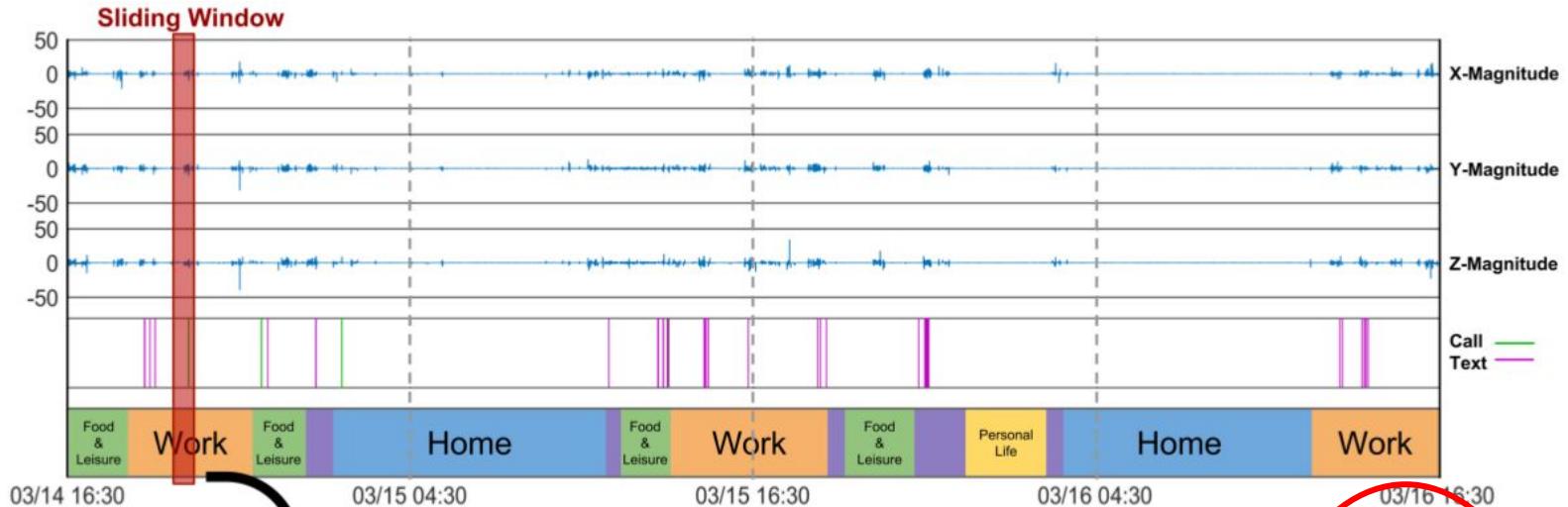






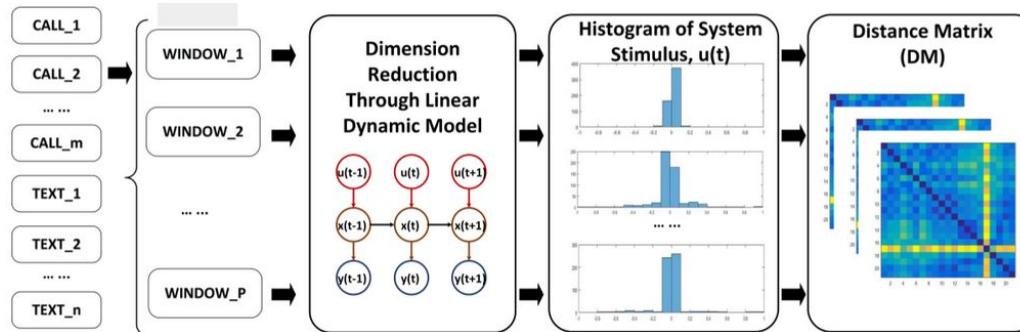






A MCS-based Framework: Understanding Behaviors and Mental Health Status

- Feature extraction



Term	Definition
<i>Call_Proportion</i>	The proportions of phone calls at different locations
<i>Text_Proportion</i>	The proportions of text messages at different locations
\overline{FAC}_1	The average of the mean values of all distance matrices ($DM(i)$) belonging to a subject
\overline{FAC}_2	The average of the standard deviations of all distance matrices ($DM(i)$) belonging to a subject
<i>MC</i>	The metric for a phone call event
<i>MT</i>	The metric for a text message event



A MCS-based Framework: Metrics

- In real-world human-subjects studies, we can objectively measure humans' behaviors in a natural setting
- From the objectively collected data, we can extract meaningful features
- We can find features that have a significant correlation with mental health status ($p<0.05$)



A MCS-based Framework: Preliminary Results

- Human study of 52 participants
 - *Sensus*
 - Duration: 14 days
 - SIAS: mean = 35.02, std = 12.10
 - Correlations between behavioral dynamics and social anxiety levels under different social contexts
-

Matrix feature	Call (MC)		Text (MT)	
	Pearson r	p-value	Pearson r	p-value
FAC_1	0.2867	0.0457	0.1961	0.1634
FAC_2	0.3041	0.0336	0.2342	0.0946



A MCS-based Framework: Preliminary Results

- Correlations between behavioral dynamics and social anxiety levels under different social contexts

Location	Call_Proportion					Text_Proportion				
	Pearson r	p-value	\bar{x}	σ		Pearson r	p-value	\bar{x}	σ	
Work	-0.1806	0.2142	0.0935	0.1074		-0.2511	0.0725	0.1441	0.1040	
Home	0.3983	0.0045	0.3868	0.2484		0.4059	0.0028	0.3989	0.2128	
Food & leisure	-0.2342	0.1053	0.1188	0.1551		-0.0882	0.5340	0.1412	0.1423	
Personal life	0.1234	0.3982	0.0138	0.0346		-0.2917	0.0359	0.0166	0.0228	
Transition	-0.0715	0.6141	0.3200	0.1812		-0.0707	0.6045	0.2381	0.1153	

*Refer to the paper for more details: Jiaqi Gong, Yu Huang, Philip I Chow, Karl Fua, Matthew Gerber, Bethany Teachman, Laura Barnes. [Understanding Behavioral Dynamics of Social Anxiety Among College Students Through Smartphone Sensors](#). *Information Fusion*, 49:57–68, September 2019.



