The ToMCAT Dataset

https://tomcat.ivilab.org

Adarsh Pyarelal

NeurIPS 2023 Datasets and Benchmarks Track

THE UNIVERSITY OF ARIZONA

Contributions

Contributions

Multimodal human-machine teaming dataset
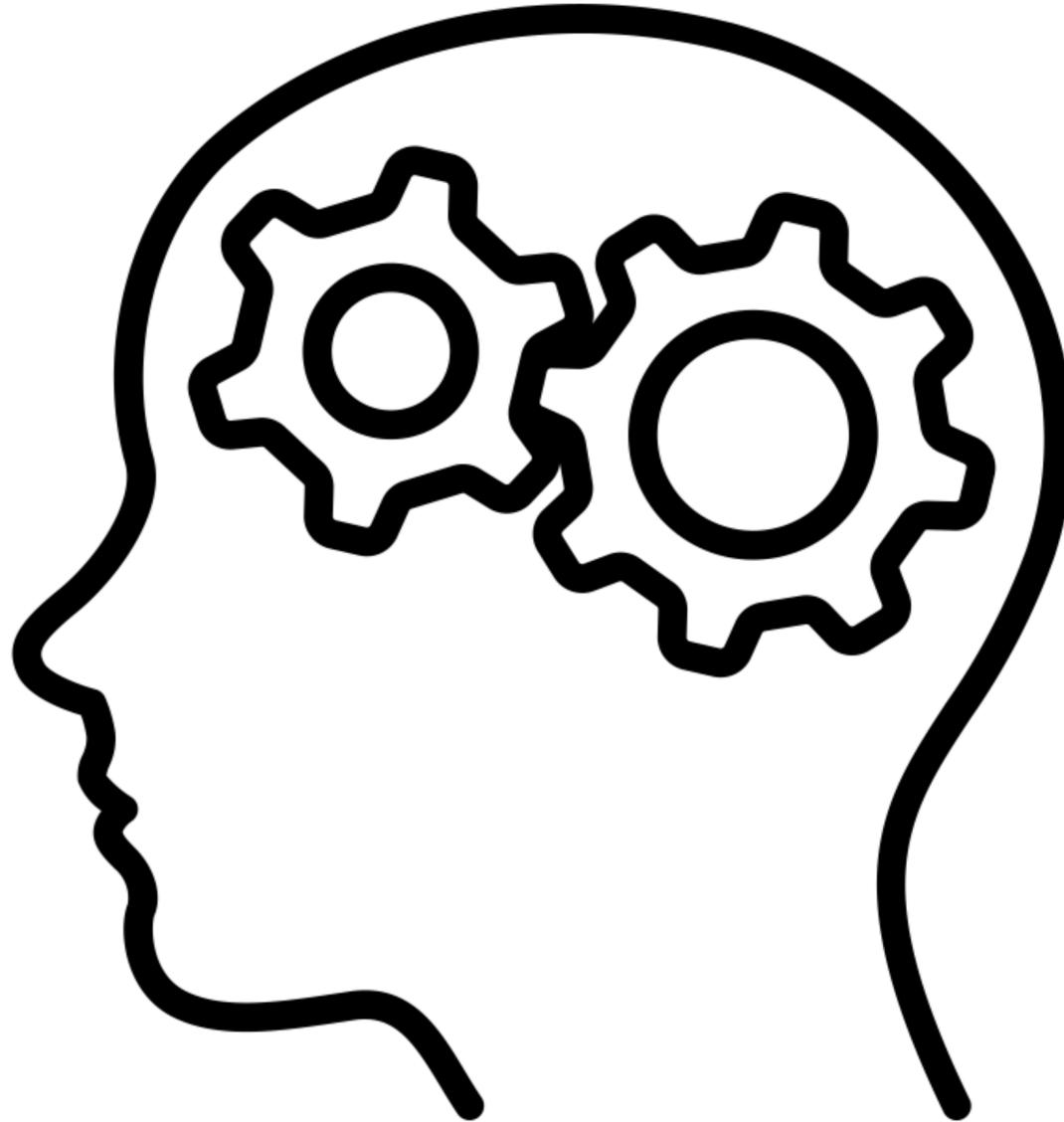
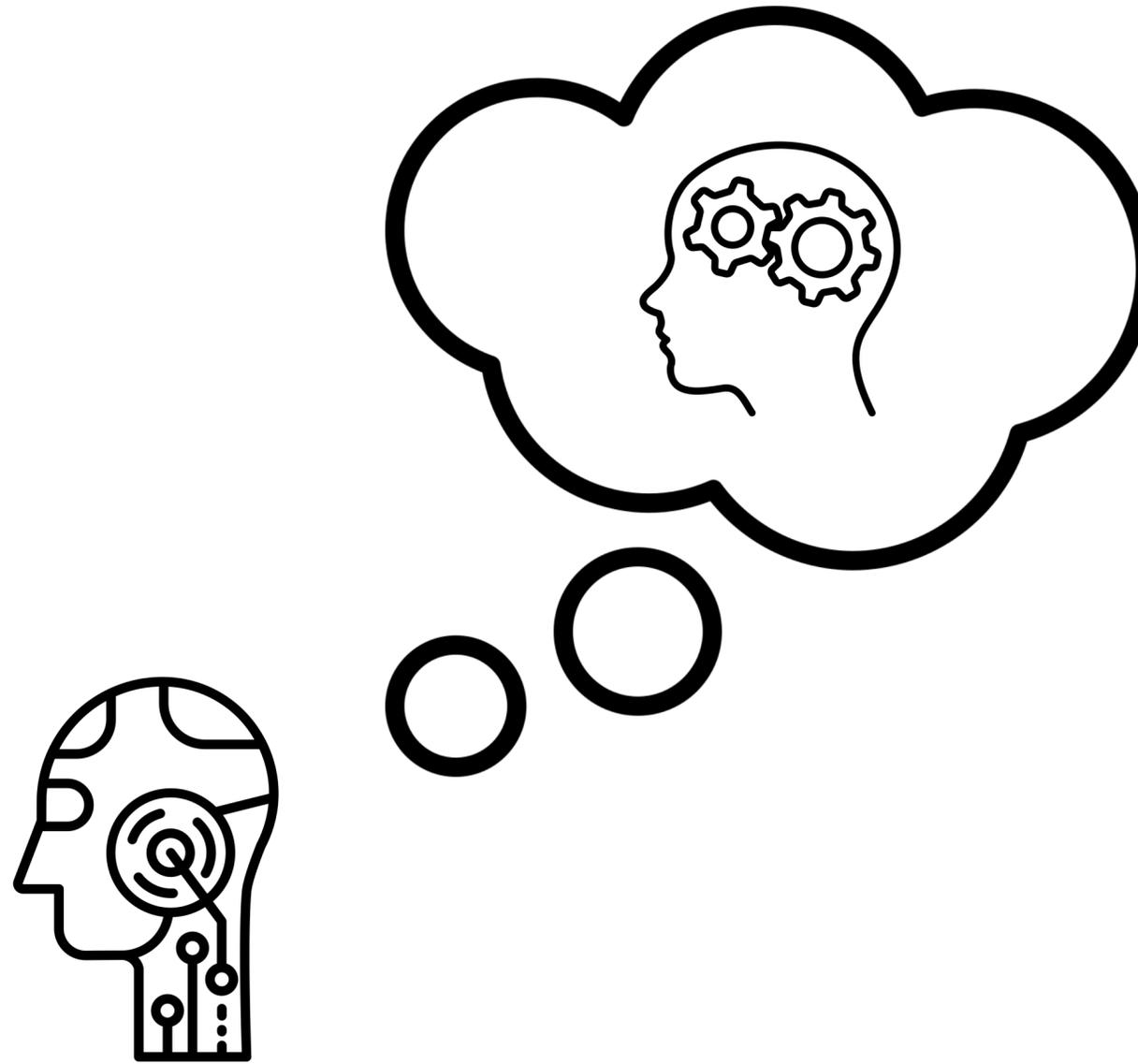# Contributions

