Incorporating Gaze into Social Navigation

Justin W. Hart, Reuth Mirsky, Xuesu Xiao, and Peter Stone

Building-Wide Intelligence Project
Learning Agents Research Group
Department of Computer Science
The University of Texas at Austin

Building-Wide Intelligence

- Real-world deployment
- Fleet of 6 autonomous service robots
- Tasks
  - Object delivery
  - Providing directions
  - Messages
- Provide services to building occupants
- Run constantly during workday

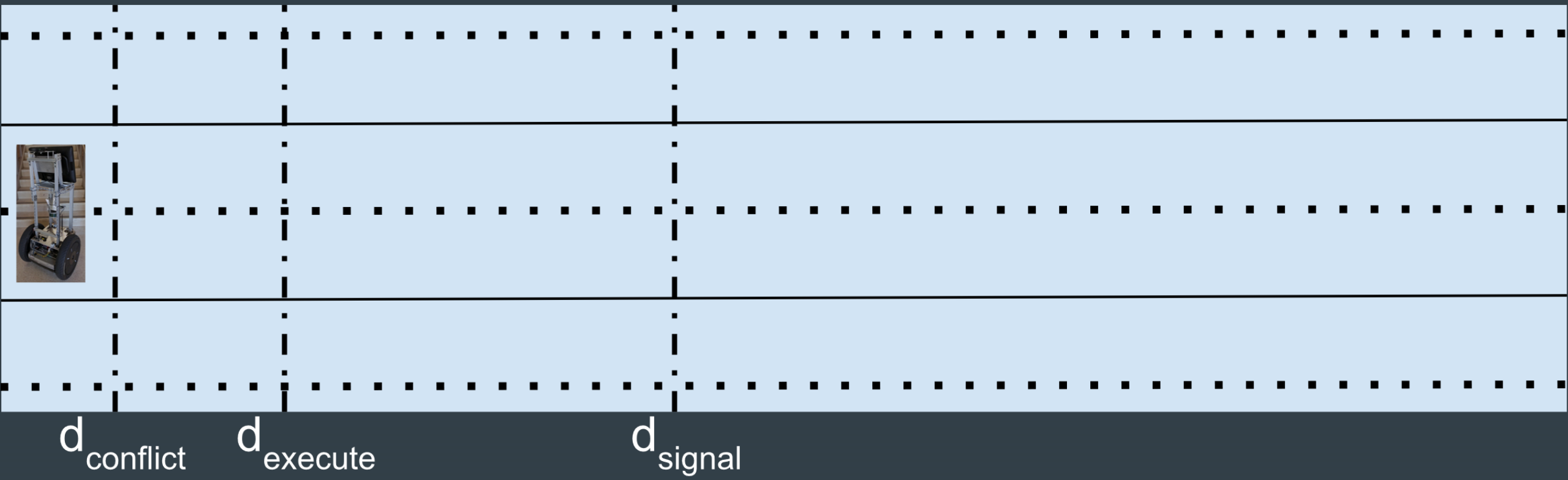
Using Turn Signals

Right Turn

Left Turn

R. Fernandez, N. John, S. Kirmani, J. Hart, J. Sinapov, and P. Stone **Passive Demonstrations of Light-Based Robot Signals for Improved Human Interpretability**. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2018)*. Nanjing China, August 2018.

Modeling the Hallway – 3 Traffic Lanes



d_{signal} = 7m – Distance the robot indicates lane change
$d_{execute}$ = 2.75m – Distance the robot changes lanes
$d_{conflict}$ = 1m – Distance the robot stops

This is designed to be difficult!
2.75m was tuned to be the last possible moment
Robot always goes left. People expect it to go right

LED turn signals are non-obvious

- Users will conflict *unless they understand the signal*
  - For turn signals - 100% of users conflict with the robot

- Introduced a “passive demonstration.”
  - Robot makes a lane change, showing the signal, before needing the signal
  - Introduction of a “passive demonstration” is sufficient for 90% of users to understand turn signal.

R. Fernandez, N. John, S. Kirmani, J. Hart, J. Sinapov, and P. Stone **Passive Demonstrations of Light-Based Robot Signals for Improved Human Interpretability**. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2018)*. Nanjing China, August 2018.

Passive Demonstrations of Light-Based Robot Signals for Improved Human Interpretability