Monitoring Mental Health Using Mobile Crowdensing

- Recall: Can we monitor humans' mental health status **objectively** via their everyday behaviors in a **natural setting**?

Yes, we can.



Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representations of data structures



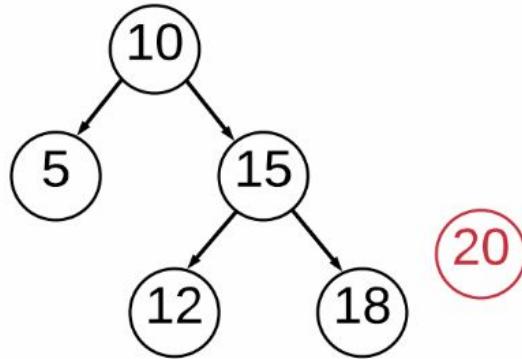
Comparing prose writing and code writing



Understanding bias in code reviews

Understanding the Neural Representations of Data Structure Manipulations

- How do human brains represent data structures? Is it more like **text** or more like **3D objects**?

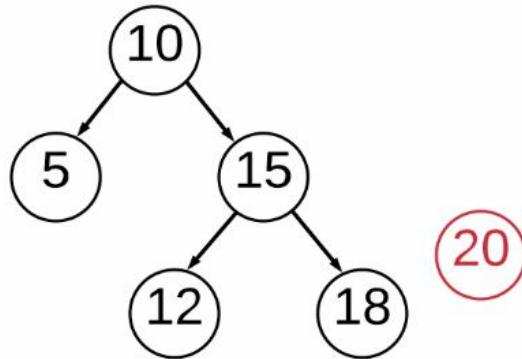


imgflip.com

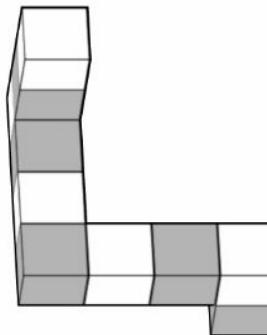


Understanding the Neural Representations of Data Structure Manipulations

- How do human brains represent data structures? Is it more like **text** or more like **3D objects**?



imgflip.com



imgflip.com

Understanding the Neural Representations of Data Structure Manipulations

- 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



Understanding the Neural Representations of Data Structure Manipulations

- **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*



Understanding the Neural Representations of Data Structure Manipulations

- Experimental design: 2 tasks
 - Data structure manipulations
 - List/Array operations
 - Tree operations
 - Mental rotations: 3D objects

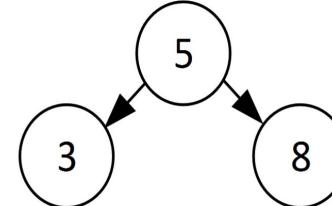
What is the minimum number of swaps required to make the given array sorted?

Indices	0	1	2	3	4	5
nums	0	6	7	4	8	10

A. 1

B. 2

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



A. 5, 3, 8

B. 8, 3, 5



Understanding the Neural Representations of Data Structure Manipulations

- Experimental design: 2 tasks
 - Data structure manipulations
 - List/Array operations
 - Tree operations
 - Mental rotations: 3D objects

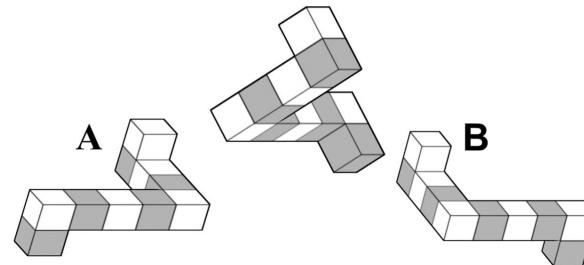
What is the minimum number of swaps required to make the given array sorted?

Indices	0	1	2	3	4	5
nums	0	6	7	4	8	10

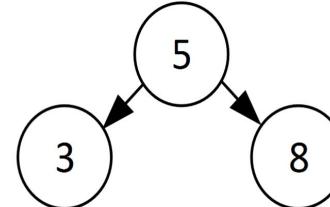
A. 1

B. 2

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



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



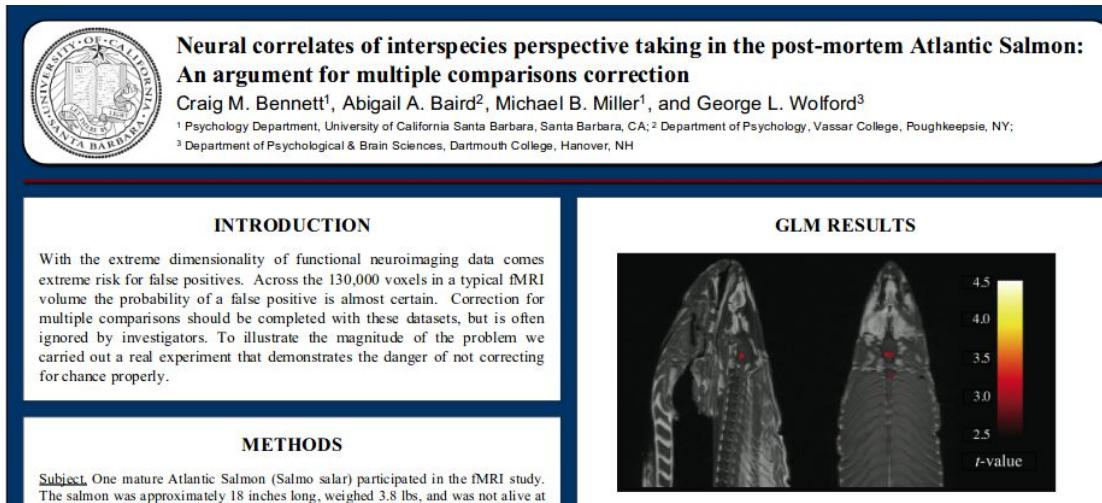
A. 5, 3, 8

B. 8, 3, 5



Understanding the Neural Representations of Data Structure Manipulations

- Data analysis: we need to be careful
 - Spurious correlations due to multiple comparison



Understanding the Neural Representations of Data Structure Manipulations

- Data analysis: we need to be careful
- fMRI and fNIRS use the same high-level 3-step analysis approach
 - False discovery rate correction for multiple comparisons (FDR)