Multimodal human-machine teaming dataset

Illustrative experiments using brain data

- Predicting affect from brain scan data

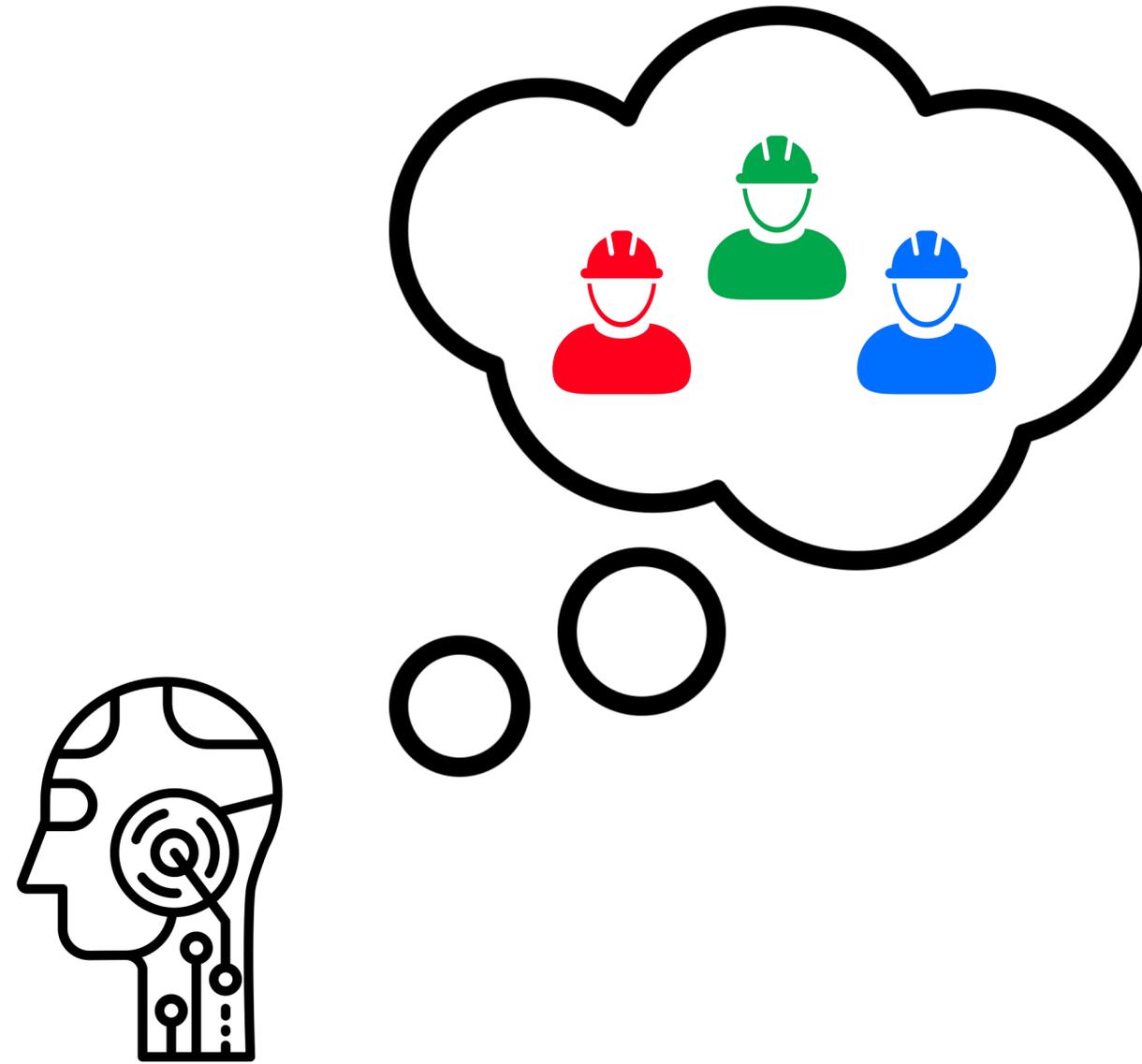- Linking temporal correlation of brain signals with team performance