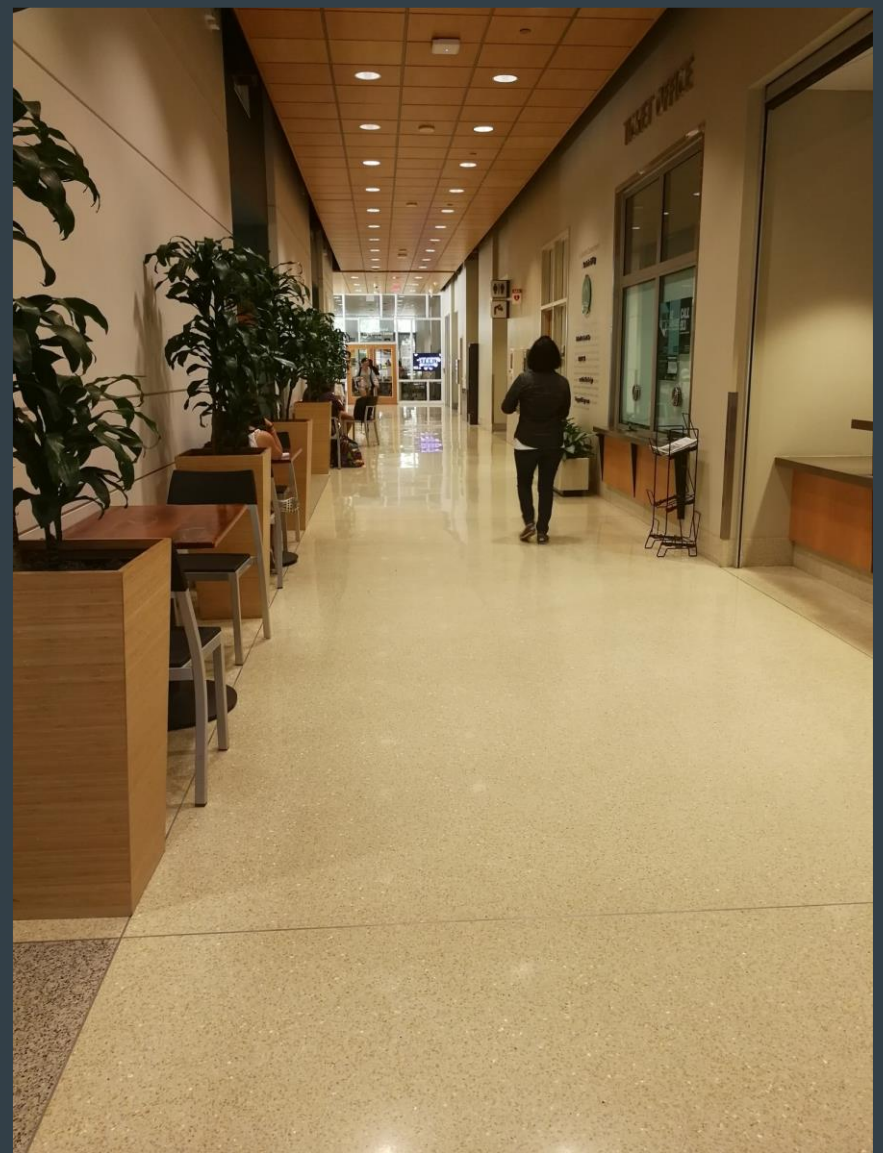
Rolando Fernandez, Nathan John, Sean Kirmani, Justin Hart,
Jivko Sinapov, and Peter Stone

What about gaze?

- Interpreting gaze may be “hard-wired” in the brain (Emery 2000)
- Gaze communicated “implicitly” (Admoni & Scassellati 2017)
- Head pose (often a proxy for gaze) can be used as a predictor of a person’s intended trajectory (Unhlelkar, Perez-D’Arino, & Shah, 2015)

Validating gaze & navigation in a human study

- Controlled confederates' gaze pattern
  - Congruent – Walk & look in same direction
  - Incongruent – Walk & look in opposite direction
  - No Gaze – Confederate looked at their cell phone
- Observed 220 interactions (130 F / 90 M)
- Annotated navigational conflicts
  - Instances where people bumped into each other, or nearly bumped into each other
- Congruent – 25% conflict
- Incongruent – 41% conflict*
- No Gaze – 28% conflict