Understanding the Neural Representations of Data Structure Manipulations

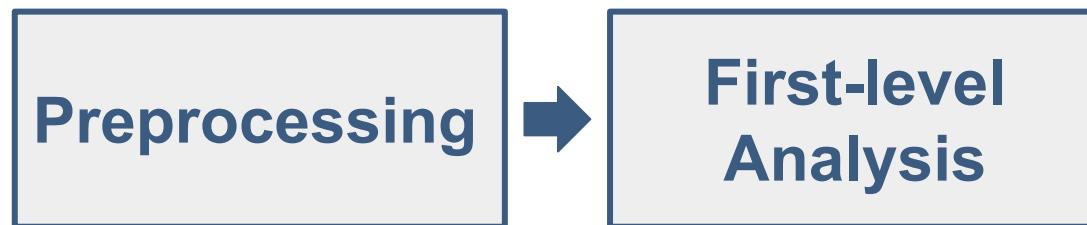
- Data analysis: we need to be careful
- fMRI and fNIRS use the same high-level 3-step analysis approach
 - False discovery rate correction for multiple comparisons (FDR)

Preprocessing



Understanding the Neural Representations of Data Structure Manipulations

- Data analysis: we need to be careful
- fMRI and fNIRS use the same high-level 3-step analysis approach
 - False discovery rate correction for multiple comparisons (FDR)



Understanding the Neural Representations of Data Structure Manipulations

- Data analysis: we need to be careful
- fMRI and fNIRS use the same high-level 3-step analysis approach
 - False discovery rate correction for multiple comparisons (FDR)



Neural Representations of Data Structures: Metrics

- Following the best practices in medical imaging, we can find significant relationship between data structure manipulations and spatial ability ($p<0.01$).
- We can find significant relationships regarding the difficulty levels of tasks.



Neural Representations of Data Structures: Preliminary Results

- Experiment setup and data

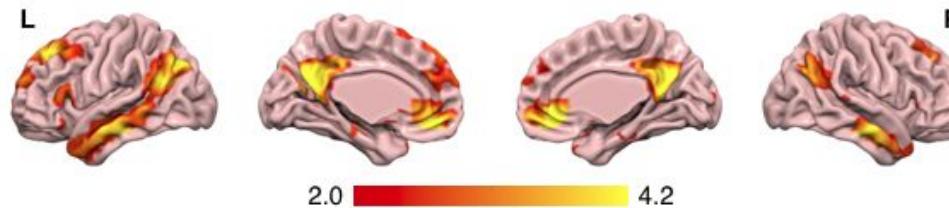
- 76 participants: 70 valid
 - fMRI: 30
 - fNIRS: 40
 - Two hours for each participant: 90 stimuli, qualitative post-survey

De-identified data is public: <https://web.eecs.umich.edu/weimerw/fmri.html>



Neural Representations of Data Structures: Preliminary Results

- Data structure manipulations involve spatial ability
 - fMRI: more similarities than differences ($p<0.01$)
 - fNIRS: activation in the same brain regions ($p<0.01$)



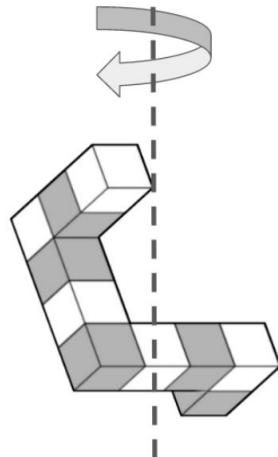
Mental Rotation vs. Tree



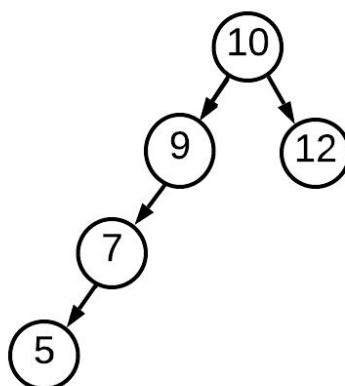
Neural Representations of Data Structures: Preliminary Results

- The brain works even **harder** for **more difficult** data structure tasks
 - Difficulty measurement
 - Mental rotations: angle of rotation
 - Data structure: size

Rotation Angle = 20°



N = 5



N= 5

Indices	0	1	2	3	4
nums	1	5	6	7	10

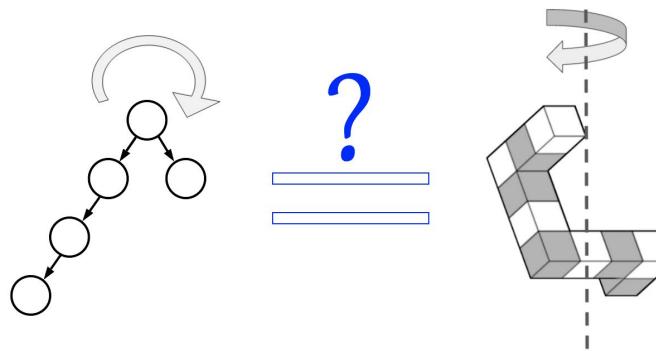


Neural Representations of Data Structures: Preliminary Results

- The brain works even **harder** for **more difficult** data structure tasks
 - Difficulty measurement
 - Mental rotations: angle of rotation
 - Data structure: size
 - **fMRI**: the rate of extra work in your brain is higher for data structure tasks than it is for mental rotation tasks
 - **fNIRS: no significant findings** for the effect of task difficulty

Neural Representations of Data Structures: Preliminary Results

- How Do **Self-reporting** and Neuroimaging Compare?
 - Self-reporting may **not be reliable**
 - Medical imaging found mental rotation and data structure tasks are very similar
 - 70% of human participants believe there is no connection!



Understanding the Neural Representations of Data Structure Manipulations

- Recall: How do human brains represent data structures? Is it more like text or more like 3D objects?

Data structure manipulations and mental rotations (spatial ability) involve very similar brain regions.



Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representations of data structures



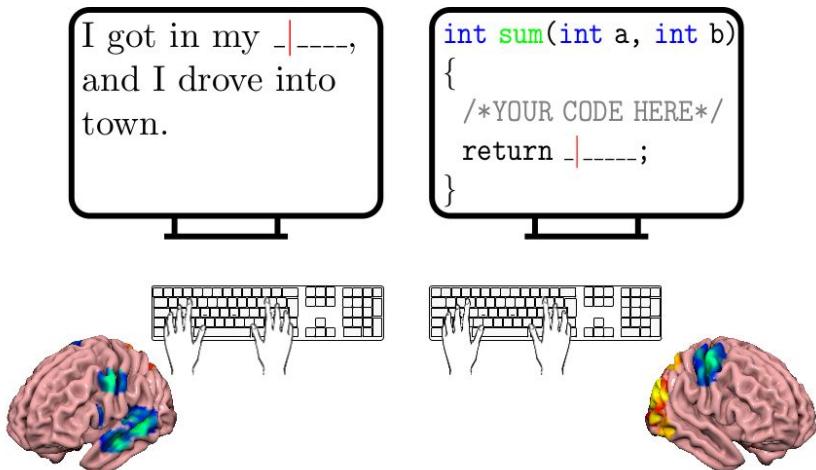
Comparing prose writing and code writing



Understanding bias in code reviews

Comparing Code Writing and Prose Writing

- Are code writing and prose writing **similar** neural activities? Do I have to be **good at English writing** to **become a good software developer?**



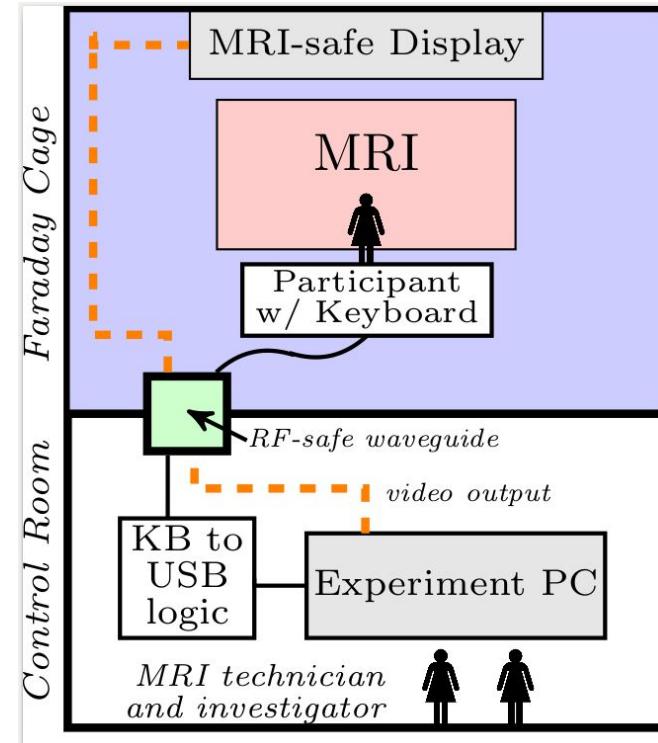
Comparing Code Writing and Prose Writing

- fMRI: penetration power
- Challenges
 - fMRI-safe bespoke keyboard
 - QWERTY keyboard
 - Allow typing and editing
 - Design writing stimuli
 - Prose writing
 - Code writing



Comparing Code Writing and Prose Writing

- fMRI: penetration power
- Challenge: fMRI-safe bespoke keyboard
 - QWERTY keyboard
 - Allow typing and editing



Comparing Code Writing and Prose Writing

- Challenge: Stimuli design
 - Two categories of tasks for **code writing** and **prose writing**
 - Fill in the blank (FITB)

```
/*Complete the sentence
 * such that the sentence
 * makes sense*/
Brian was so fond of
his dog that their
brief █ left him
not just saddened, but
in a state of sorrow.
```

Prose - FITB

```
/*Complete the definition of
 * the function such that it
 * receives an integer parameter
 * and returns the absolute
 * value of the parameter.*/
int absoluteValue(int num1)
{
    /* YOUR CODE HERE */
    return absoluteValue;
}
```

Code - FITB



Comparing Code Writing and Prose Writing

- Challenge: Stimuli design

- Two categories of tasks for **code writing** and **prose writing**
- Fill in the blank (FITB)
- Long response (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 |
```

Prose - LR

```
1 |
```

Code - LR



Comparing Code Writing and Prose Writing

- Experimental design: 2 categories of tasks for **code writing** and **prose writing**
 - Code writing tasks: [Turing's Craft](#)
 - Prose writing tasks: **SAT**



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.



Code Writing vs. Prose Writing: Metrics

- We can have a bespoke QWERTY keyboard that can safely work in fMRI machine
- We can find significant relationship between code writing and prose writing ($p<0.01$)
 - General relationship
 - Relationship between different types of tasks (i.e., FITB and LR)



Code Writing vs. Prose Writing: Preliminary Results

- **IRB approved**
- Bespoke keyboard
 - **Finished** deployment and **passed safety tests**
- Data collection is **done**
 - 30 participants
 - Two hours for each participant: 52 stimuli
 - For both code writing and prose writing:
 - FITB: 17
 - LR: 9



Proposal Overview: Four Components



Monitoring mental health using mobile crowdsensing



Understanding the neural representations of data structures



Comparing prose writing and code writing



Understanding bias in code reviews

Understanding Bias in Code Reviews

-

Code reviews

- The systematic inspection, analysis, evaluation, and revision of code.
- The latent defect discovery rate of formal code review can be 60%-65%.

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
  @@ -32,7 +32,7 @@ int majorityElement(vector<int> &num) {
    32      32          cnt++;
    33      33          }else{
    34      34              majority == num[i] ? cnt++ : cnt --;
  35      -          if (cnt >= num.size()/2) return majority;
  35      +          if (cnt > num.size()/2) return majority;
    36      36      }
    37      37      }
  38      38      return majority;
  @@
```



Understanding Bias in Code Reviews

- Code reviews
 - The systematic inspection, analysis, evaluation, and revision of code.
 - The latent defect discovery rate of formal code review can be 60%-65%.



Russell Keith-Magee
freakboy3742

Sponsor

Overview **Repositories 41** Projects 0 Stars 11 Followers 1k Following 0

Find a repository...

beeware

Forked from beeware/beeware
A simplified command line user interface to the BeeWare suite.
Python 12 Other Updated 2 days ago

briefcase

Forked from beeware/briefcase
Tools to support converting a Python project into a standalone native application.