Theory of Mind

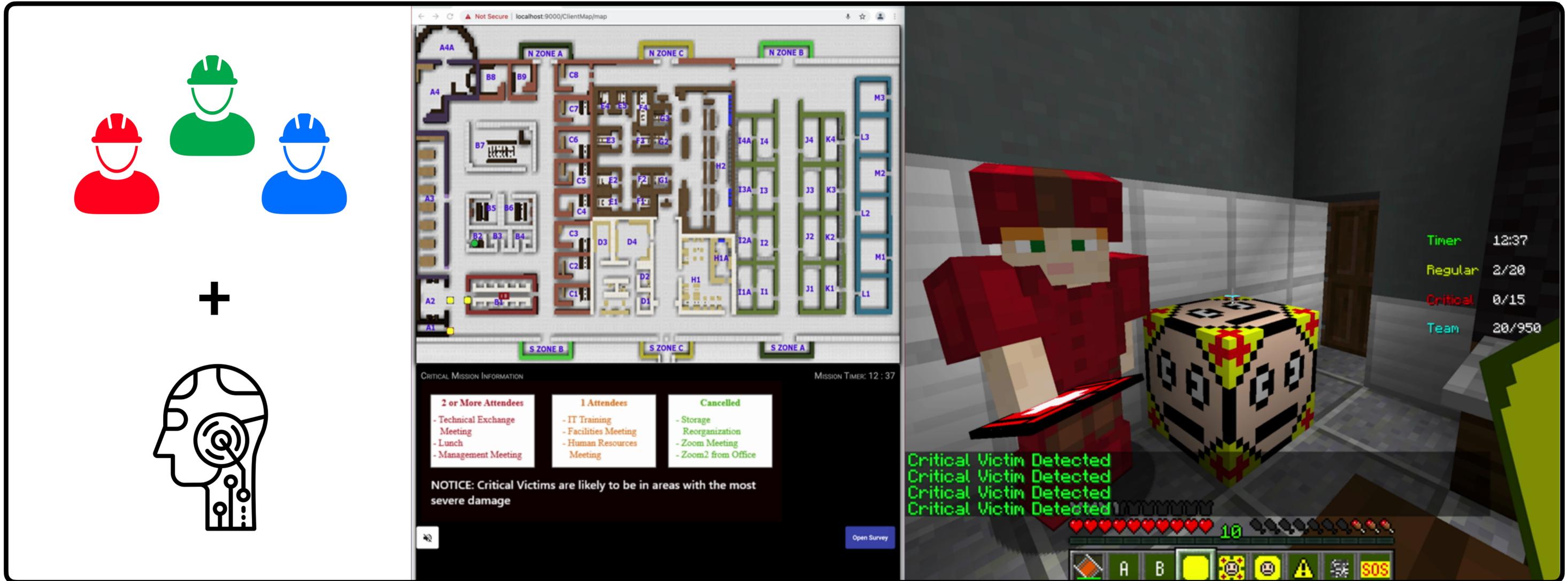# Machine Theory of Mind

Machine Theory of Teams

Teams of humans…
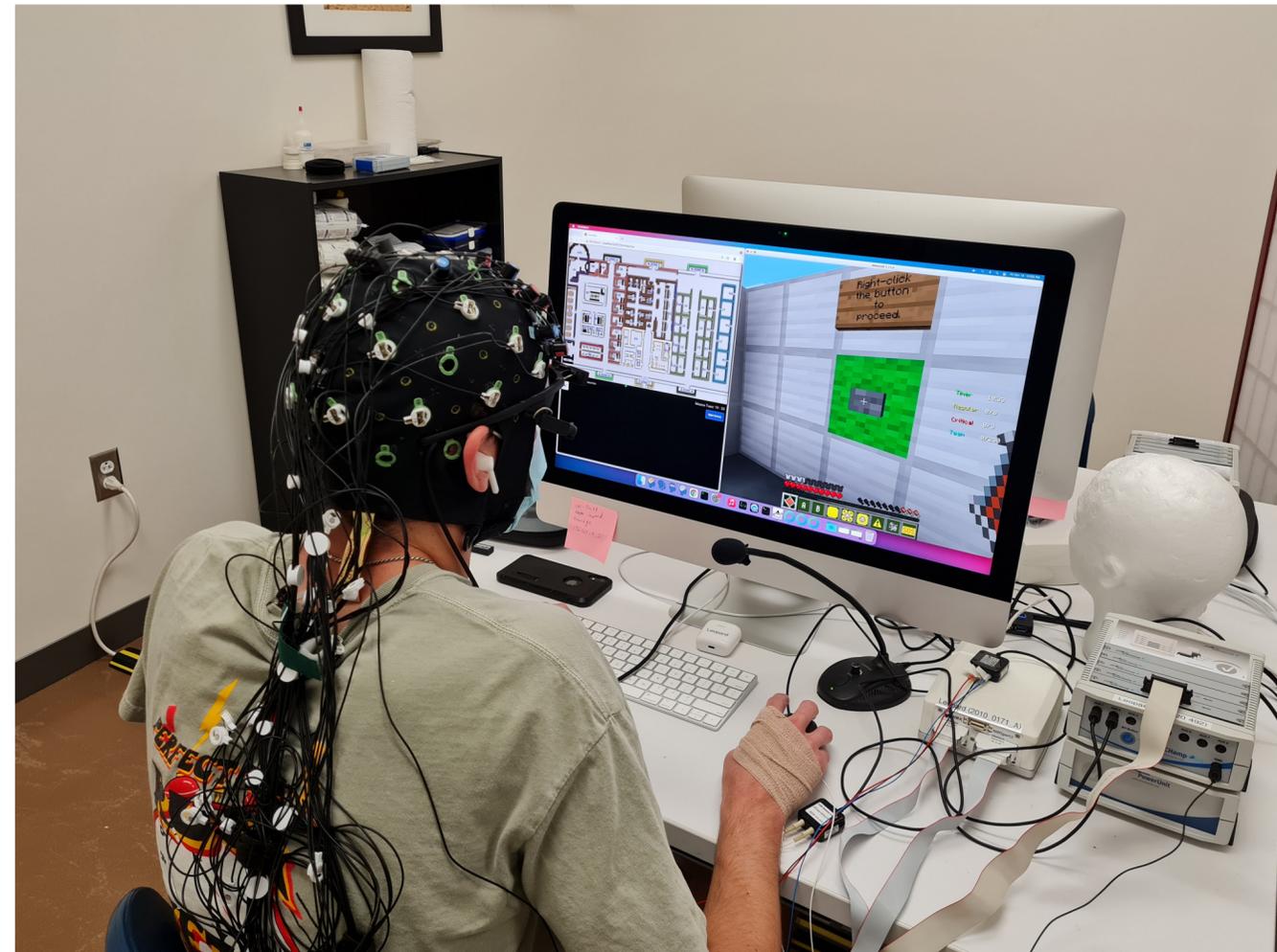
...working with an AI agent...