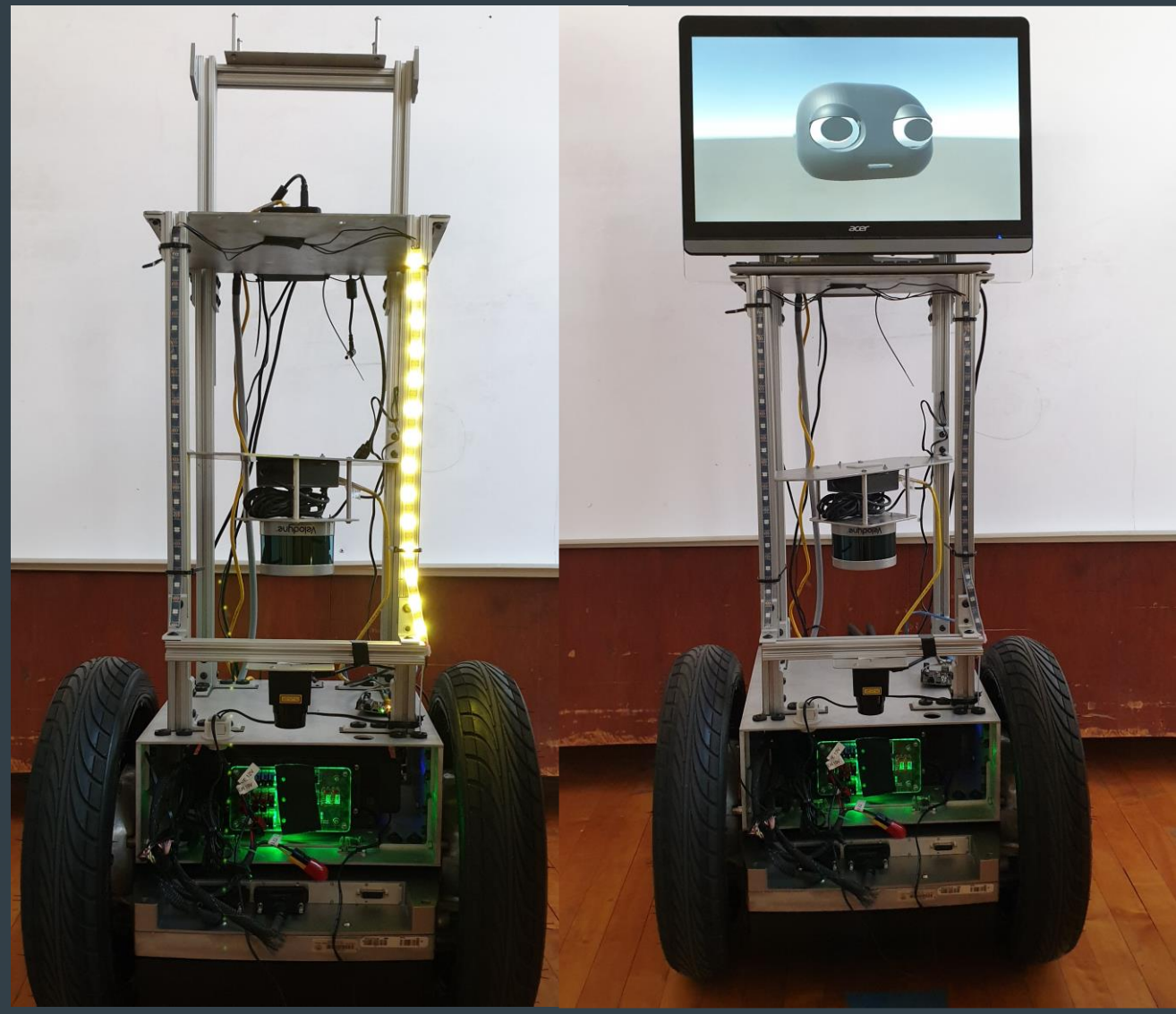


Human Experiment

Condition 1 - Congruent Look

Does gaze work better than turn signals?

- Contrast LED turn signal against a gaze cue
- Gaze as a virtual agent version of the Maki 3D printable robot head, rendered in Unity
- Gaze cue implemented as a head rotation to the “lane” the robot intends to shift to



Robot Experiment

Condition 1 - LED

Human-Robot Study

- Recruited 38 participants
  - Excluded 11, due to head not being displayed on the screen
  - LED Condition – 11 participants
  - Gaze Condition – 16 participants
- Results
  - LED Condition – 100% conflict
  - Gaze Condition – 50% conflict


Robot Experiment
Condition 1
(LED)
Conflict

Does Gaze Work Better than Turn Signals

- In a previous study, we found that LED turn signals were not readily interpreted. (0%)
  - Introduced “passive demonstration” which improved performance to 90%
