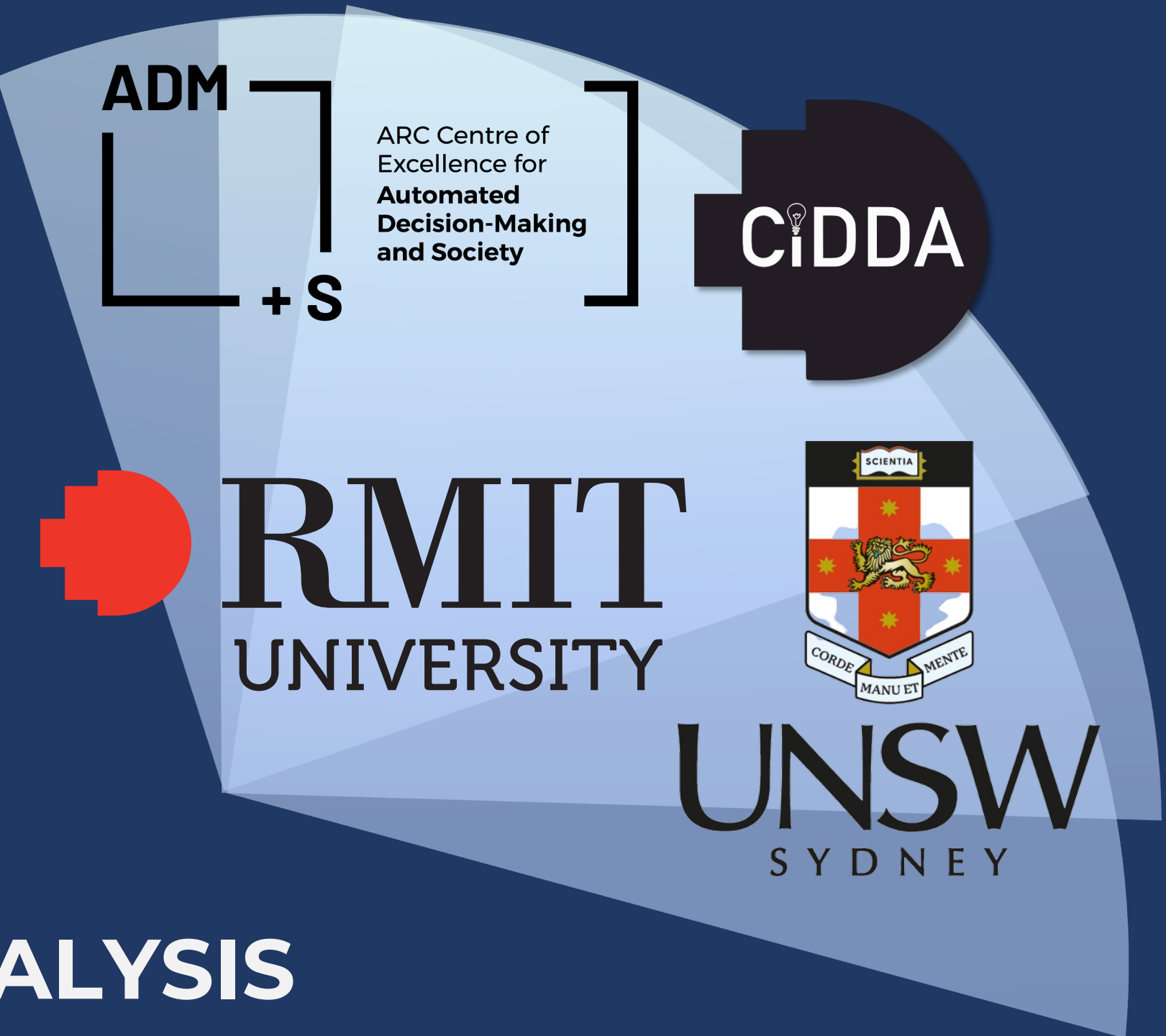


Examining the Impact of Uncontrolled Variables on Physiological Signals in User Studies for Information Processing