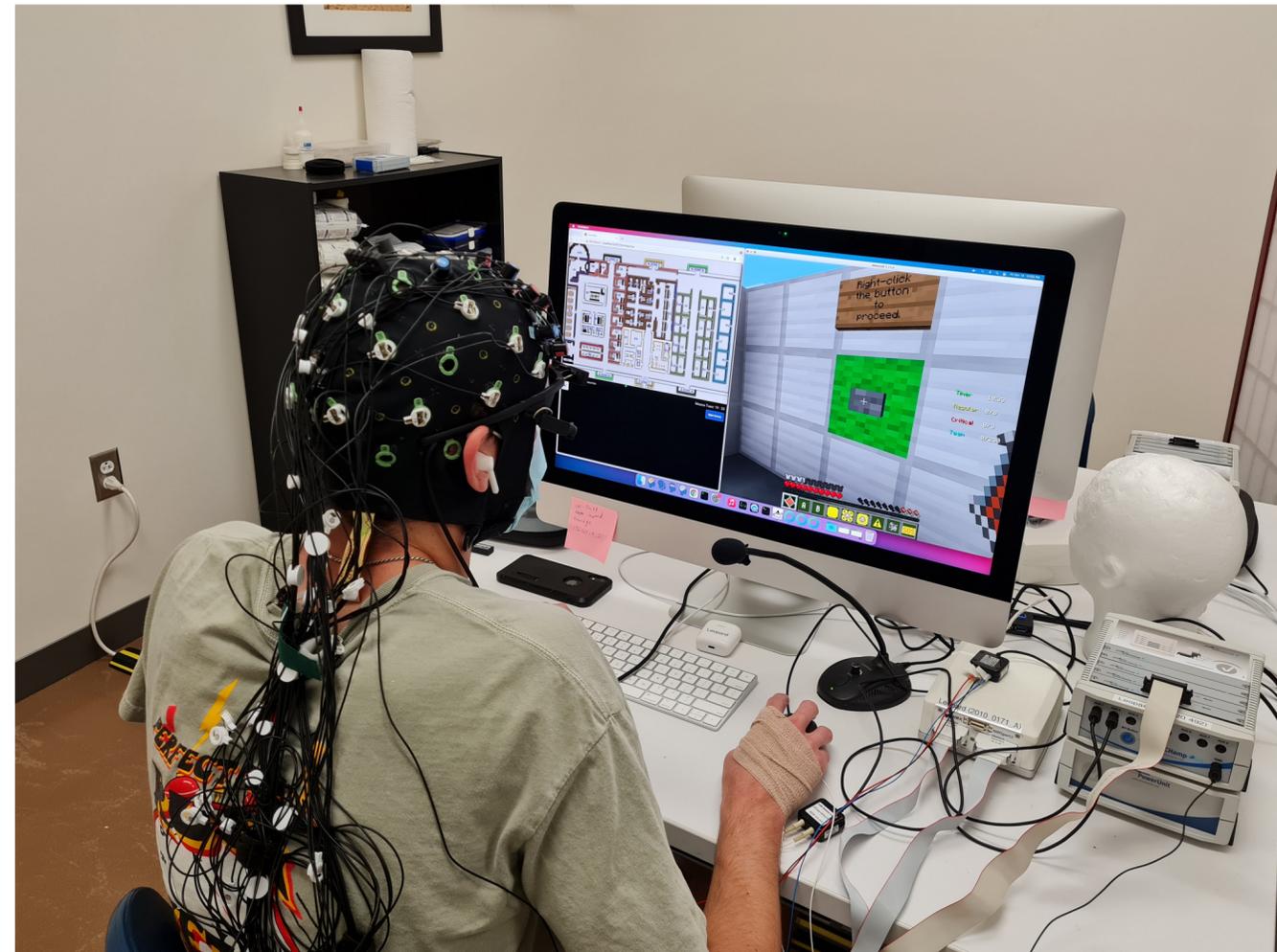+

...in a Minecraft-based testbed*

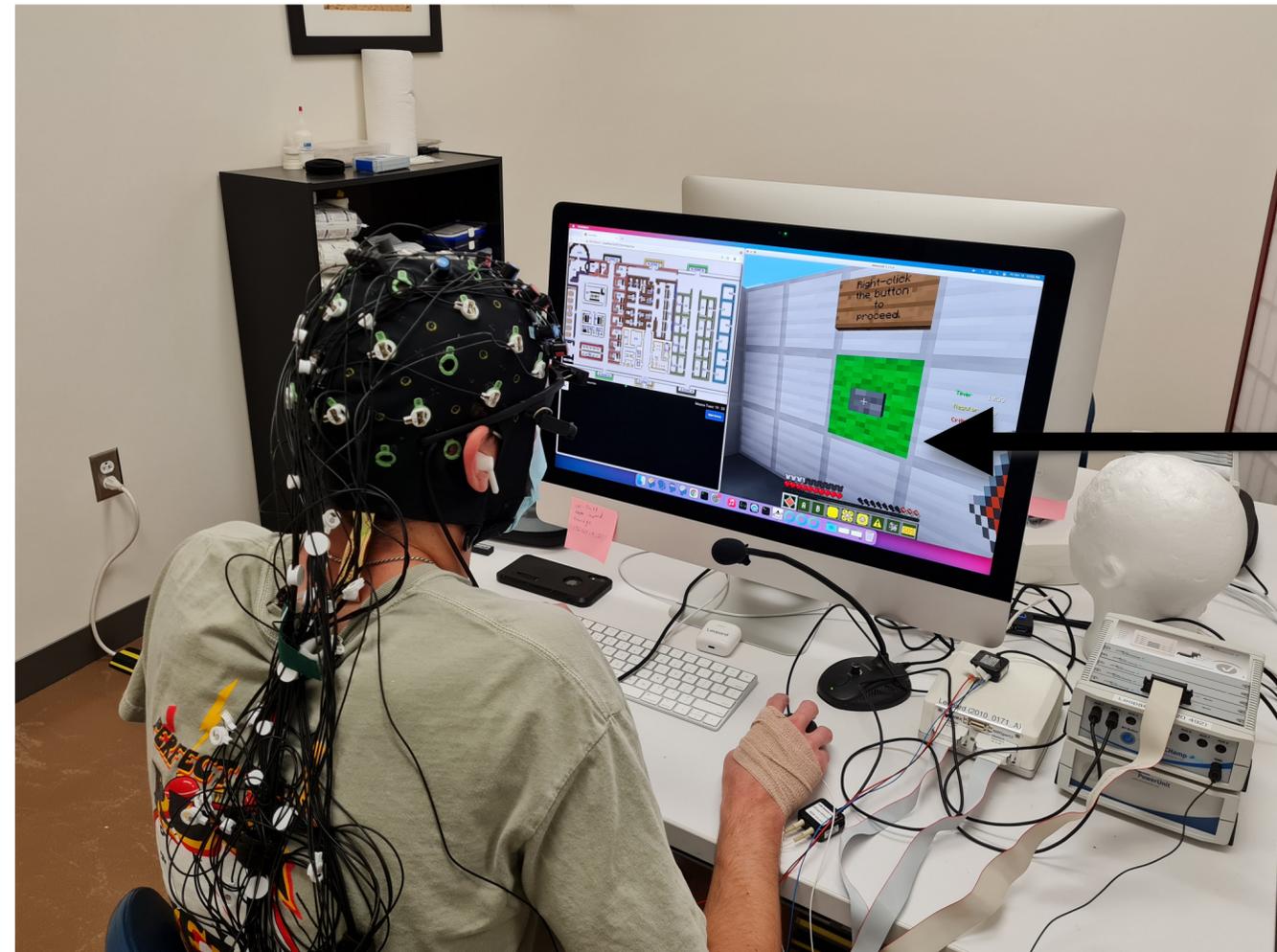# Modalities in the ToMCAT dataset
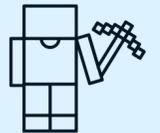
Modalities in the ToMCAT dataset
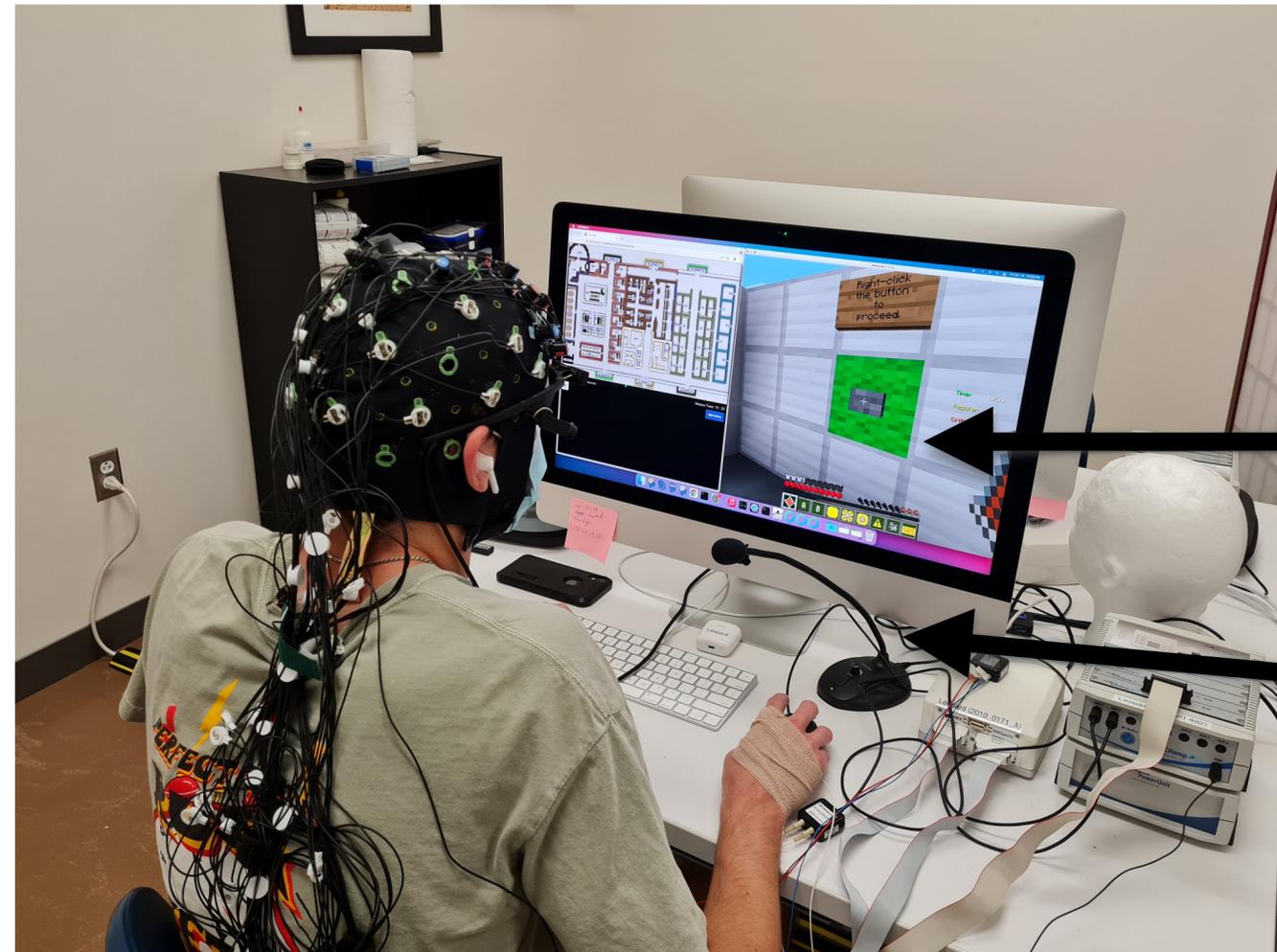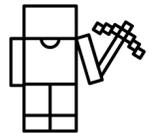