Understanding Bias in Code Reviews

- Code reviews
 - The systematic inspection, analysis, evaluation, and revision of code.
 - The latent defect discovery rate of formal code review can be 60%-65%.
- Bias in code reviews
 - Code source
 - Gender



Russell Keith-Magee
freakboy3742

[Sponsor](#)

Overview **Repositories 41** Projects 0 Stars 11 Followers 1k Following 0

Find a repository...

beeware

Forked from beeware/beeware
A simplified command line user interface to the BeeWare suite.
Python 12 Other Updated 2 days ago

briefcase

Forked from beeware/briefcase
Tools to support converting a Python project into a standalone native application.



Understanding Bias in Code Reviews

- Code reviews
 - The systematic inspection, analysis, evaluation, and revision of code.
 - The latent defect discovery rate of formal code review can be 60%-65%.
- Bias in code reviews
 - Code source
 - Gender
 - Automated software repair tools



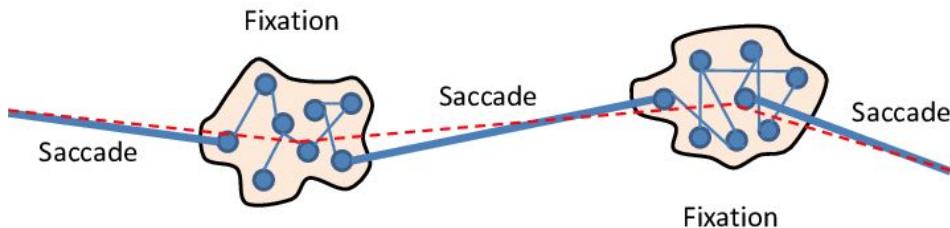
Understanding Bias in Code Reviews

- How does **author information** affect software developers' **decision making** in code reviews?
- Do software developers have **gender bias** in code reviews?
- Do software developers have **bias against machine-generated code patches**?



Understanding Bias in Code Reviews

- Neural activities in code reviews: fMRI
- Visual focus in code reviews: eye tracking
 - Fixations and saccades
 - Attention over different Area of Interests (AOI)
 - Comment
 - Code changes
 - Author information

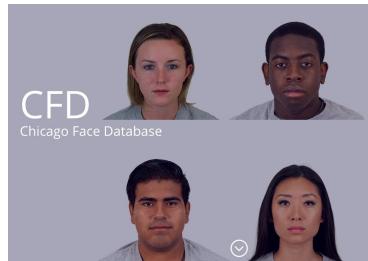


```
import java.util.Scanner;  
  
public class eyeTrack2 {  
  
    public static void main(String[] args)  
    {  
        Scanner in = new Scanner(System.in);  
        System.out.print("First number: ");  
        int num1 = in.nextInt();  
        System.out.print("Second number: ");  
        int num2 = in.nextInt();  
        int average = (num1 + num2) / 2;  
        System.out.print("Average: ");  
        System.out.print(average);  
    }  
}
```



Understanding Bias in Code Reviews

- Stimuli design
 - Pull requests from real world open source C and C++ projects (e.g., GitHub)
 - **Relabel** the author information
 - Pictures from **Chicago Face Database**
 - Controlling age, race, attractiveness and facial expressions
 - Avatar picture to represent automated software repair tools



Understanding Bias in Code Reviews

- Stimuli design
 - Pull requests from real world open source projects (C and C++) (e.g., GitHub)
 - Relabel the author information
 - Pictures from **Chicago Face Database**
 - Controlling age, race, attractiveness and facial expressions
 - Avatar picture to represent automated software repair tools
 - We will not tell the participants about the relabeling and the purpose of investigating the author bias in code reviews.
 - Avoid social desirability bias



Understanding Bias in Code Reviews

- Stimuli design
 - Simulating a real-world code review interface

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

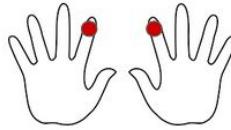
Commit message

2 algorithms/cpp/majorityElement/majorityElement.cpp

```
2 32    @@ -32,7 +32,7 @@ int majorityElement(vector<int> &num) {
32      32          cnt++;
33      33          }else{
34      34              majority == num[i] ? cnt++ : cnt--;
35      35      -          if (cnt >= num.size()/2) return majority;
36      35      +          if (cnt > num.size()/2) return majority;
37      36          }
38      37          }
39      38      return majority;
```



Owner: [REDACTED]



Accept Reject



Understanding Bias in Code Reviews

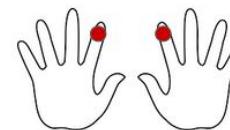
- Stimuli design
 - Simulating a real-world code review interface

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

Code changes

```
2 algorithms/cpp/majorityElement/majorityElement.cpp
  @@ -32,7 +32,7 @@ int majorityElement(vector<int> &num) {
    32      32          cnt++;
    33      33      }else{
    34      34          majority == num[i] ? cnt++ : cnt--;
  35      -          if (cnt >= num.size()/2) return majority;
  35      +          if (cnt > num.size()/2) return majority;
    36      36      }
    37      37      }
    38      38      return majority;
```

Owner: [REDACTED]



Accept Reject



Understanding Bias in Code Reviews

- Stimuli design
 - Simulating a real-world code review interface

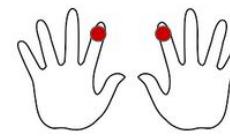
Delete the equal mark in case the array is like
 $\{x,x,x\dots(n),y,y,y,y\dots(n+1)\}$

Author image

2 algorithms/cpp/majorityElement/majorityElement.cpp

```
2 #include <iostream>
3 #include <vector>
4
5 int majorityElement(vector<int> &num) {
6     int cnt = 0;
7     int majority = num[0];
8
9     for (int i = 1; i < num.size(); i++) {
10        if (num[i] == majority) {
11            cnt++;
12        } else {
13            cnt--;
14        }
15        if (cnt >= num.size() / 2) {
16            majority = num[i];
17        }
18    }
19
20    return majority;
21}
```

Owner: [REDACTED]



Accept Reject



Bias in Code Reviews: Metrics

- We are able to involve author deception in the stimuli design (IRB permission)
- We are able to recruit approximately gender-balanced group of participants
- We are able to obtain significant relationship between the brain activities of code reviews with different author information ($p<0.01$)
- We are able to observe significant similarities or differences of the visual focus and strategies for code reviews with different author information ($p<0.01$)