- Performed a human study verifying the importance of gaze in deconflicting navigational trajectories
- Performed a robot study showing that gaze worked 50% of the time

Using Gaze Instead of Turn Signals

- In a previous study, we found that LED turn signals were not readily interpreted. (0%)
  - Introduced “passive demonstration” which improved performance to 90%
- Performed a human study verifying the importance of gaze in deconflicting navigational trajectories
- Performed a robot study showing that gaze worked 50% of the time

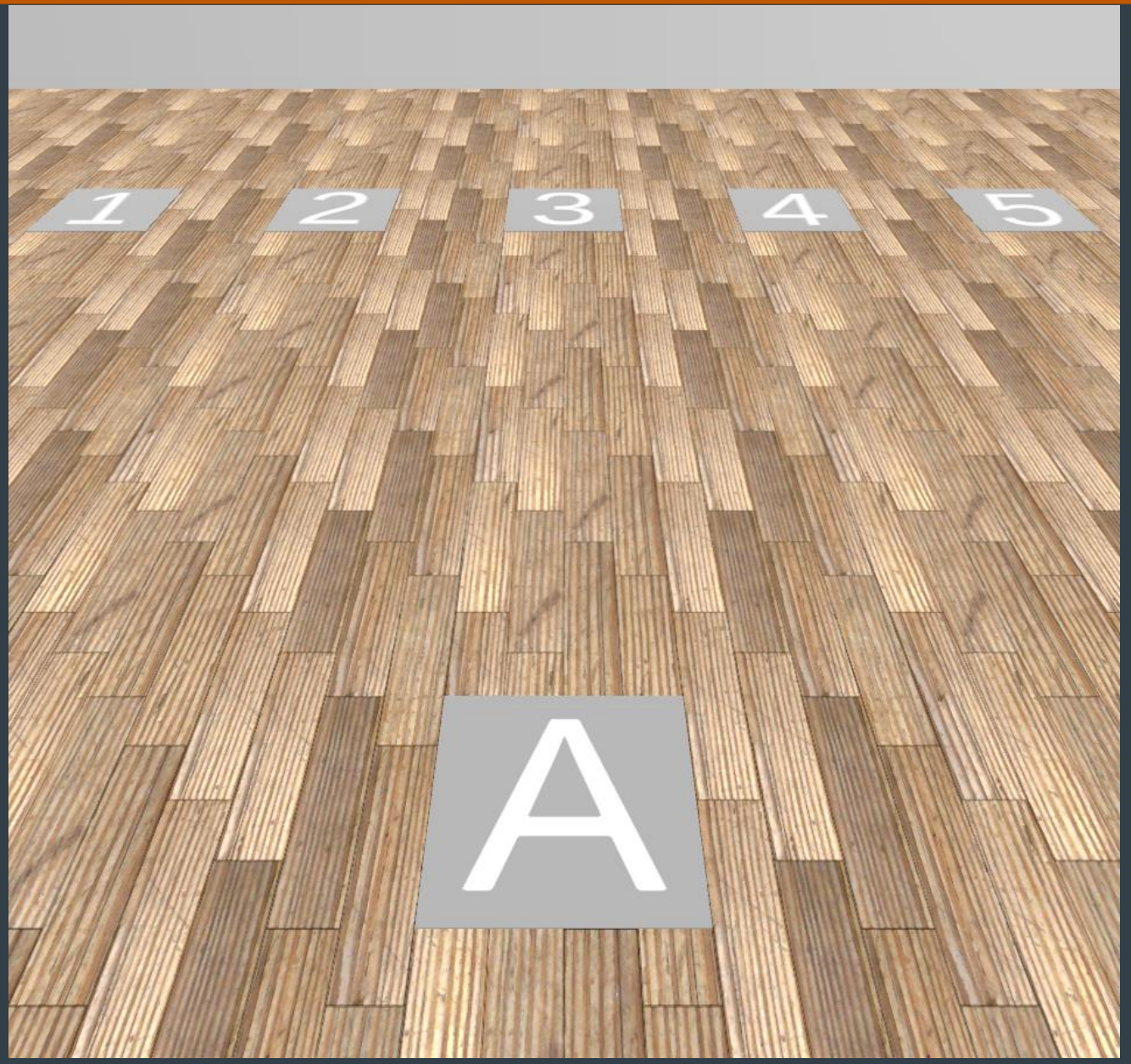
Interpreting Gaze

- By analyzing gaze we can interpret future walking trajectories
- Previous studies have looked at head orientation, but gaze precedes head orientation
  - *This has not been demonstrated as a general principle, but is reflected in our results.*