Questionnaires

Modalities in the ToMCAT dataset

Gameplay data

Questionnaires

Modalities in the ToMCAT dataset

Gameplay data

Audio
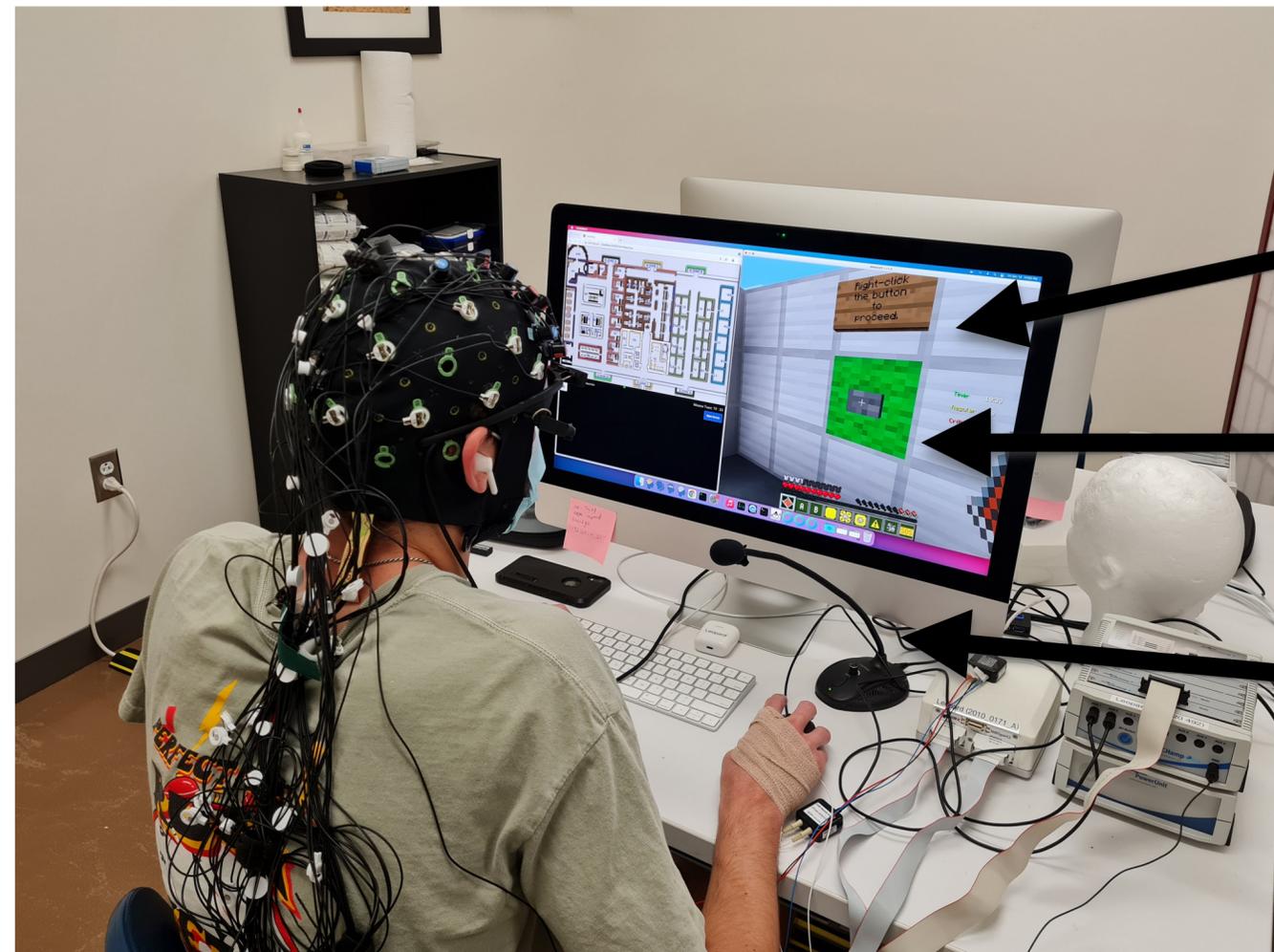
Questionnaires

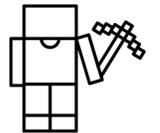# Modalities in the ToMCAT dataset

Modalities in the ToMCAT dataset
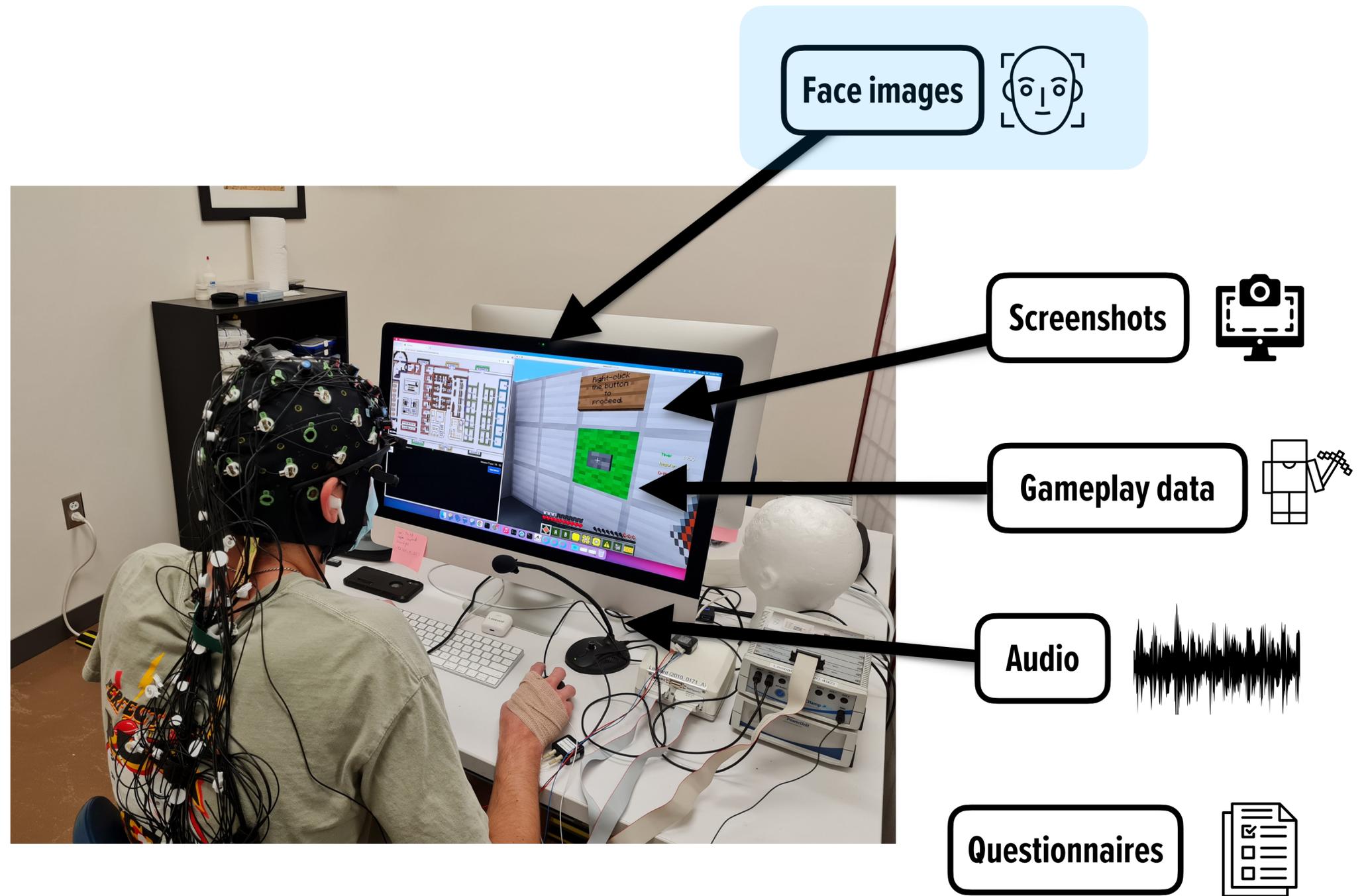

Face images

Screenshots

Gameplay data

Audio

Questionnaires

Modalities in the ToMCAT dataset

Modalities in the ToMCAT dataset

Gaze

Face images

EEG

fNIRS

Screenshots

Gameplay data

Audio

Questionnaires

Modalities in the ToMCAT dataset

Gaze

Face images

EEG

fNIRS

Skin conductance

Screenshots

Gameplay data

Audio

Questionnaires

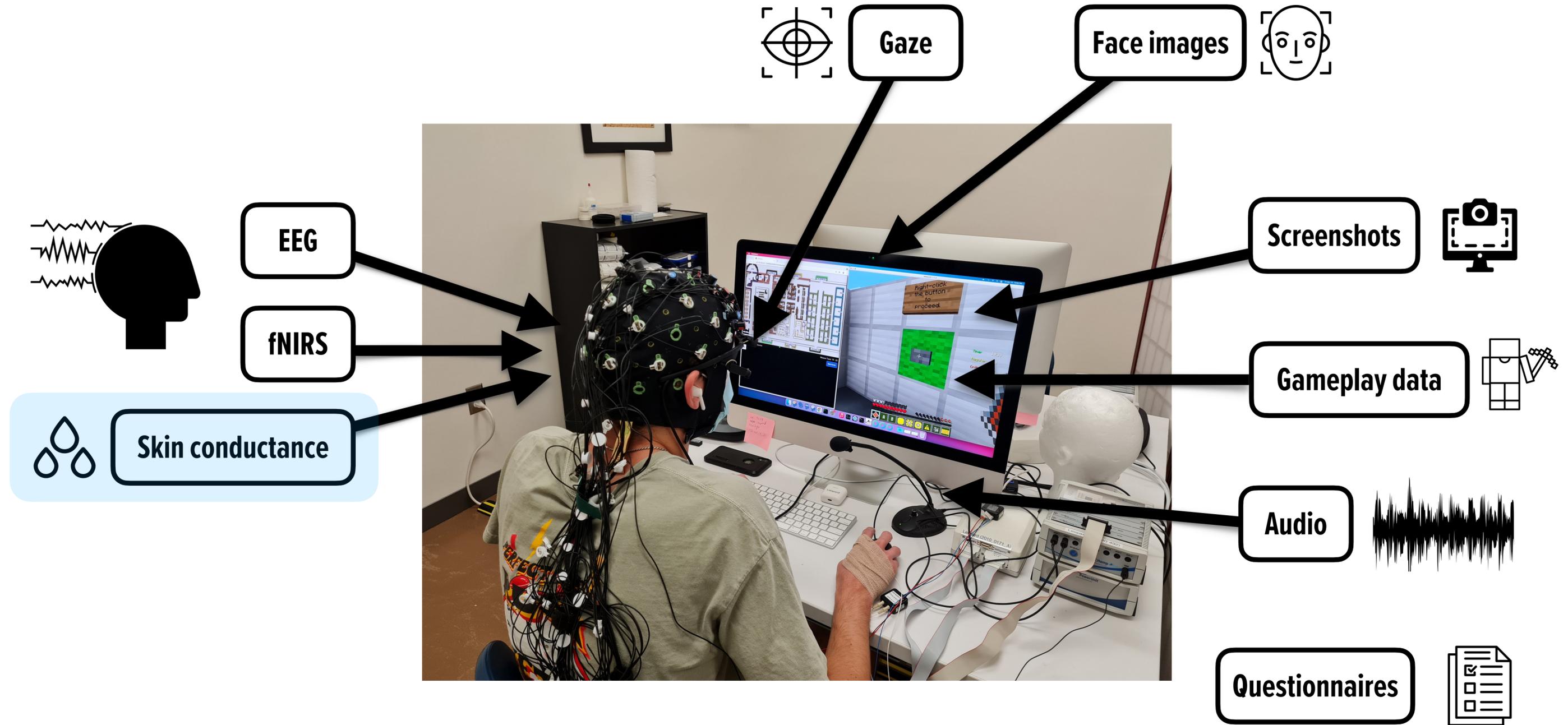# Modalities in the ToMCAT dataset

Gaze

Face images

EEG

fNIRS

Skin conductance

EKG

Screenshots

Gameplay data

Audio

Questionnaires

Participant layout

Leopard

Cheetah

Divider

Tiger

Lion

Pre-session

Consent form → Cap measurement → Speech Elicitation Tasks → Questionnaires

Group session

Affective tasks

Rest state task → Finger-tapping Task → Individual Affective Task → Team Affective Task