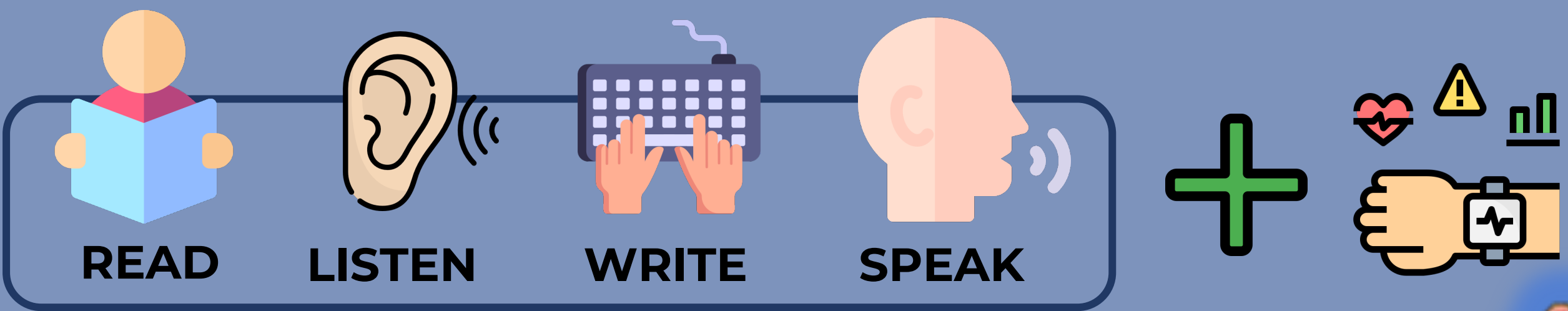
Kaixin Ji^{1,3}, Damiano Spina^{1,3}, Danula Hettiachchi^{1,3}, Flora Salim^{2,3}, Falk Scholer^{1,3}

¹RMIT University, ²UNSW University, ³ARC Centre of Excellence for Automated Decision Making + Society



MOTIVATION

- **Laboratory** user studies require controlling for factors influenced by each individual participating in the experiment, and their perception of the information and task.
- **Information-seeking models** characterize how users interact with a variety of systems, e.g., screen-based search, spoken conversational search, multimedia platforms.
- What **physiological signals** (captured by wearable devices) can tell us about how users engage in Information Processing Activities (IPAs)?



Information-Processing Activities (IPAs)

AIM

Can we predict the specific IPA the user performed by feeding a machine-learning model with the signals obtained from the sensors?

- **But** signals are **sensitive** to noise:

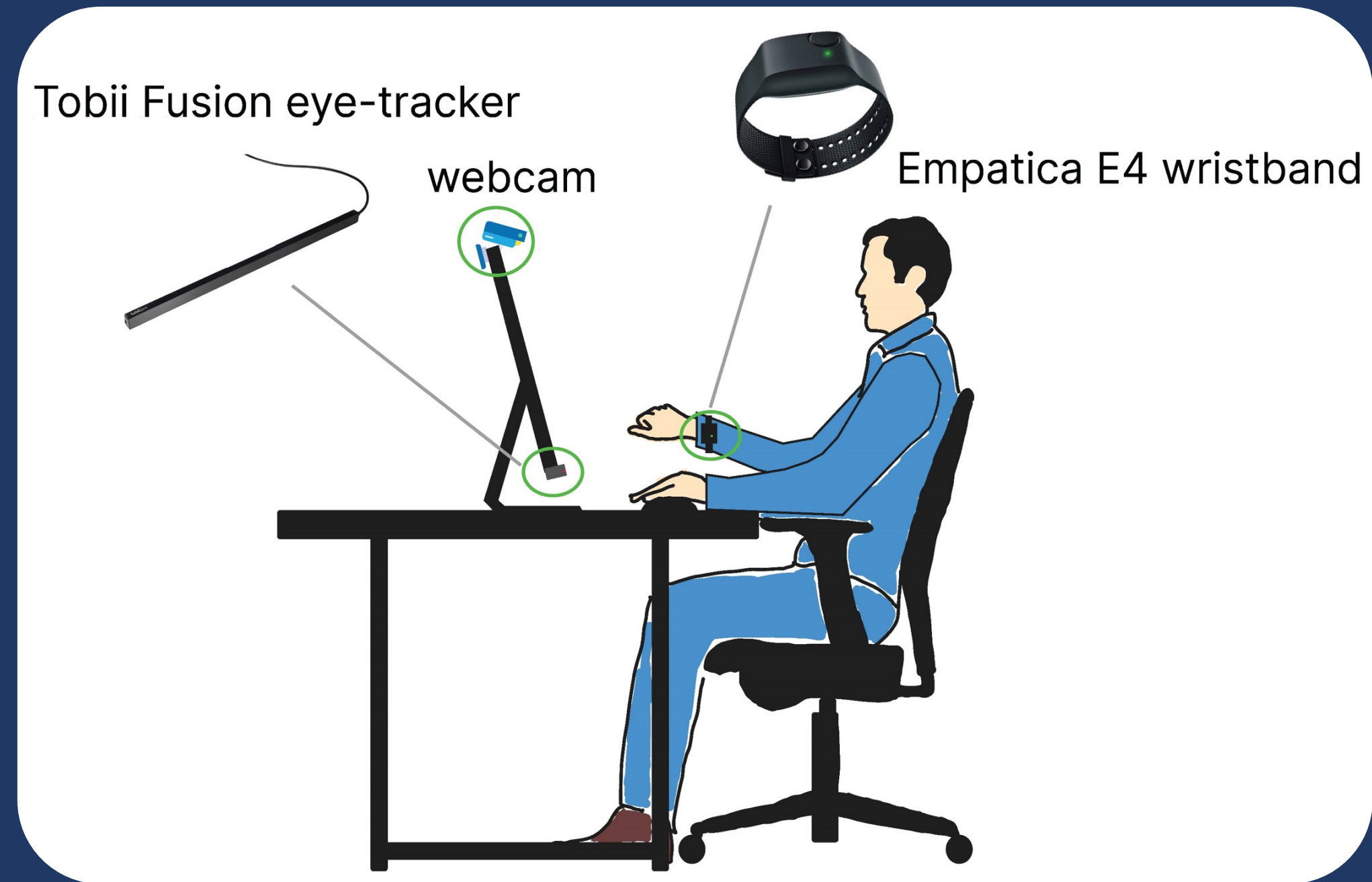
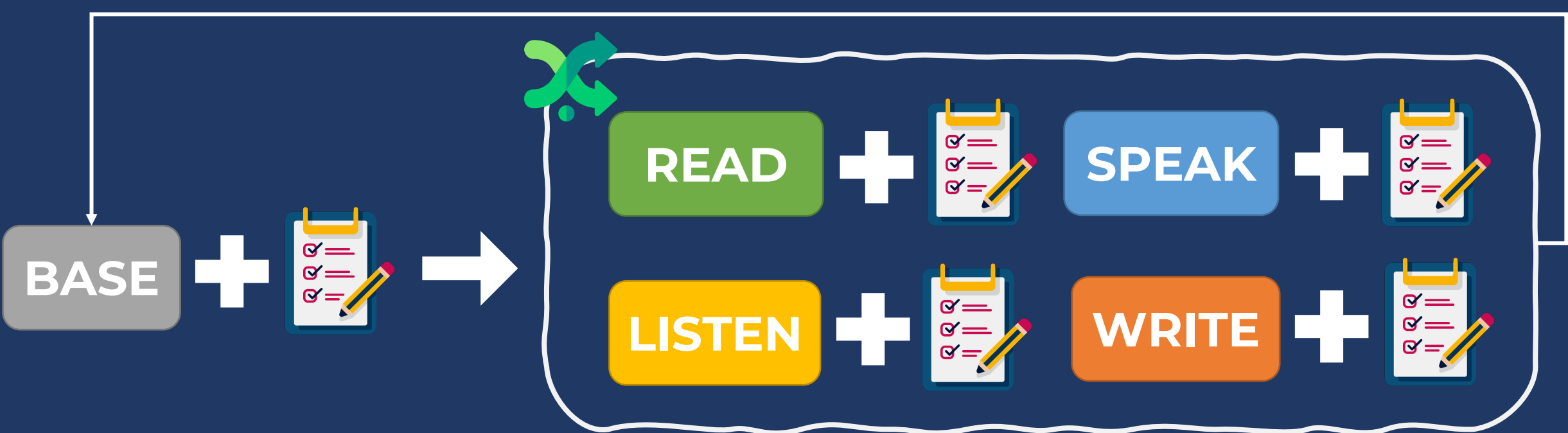
how variables (controlled and uncontrolled) influence the ML model's performance?

and

whether this can be used as a mechanism to scrutinize the validity of our experimental design?