- Virtual Reality

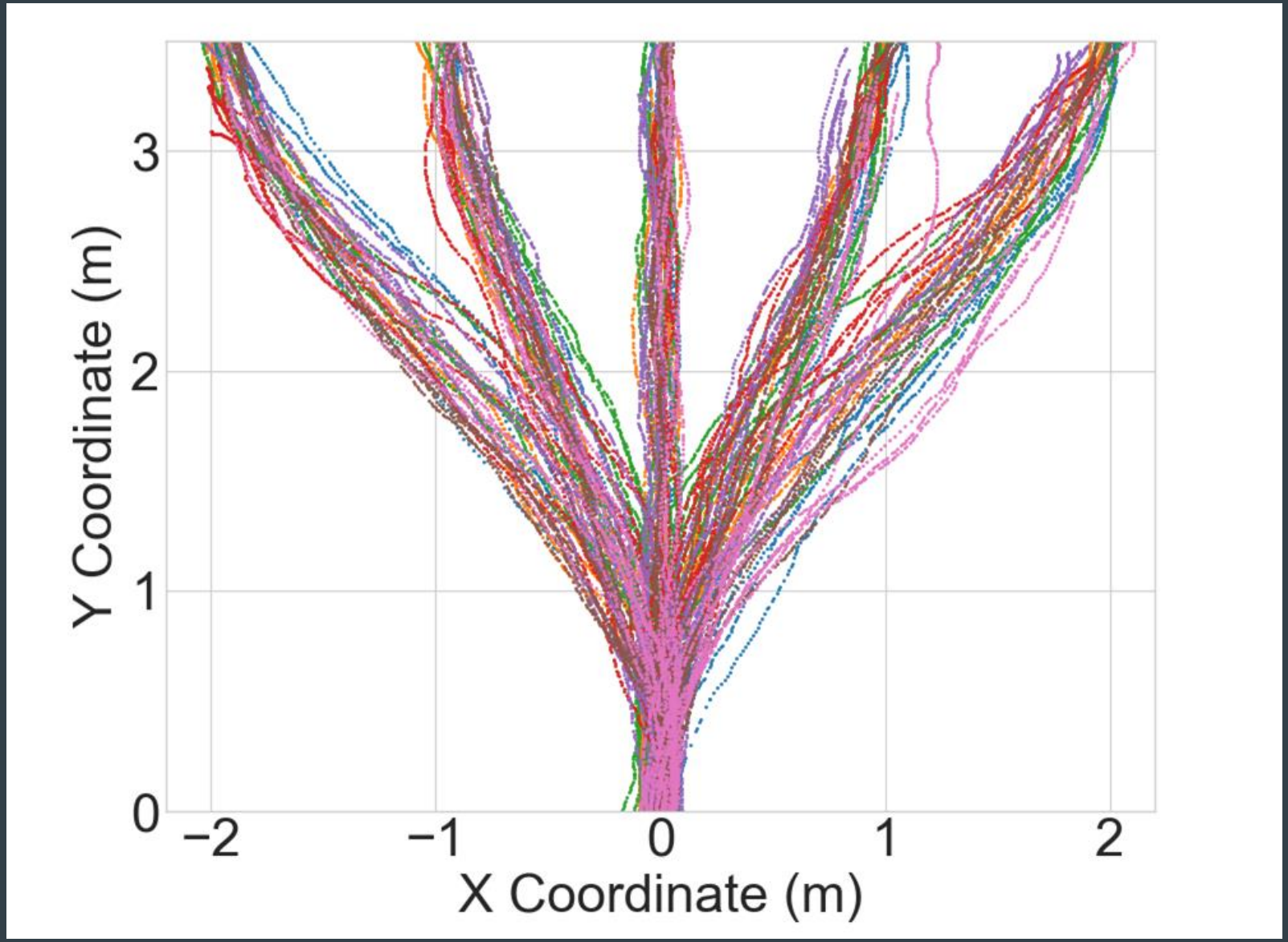
- Body Tracking
- Gaze Tracking

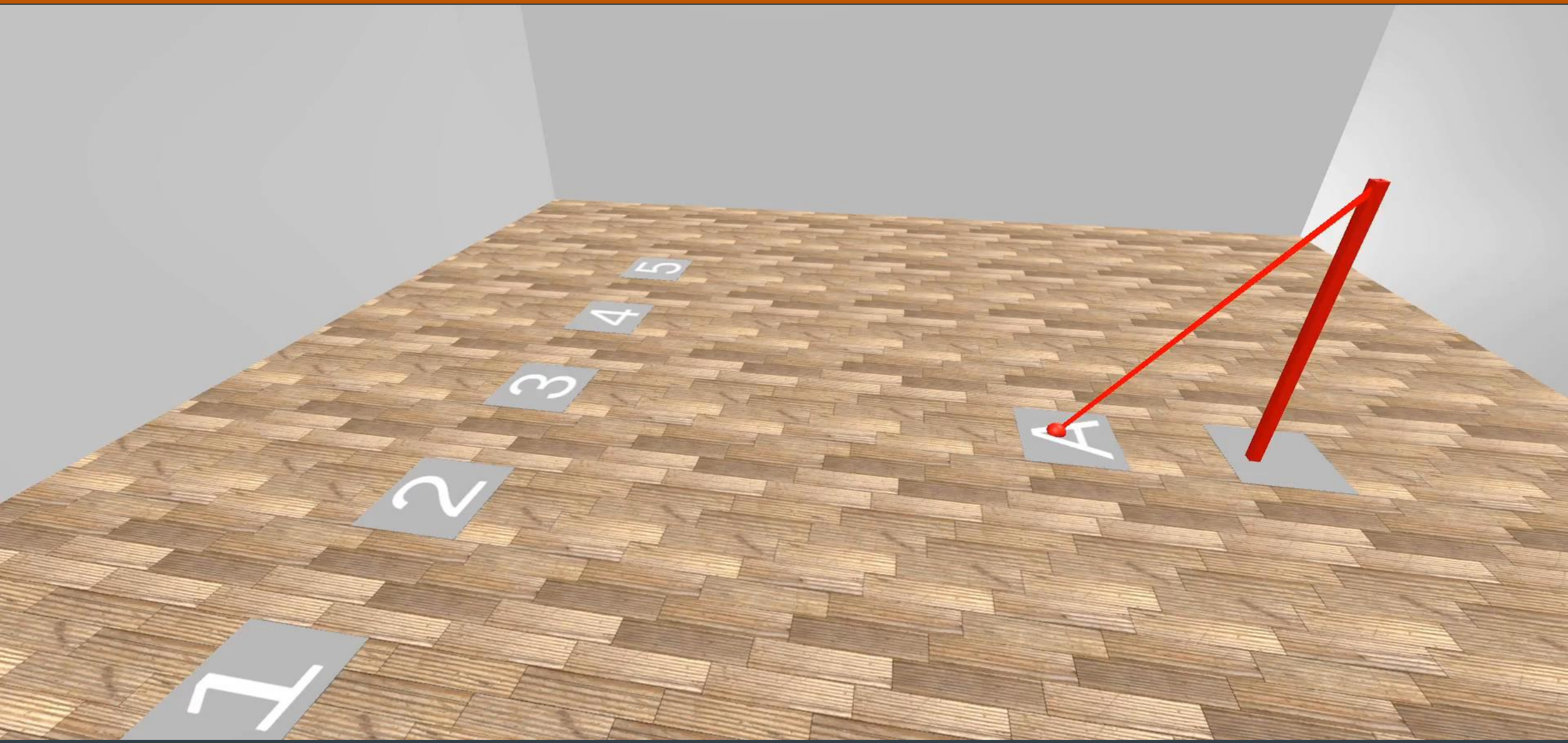
- Study in virtual reality
  - Participants navigate to one of 5 targets
  - Tracking
    - Gaze
    - Head orientation
    - Position in room
- 7 participants – 25 trials each
  - Participants were members of the lab due to COVID-19 protocols
  - 6 male, 1 female
  - Study took ~4 minutes to complete