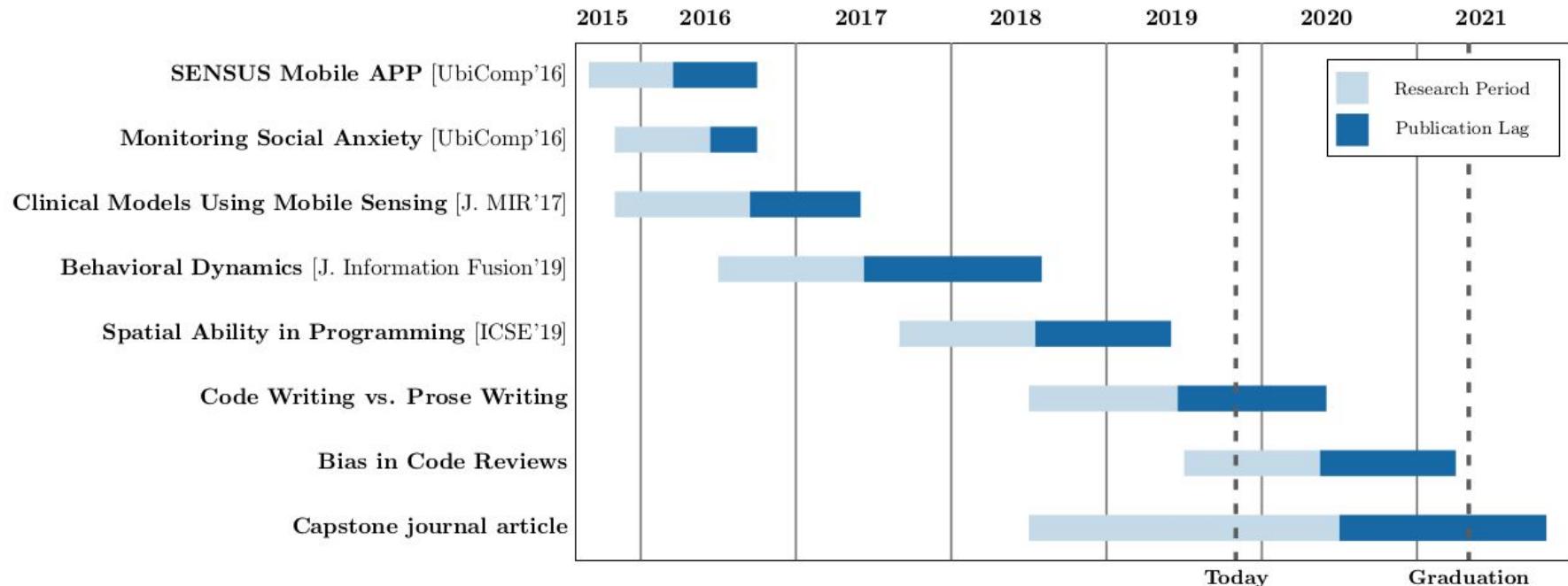


Bias in Code Reviews: Preliminary Results

- **Stimuli design is done**
 - Two sets of stimuli: 60 stimuli each
 - Randomly assign author pictures into three groups
 - 20 men
 - 20 women
 - 20 machine
 - Relabel each set with different code-author combinations
 - Control code quality
- **IRB approved**
- The fMRI lab has a **built-in eye tracker**
- fMRI lab pilot grant to support this study



Ph.D. Timeline



Publications: Supporting this Proposal

1. **Distilling Neural Representations of Data Structure Manipulation using fMRI and fNIRS.** Yu Huang, Xinyu Liu, Ryan Krueger, Tyler Santander, Xiaosu Hu, Kevin Leach, Westley Weimer. *41st ACM/IEEE International Conference on Software Engineering (ICSE 2019)*. *Distinguished Paper Award*
2. **Understanding Behavioral Dynamics of Social Anxiety Among College Students Through Smartphone Sensors.** Jiaqi Gong, Yu Huang, Philip I Chow, Karl Fua, Matthew Gerber, Bethany Teachman, Laura Barnes. *Information Fusion*, 49:57–68, September 2019.
3. **Discovery of Behavioral Markers of Social Anxiety From Smartphone Sensor Data.** Yu Huang, Jiaqi Gong, Mark Rucker, Philip Chow, Karl Fua, Matthew S. Gerber, Bethany Teachman, and Laura E. Barnes. *The 1st Workshop on Digital Biomarkers, DigitalBiomarkers '17*, pages 9–14, New York, NY, USA, ACM.
4. **Using Mobile Sensing to Test Clinical Models of Depression, Social Anxiety, State Affect, and Social Isolation Among College Students.** Philip I. Chow, Karl Fua, Yu Huang, Wesley Bonelli, Haoyi Xiong, Laura E. Barnes, and Bethany Teachman. *J Med Internet Res*, 19(3):e62, Mar 2017.
5. **Assessing Social Anxiety Using GPS Trajectories and Point-of-Interest Data.** Yu Huang, Haoyi Xiong, Kevin Leach, Yuyan Zhang, Philip Chow, Karl Fua, Bethany A Teachman, and Laura E Barnes. *In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 898–903.
6. **Sensus: a Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies.** Haoyi Xiong, Yu Huang, Laura E Barnes, and Matthew S Gerber. *In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16*, pages 415–426.
7. **Demons: an Integrated Framework for Examining Associations Between Physiology and Selfreported affect Tied to Depressive Symptoms.** Philip Chow, Wesley Bonelli, Yu Huang, Karl Fua, Bethany A Teachman, and Laura E Barnes. *In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 1139–1143.

Publications: Others

8. **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.
9. **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 .
10. **Monitoring Social Anxiety From Mobility and Communication Patterns.** Mehdi Boukhechba, Yu Huang, Philip Chow, Karl Fua, Bethany A. Teachman, and Laura E.Barnes. *The ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, UbiComp '17*, pages 749–753.
11. **Daehr: A Discriminant Analysis Framework for Electronic Health Record Data and an Application to Early Detection of Mental Health Disorders.** Haoyi Xiong, Jinghe Zhang, Yu Huang, Kevin Leach, and Laura E. Barnes. *ACM Trans. Intell. Syst. Technol.*, 8(3):47:1–47:21, February 2017.
12. **A Design and Theoretical Analysis of a 145 mV to 1.2 V Single-Ended Level Converter Circuit for Ultra-Low Power Low Voltage ICs.** Yu Huang, Aatmesh Shrivastava, Laura E Barnes, and Benton H Calhoun. *Journal of Low Power Electronics and Applications*, 6(3):11, 2016.
13. **M-SEQ: Early Detection of Anxiety and Depression via Temporal Orders of Diagnoses in Electronic Health Data.** Jinghe Zhang, Haoyi Xiong, Yu Huang, Hao Wu, Kevin Leach, and Laura Barnes. *In Proceedings of the 2015 IEEE International Conference on Big Data (BigData 2015)*, September 2015.
14. **A 145 mV to 1.2 V Single Ended Level Converter Circuit for Ultra-Low Power Low Voltage ICs.** Yu Huang, Aatmesh Shrivastava, and Benton H Calhoun. *In SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), 2015 IEEE*, pages 1–3.