Ping-pong tasks

Competitive Ping Pong Task → Cooperative Ping Pong Task

Minecraft tasks

Saturn A → Saturn B → Post-game survey
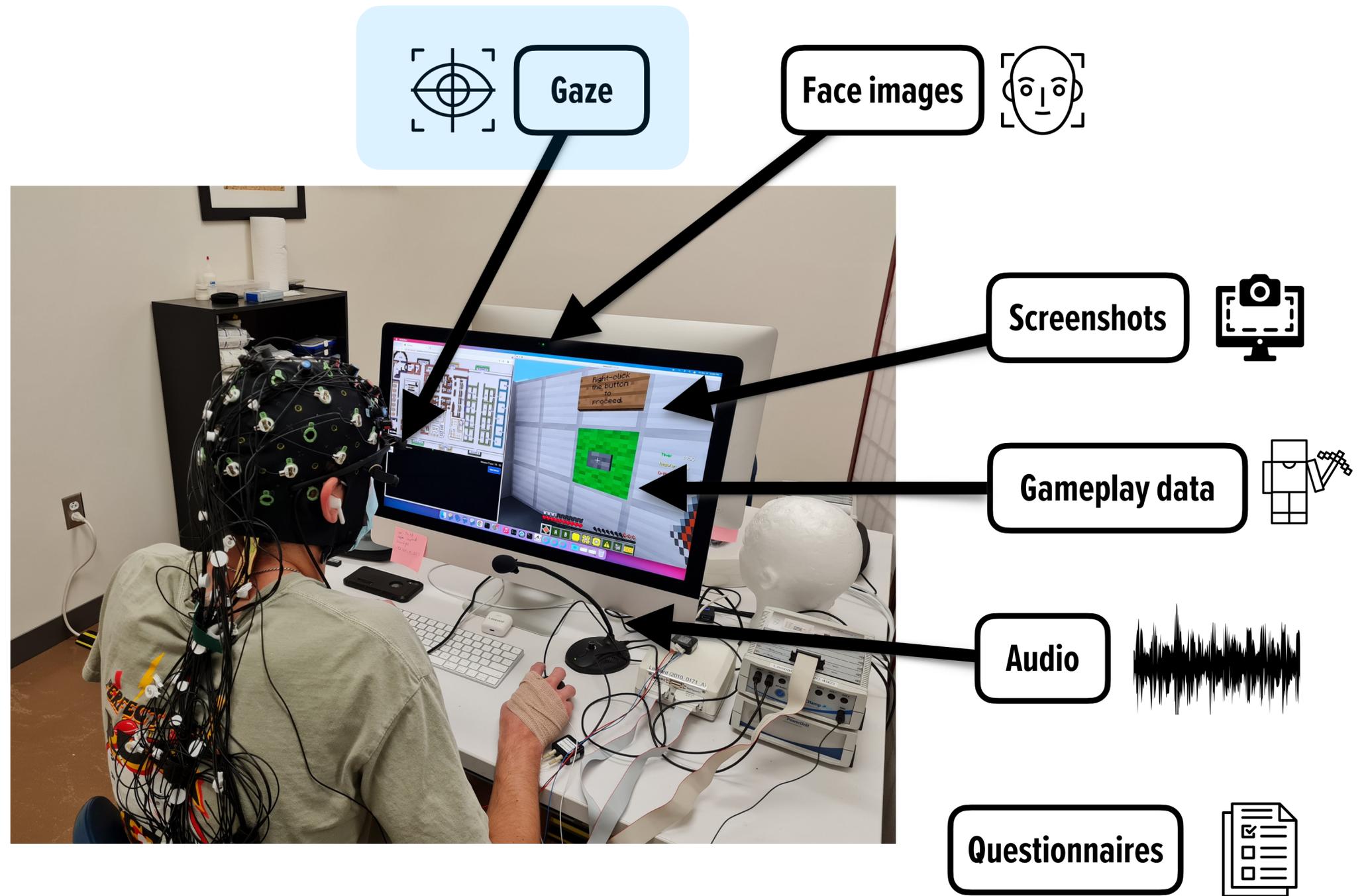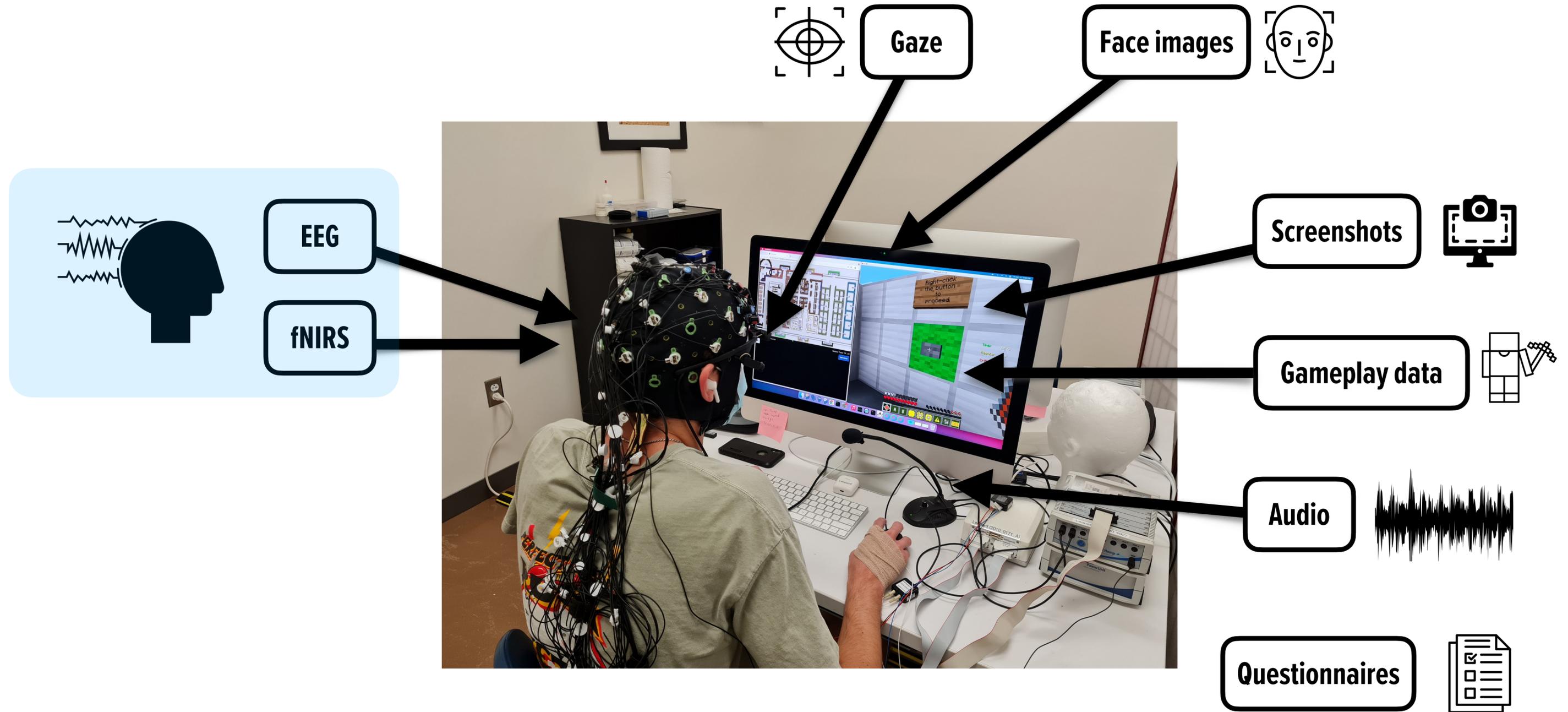