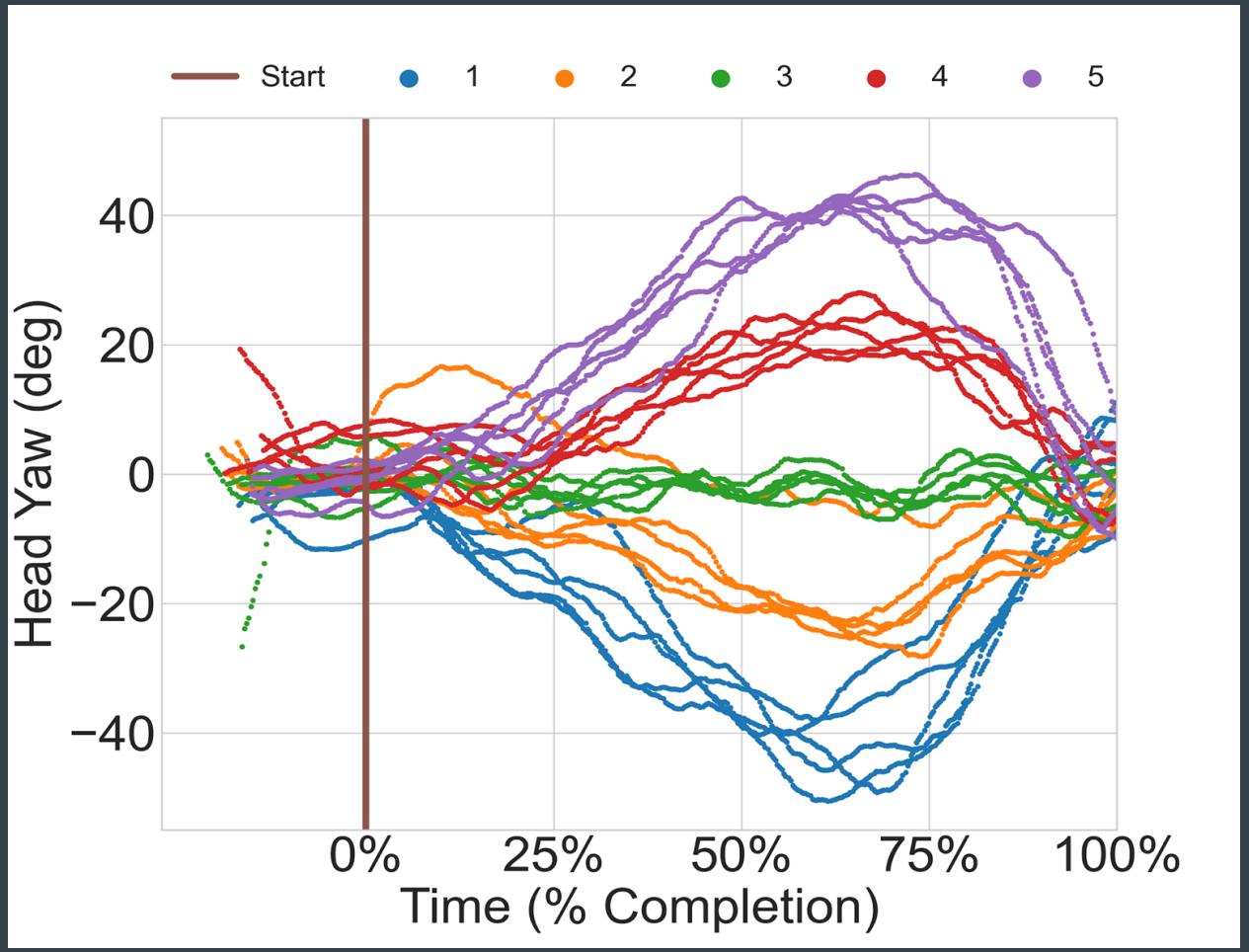

1
2
3
4

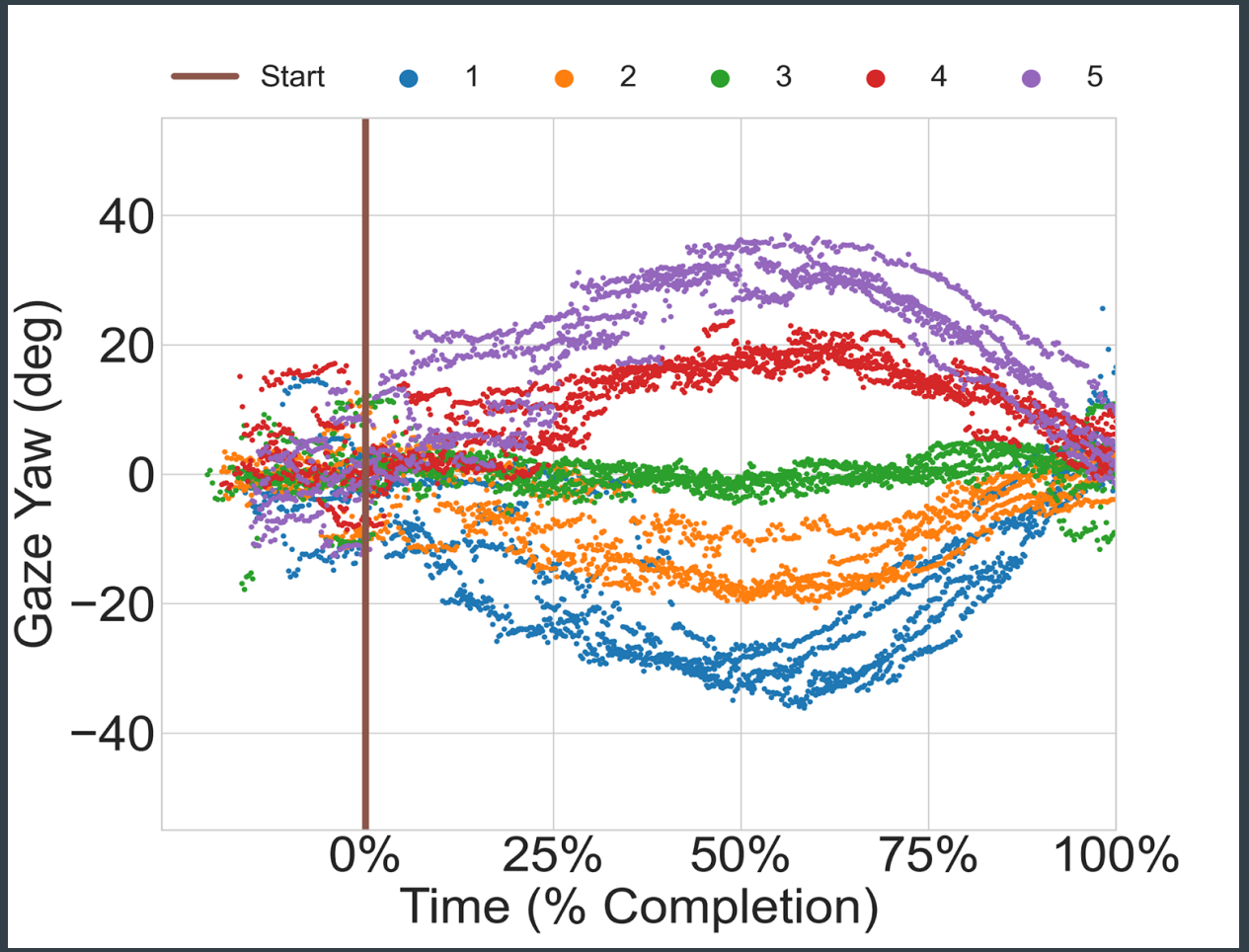
Participants walk directly to the goal



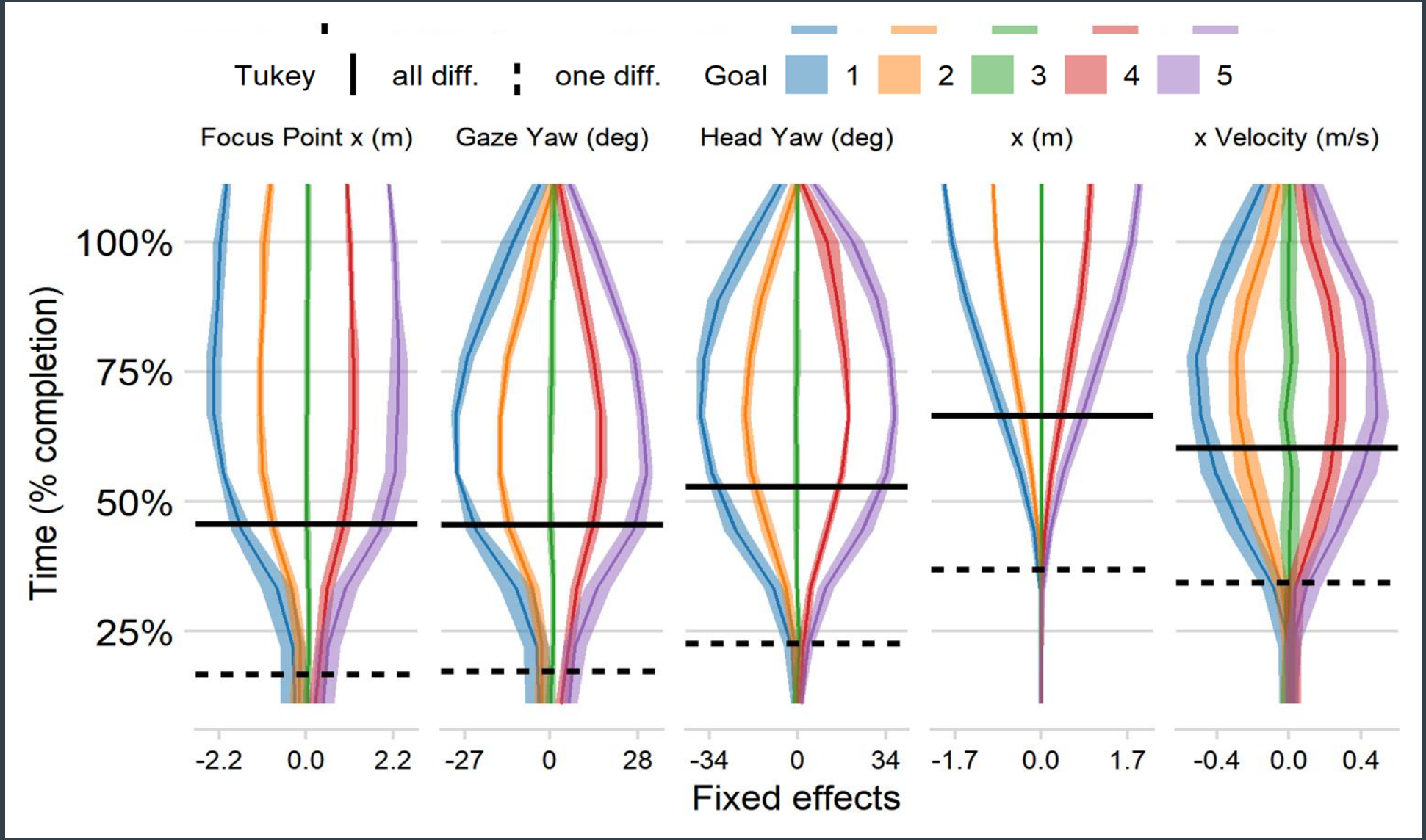

1
2
3
4
5
A

Head and gaze yaw predict walking motion

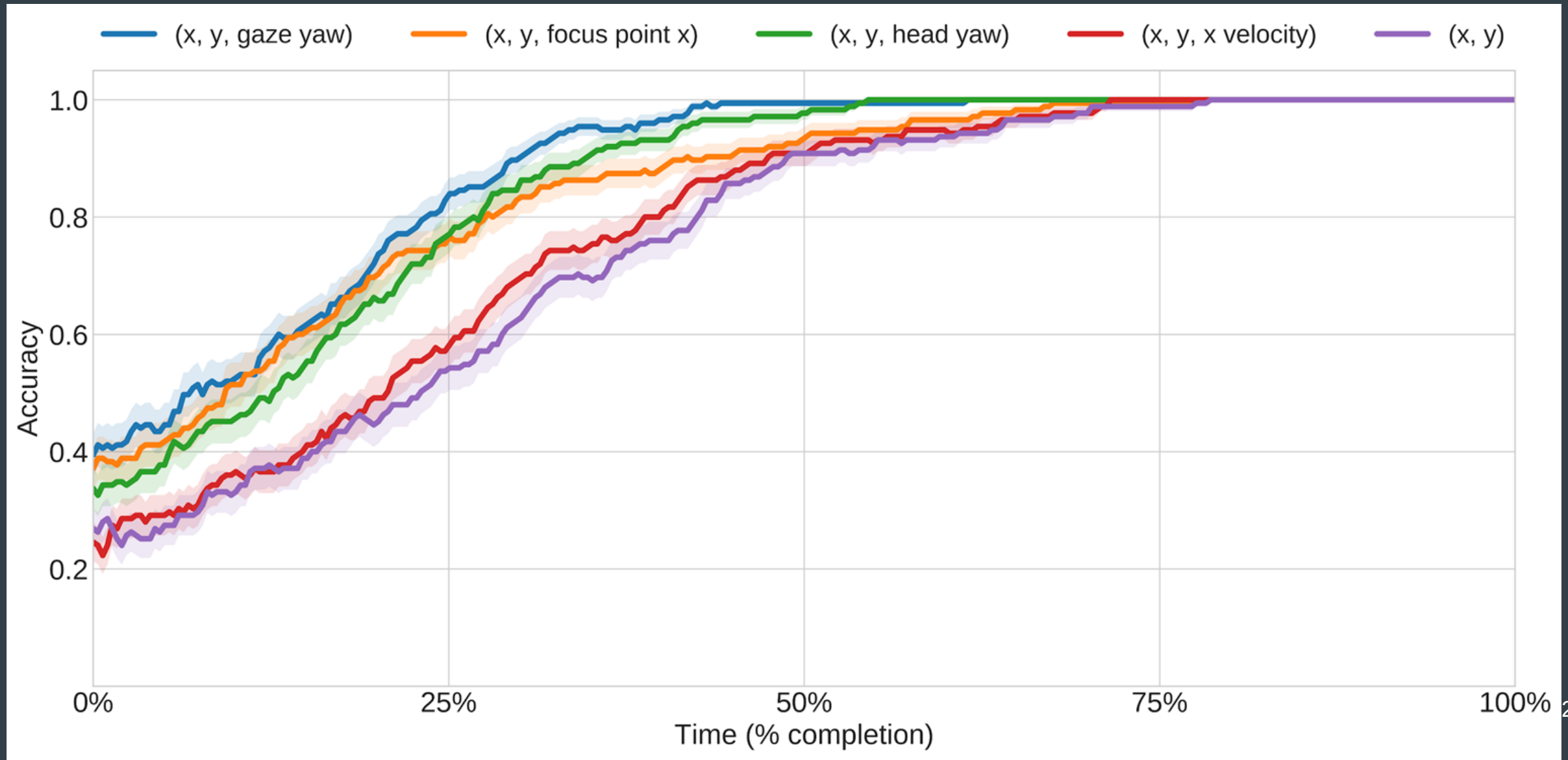