Publications: Others

15. **Optimizing Energy Efficient Low Swing Interconnect for Sub-Threshold FPGAs.** He Qi, Oluseyi Ayorinde, Yu Huang, and Benton Calhoun. *In Field Programmable Logic and Applications (FPL), 2015 25th International Conference on*, pages 1–4. IEEE, 2015.
16. **Using Island-Style Bi-directional Intra-CLB Routing in Low-Power FPGAs.** Oluseyi Ayorinde, He Qi, Yu Huang, and Benton H Calhoun. *In Field Programmable Logic and Applications (FPL), 2015 25th International Conference on*, pages 1–7. IEEE, 2015.

Broader Impact

- All the medical imaging and behavioral data will be de-identified and released publicly
- **Sensus** has been released and can be used in a wide range of human-subject studies
- Our research findings can help psychologists monitor mental health status and help computer science educators develop efficient training strategies
- Our studies provide guidelines for future study design and implementation in the community

Proposal Summary: Four Components

- Monitoring mental health using mobile crowdsensing
 - Sensus: Cross-platform, general MCS mobile application for human-subject studies
 - Understanding human behaviors and mental health status via MCS
- Understanding the neural representation of data structures
- Comparing prose writing and code writing
- Understanding bias in code reviews



Bonus Slides

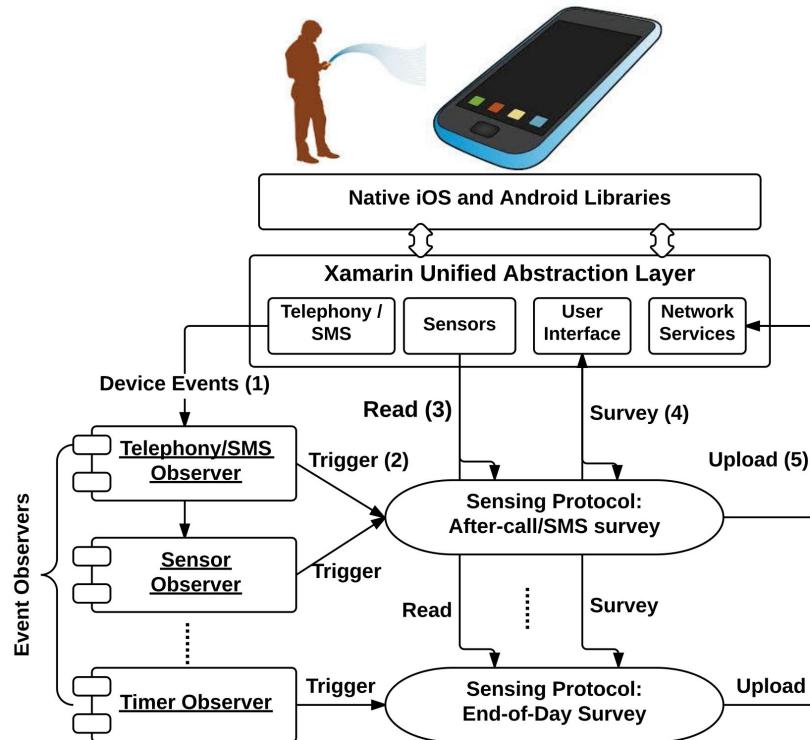


Break Down the Title

- A **fundamental understanding** of computational activities is essential to improving **productivity** and **efficiency** in software engineering
 - Self-reporting
 - Pedagogy
 - Technology transfer
 - Programming expertise

Architecture of Sensus: Mobile Runtime

- The Sensus Mobile runtime
 - Control the underlying device (probes) according to the study protocol
 - Administers researcher-designed surveys
 - Anonymizes data to varying degrees



Sensus: An Example Case

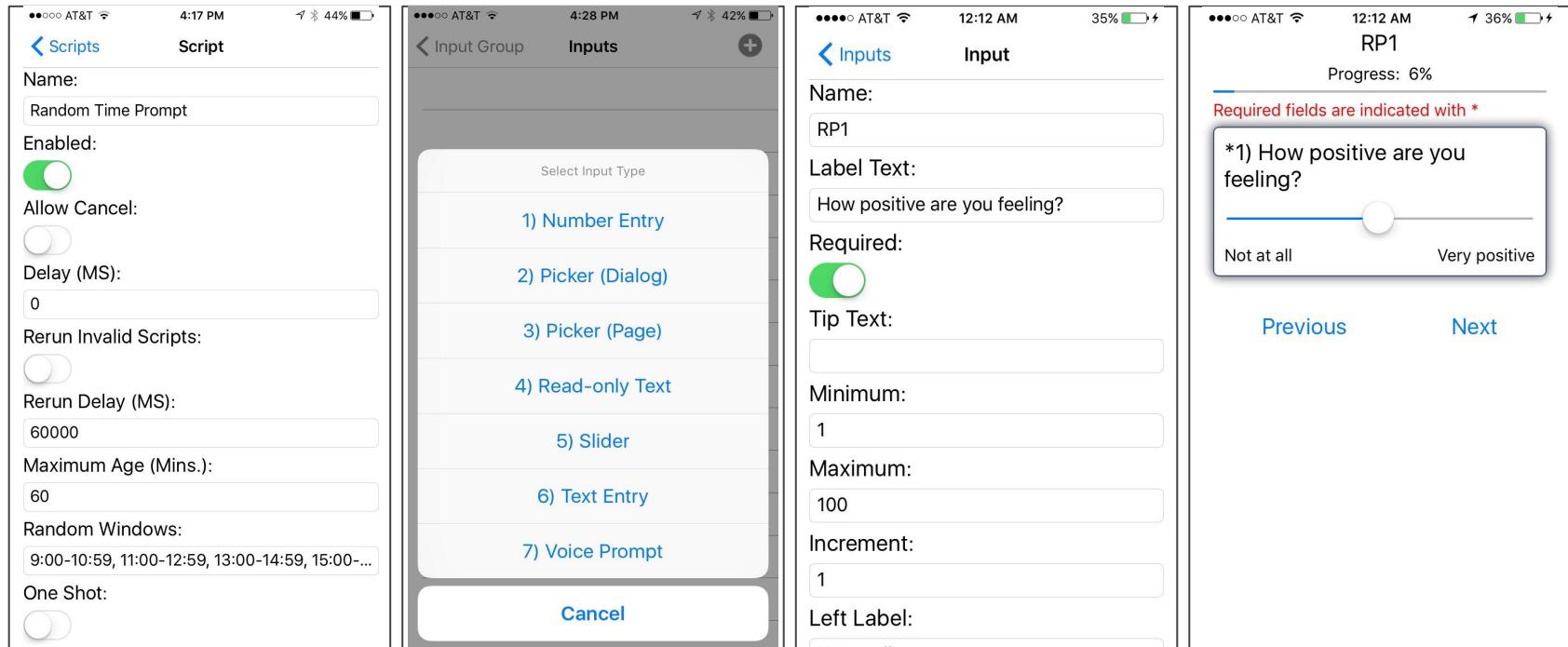
1. Researchers use Sensus to configure a protocol (protocol editor)