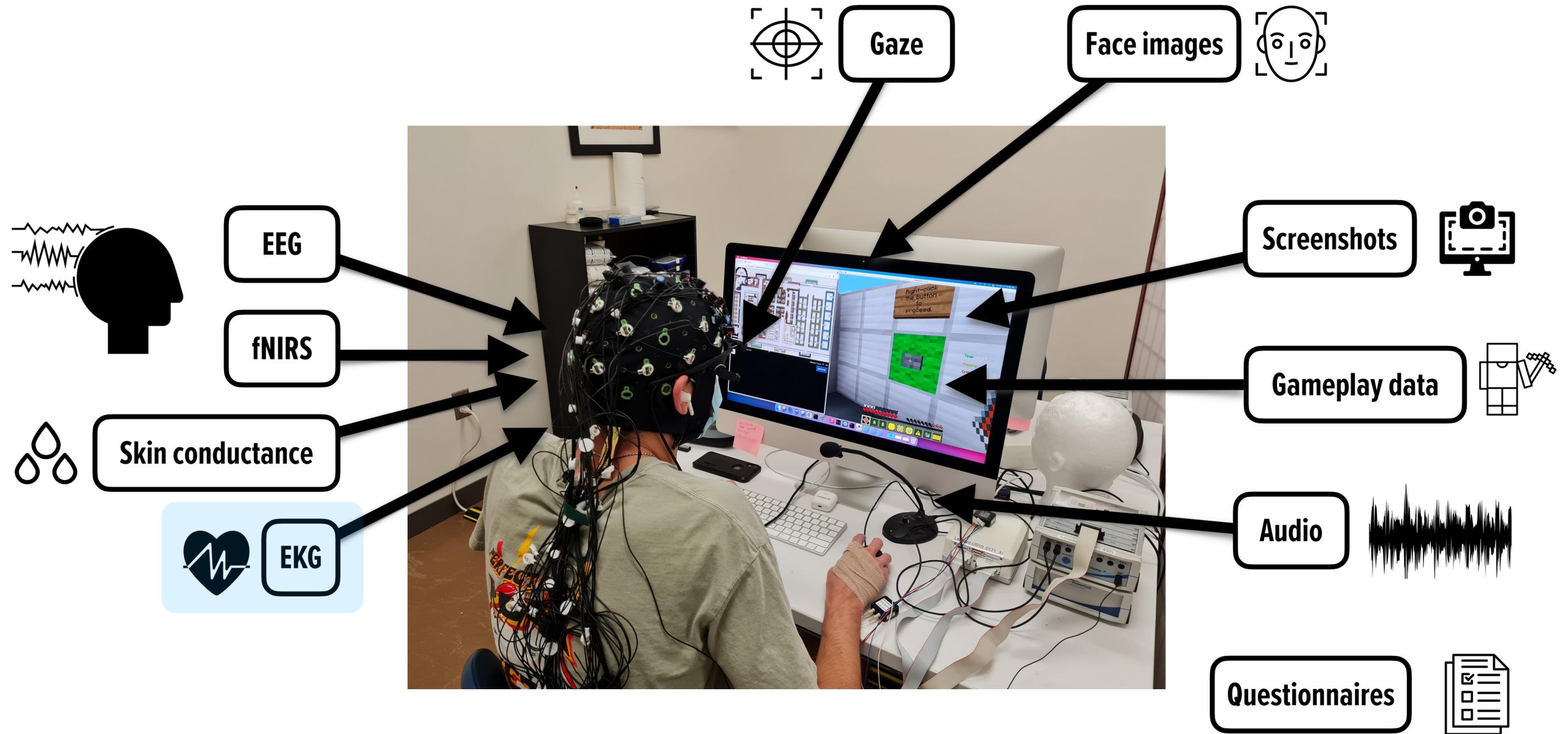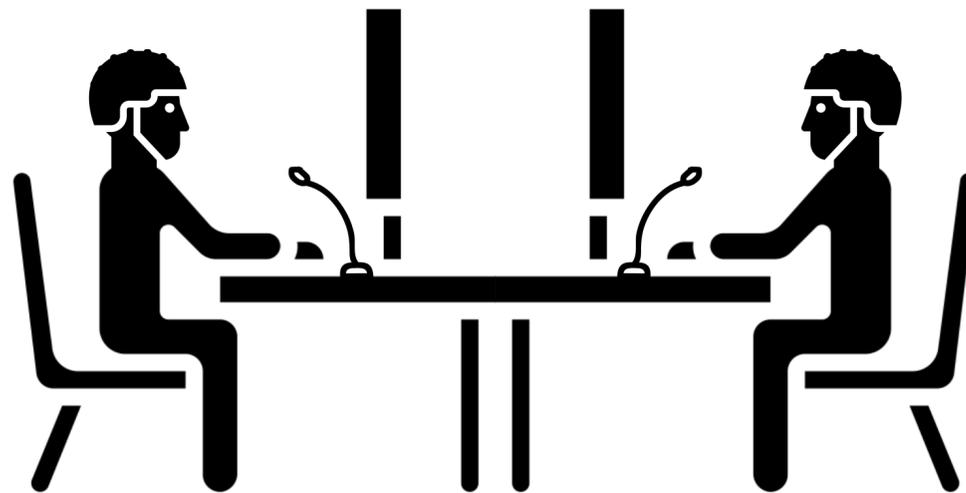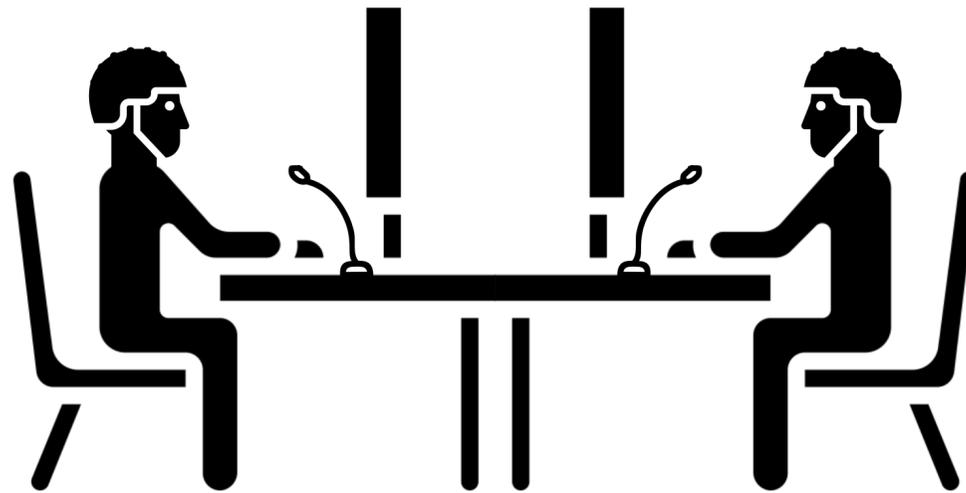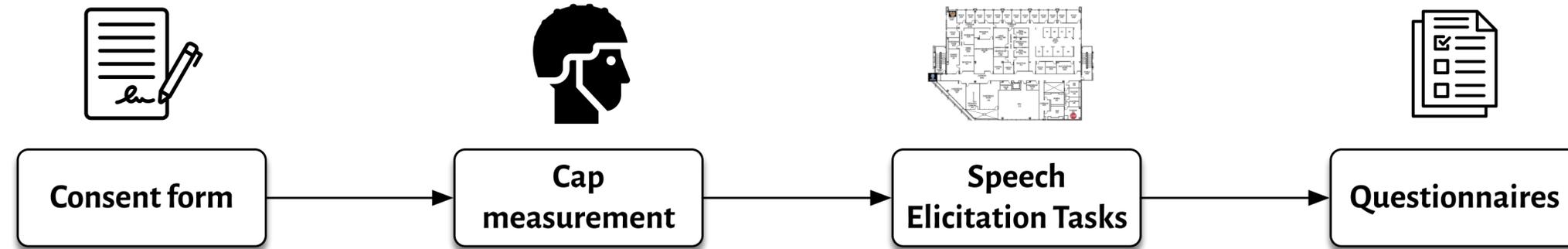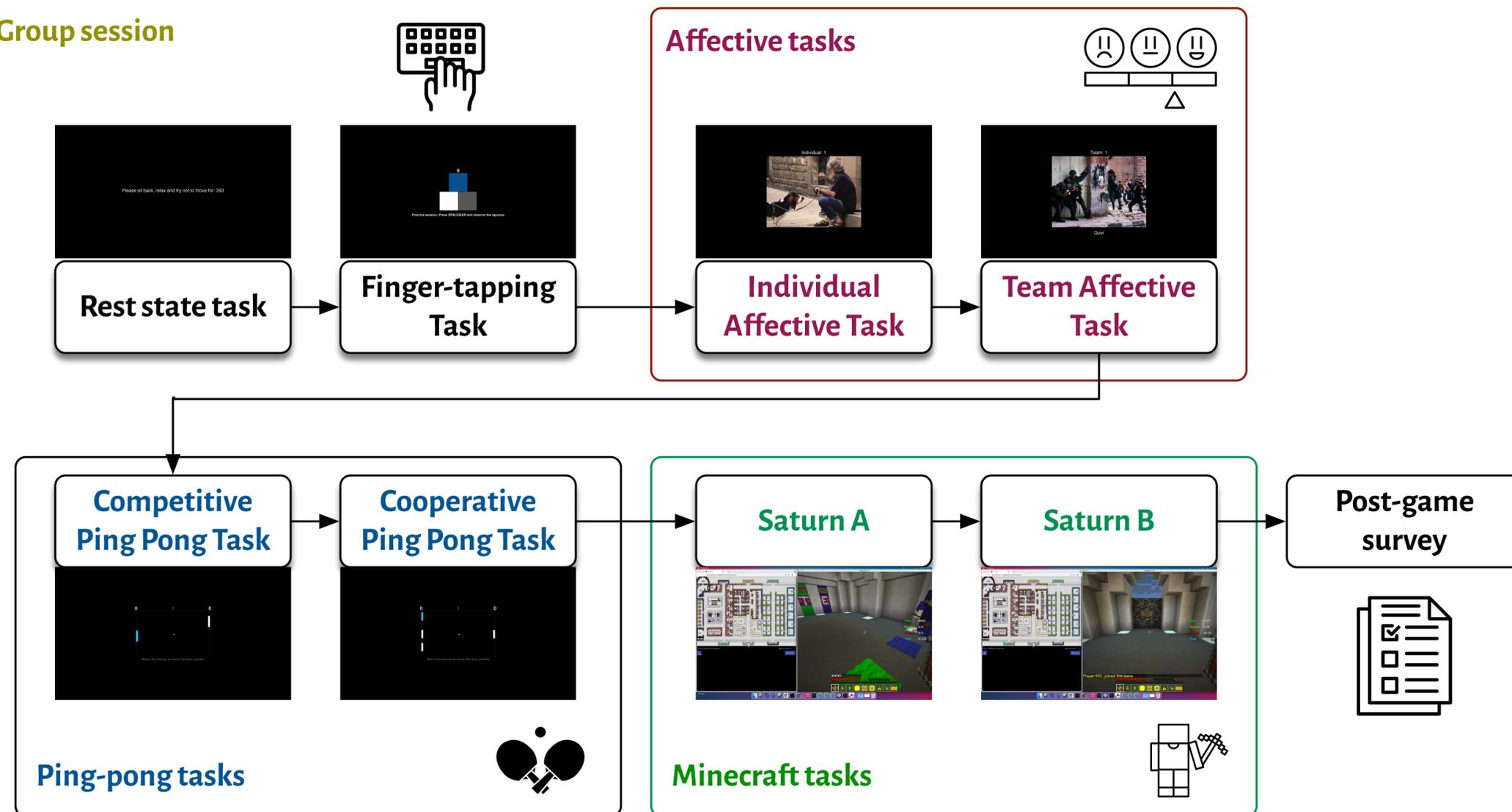
Task flow