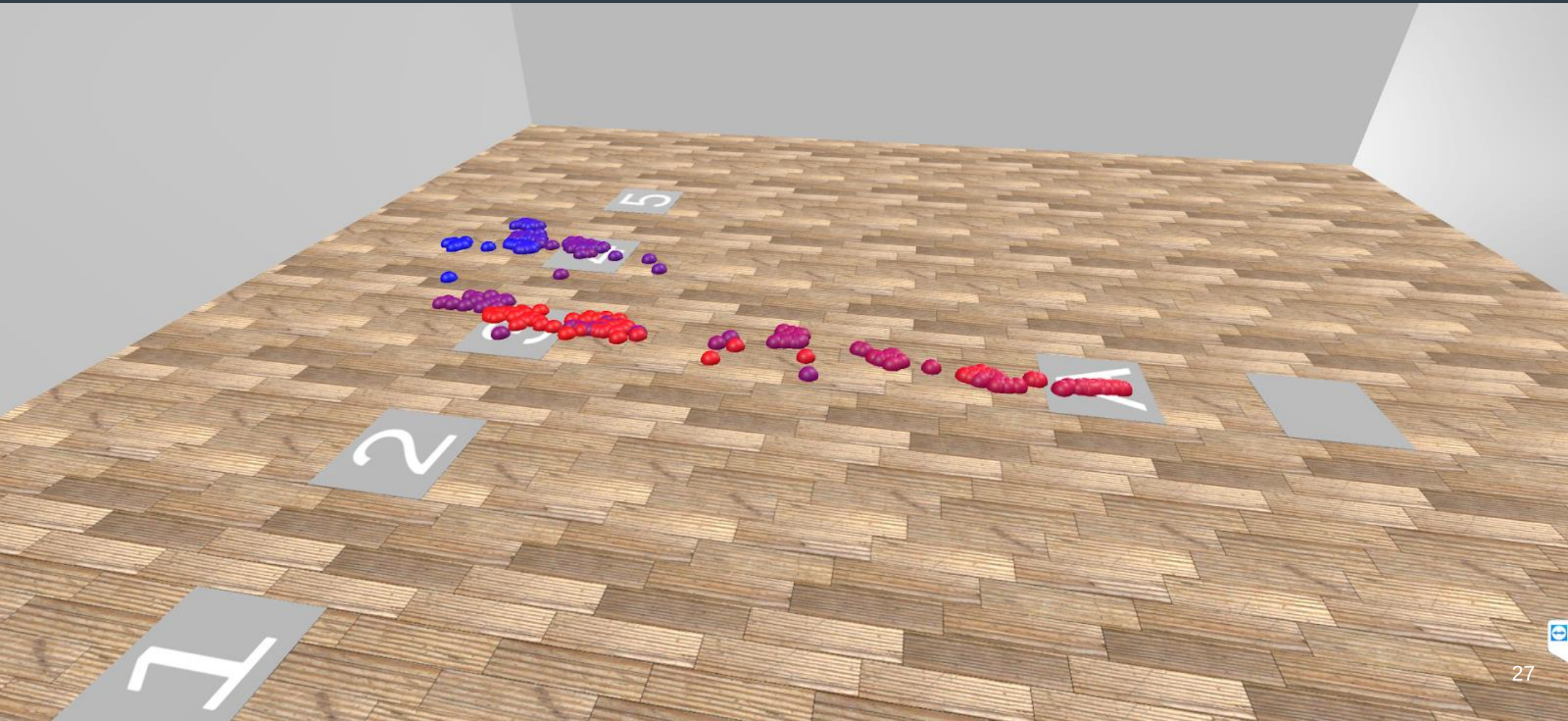
Gaze signal precedes all others



Gaze is a better predictor than other features



Gaze converges to goal target early in walking



Thank You!

Acknowledgement

This work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CPS-1739964, IIS-1724157, NRI-1925082), ONR (N00014-18-2243), FLI (RFP2-000), ARO (W911NF-19-2-0333), DARPA, Lockheed Martin, GM, and Bosch. Peter Stone serves as the Executive Director of Sony AI America and receives financial compensation for this work. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.