USER STUDY

repeat for 2 complexity (low and high)



DATA & FEATURES (7 participants)

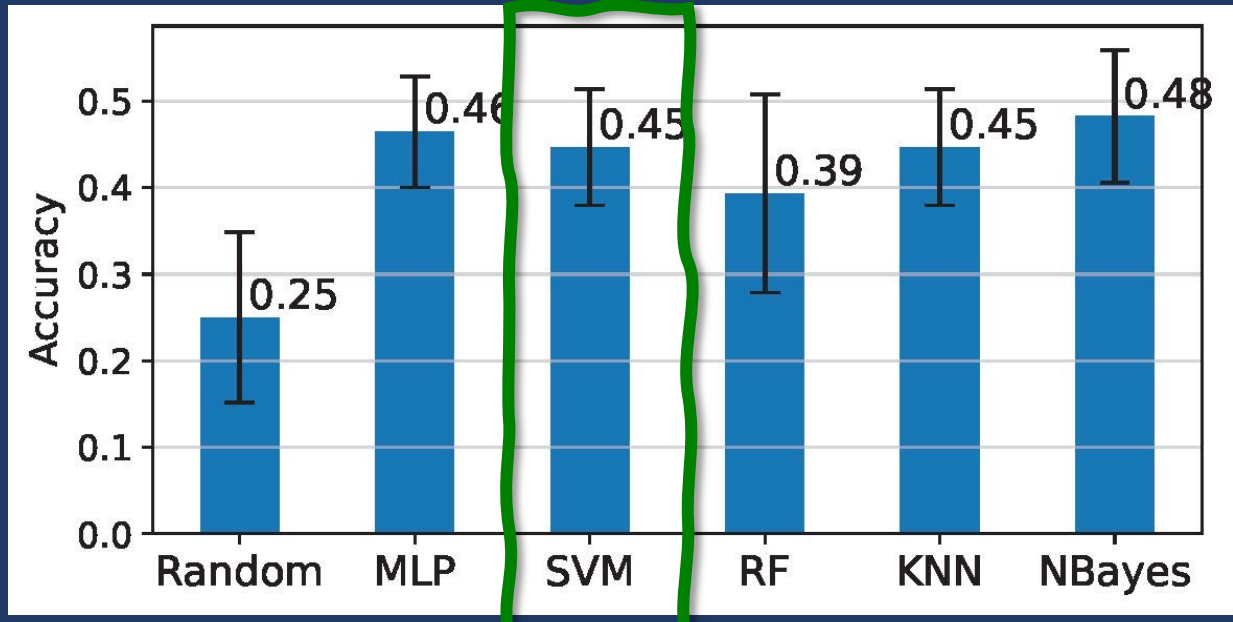
Data	Features
Electrodermal Activity	original, 1 st & 2 nd derivative
Blood Volume Pulse	
Pupil Diameter	mean, median, standard deviation

RESULTS & ANALYSIS

TASK: 4-class IPAs classification
Cross-Validation: Leave-One-Participant-Out

1. Model Selection

compare machine learning (ML) models which are commonly used on physiological data:



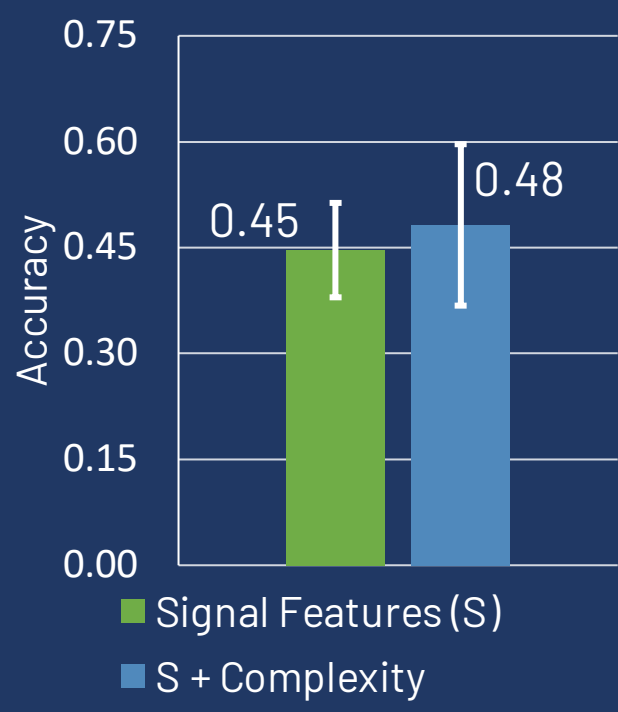
Accuracy of different multi-class ML models, using leave-one-participant-out (7 folds). Error bars indicate 95% confidence intervals (t-distribution).