The figure consists of four screenshots of the Sensus mobile application interface, illustrating the configuration of a protocol:

- Screenshot 1: Your Sensus Studies Protocol**
 - Name: SALMON
 - Shareable:
 - Description: This is a protocol for the SALMON research study.
 - Participation Horizon (Days): 10
 - Contact Email: kcf3st@virginia.edu
 - Groupable:
 - Force Reports to Remote:
 - Reward Threshold: 0.5
- Screenshot 2: Protocol - Remote Data Store**
 - Bucket: salmon-test
 - Commit Delay (MS): 1800000
 - Folder: 1000
 - Cognito Pool Id: us-east-1:f017a2f5-cf5e-4911-b433-18cc84...
 - Require WiFi:
 - Require Charging:
- Screenshot 3: Protocol - Probes**
 - All Probes
 - Acceleration (Listening)
 - Battery Level (Polling)
 - Compass Heading (Listening)
 - Facebook Profile (Polling)
 - GPS Location (Polling)
 - GPS Location (Listening)
 - Phone Call Metadata (Polling)
 - Points of Interest Proximity (Polling)
 - Points of Interest Proximity (Listening)
 - Scripted Interactions
 - Sound Level (Polling)
 - Speed (Polling)
- Screenshot 4: Probes - Probe**
 - GPS Location (Listening)
 - Enabled:
 - Store Data:
 - Max Data / Second: 1
 - Anonymization
 - Device ID: Do Not Anonymize
 - Latitude: Round to Tenths
 - Longitude: Round to Tenths
 - Timestamp: Anonymous Timeline

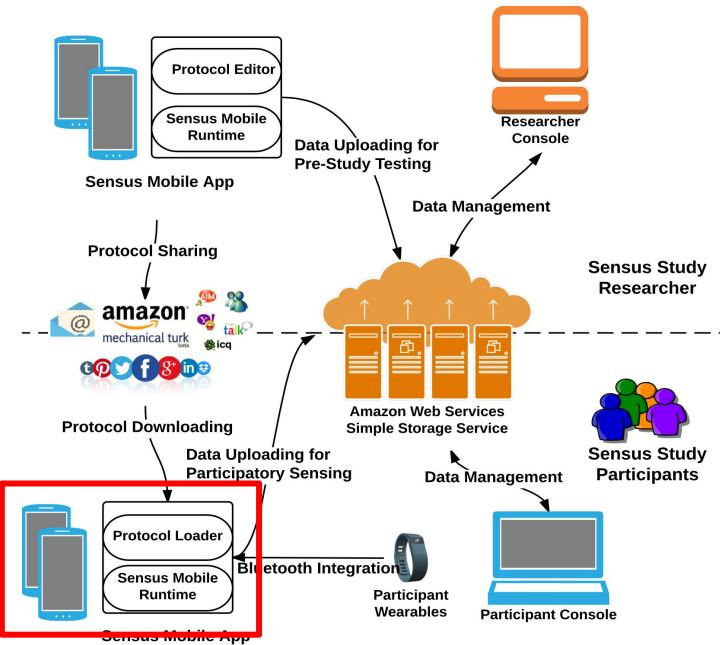
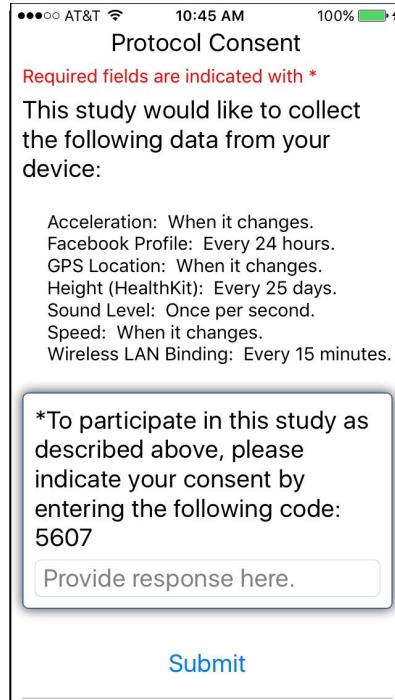
Sensus: An Example Case

1. Researchers use Sensus to configure a protocol (protocol editor)



Sensus: An Example Case

2. Participants load the protocol and start the participation (protocol loader)



Neural Representations of Data Structures: Preliminary Results

- Implications: fMRI vs. fNIRS for software engineering

	fMRI	fNIRS
Time	~2 hours	~2 hours
Penetration Power	Strong	Moderate
Cost	> \$20,000 for 36	~\$2000 for 40
Environment	Restricted	Free
Task Accuracy	Lower (85%, $p < 0.01$)	Higher (92%, $p < 0.01$)
Effort	Light	Heavy
Recruitment	Easy	Moderate (hair)