Tenure ⌄   Notion   INFO 521 ⌄   NGT ⌄   ARL Collab ⌄   ML4AI ⌄   Ramona ⌄   SKEMA ⌄   ASIST-ToMCAT ⌄   Completing F...ion & Impact   ASIST TA3 ⌄

Home Database Ethical Review Derived Data Products Updates ⌄ Errata Schema Papers and Systems Mailing List

screen_capture
Table with screen capture data.

Column descriptions

group_session
Group session ID

task
Task ID.

station
The iMac that the participant was stationed at.

participant
Participant ID. A participant ID of -1 indicates an unknown participant. All rows that do not have values for the `task` column will have their participant ID set to -1. In the future, we will label some of these inter/pre/post task rows with the correct participant ID.

timestamp_unix
The Unix timestamp of the event.

timestamp_iso8601
The ISO-8601 timestamp of the event.

filename
Name of the screenshot file.

timestamp_origin
The source of the timestamp. The priority is precision-driven. LSL time are used if available, otherwise, creation time is used if available. Modification time is used as last resort.

where task = "saturn_b"

| task ▾ | = ▾ | saturn_b | ✖ |

| - column - ▾ | = ▾ | | |

Apply

✎ View and edit SQL

This data as json, copyable, CSV (advanced)

group_session ⚙	station ⚙	participant ⚙	id ⚙	task ⚙	timestamp_unix	timestamp_iso8601 ⚙	filename ⚙	url ⚙	timestamp_origin ⚙
exp_2022_10_07_15	leopard	34	50485	saturn_b	1665189669.8526385	2022-10-08T00:41:09.852638+00:00	output050486.png		modification
exp_2022_10_07_15	leopard	34	50486	saturn_b	1665189669.9326417	2022-10-08T00:41:09.932642+00:00	output050487.png		modification

Web interface

23

Open-source software

Continued Engagement

- Scale and complexity of dataset make it infeasible to annotate every interesting label type

- Researchers can build upon this dataset by adding annotation layers, unlocking the ability to answer additional research questions.

- We **highly** encourage users to sign up for our mailing list to get updates on data issues and annotation layers, benchmarks, and documentation.
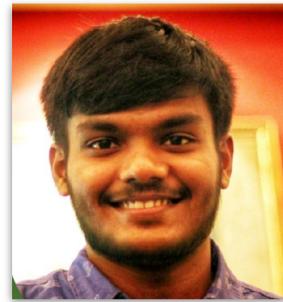
Team

Adarsh Pyarelal

Eric Duong

Caleb Jones Shibu

Paulo Soares

Savannah Boyd
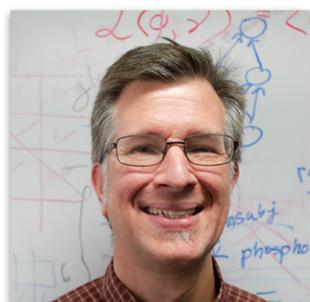
Payal Khosla

Valeria Pfeifer

Diheng Zhang

Eric Andrews

Rick Champlin

Vincent Raymond

Meghavarshini Krishnaswamy

Clayton Morrison

Emily Butler

Kobus Barnard