2. Impact of Variables

Assumption: The ML model is supposed to calibrate the inputs (i.e., performance should not change) if the variables do not impact the user study.

Controlled Variable #1: Complexity (binary)

- complexity level should be carefully controlled when designing a user study related to reading and listening.

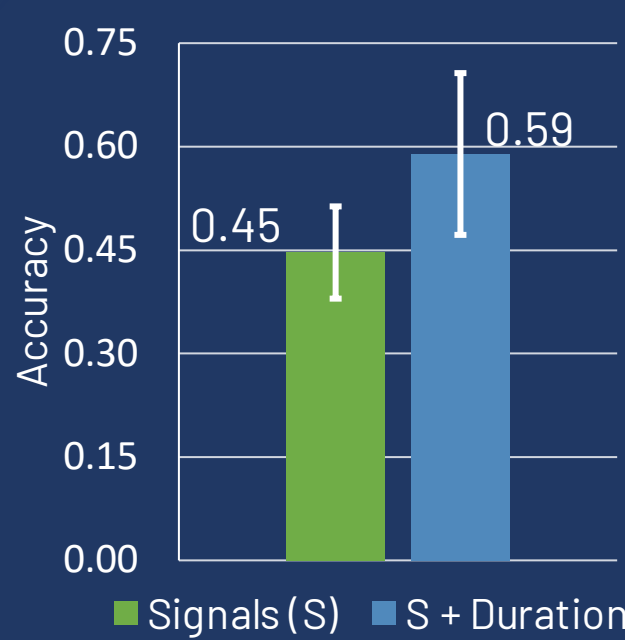


Controlled Variable #2 : Sequence (nominal)

- task sequence may not impact the model.

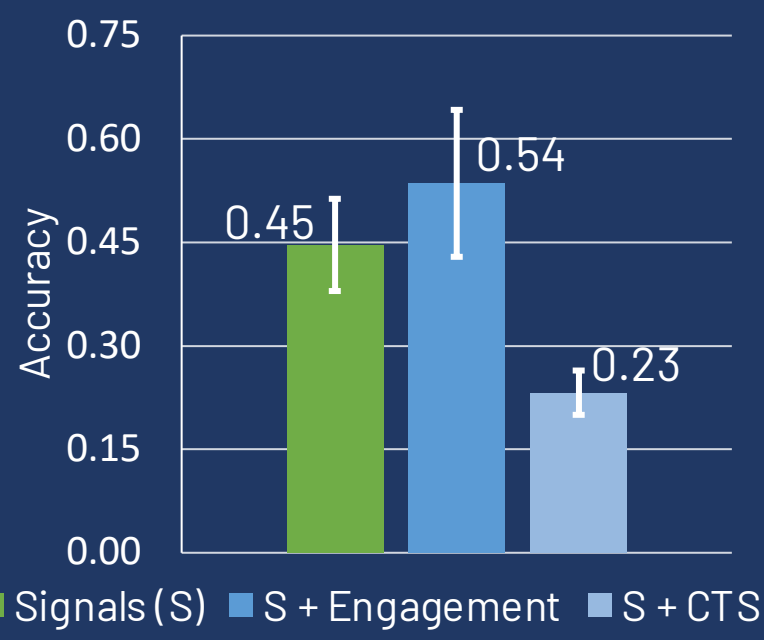


Uncontrolled Variable #1: Duration (numeric)



- IPA 'duration' may play a strong role in the user study.
- encourage consistent task durations by informing users of expected IPA durations and displaying a timer to motivate timely completion without enforcement.

Uncontrolled Variables #2: Cumulative Time Spent & #3: Engagement Score



- both results have substantial influences on the model performances.
- shorter sessions may be more suitable to avoid fatigue (e.g., having a longer break between sections).

CONCLUSION

- We can infer whether our experimental design needs refinement by examining changes in the model's classification performances.
- Given the exploratory nature of our investigation and small sample size limitations, further research is needed to understand the variables' impacts fully.
- **The methodology described allowed us to reveal the shortcomings in our experimental design early in data collection.**