VISUAL ATTENTION IN SPOKEN HUMAN-ROBOT INTERACTION (HRI)

Maria Staudte & Matthew Crocker

HUMANS...

- In Language Production
  - Referential gaze about 1 s before speech onset
- In Language Comprehension
  - Listeners gaze at referent approx. 200-500 ms after offset
- Alignment, Joint attention
  - 2 sec span between speaker's and listener's gaze
  - Joint attention speeds this up

Research Questions

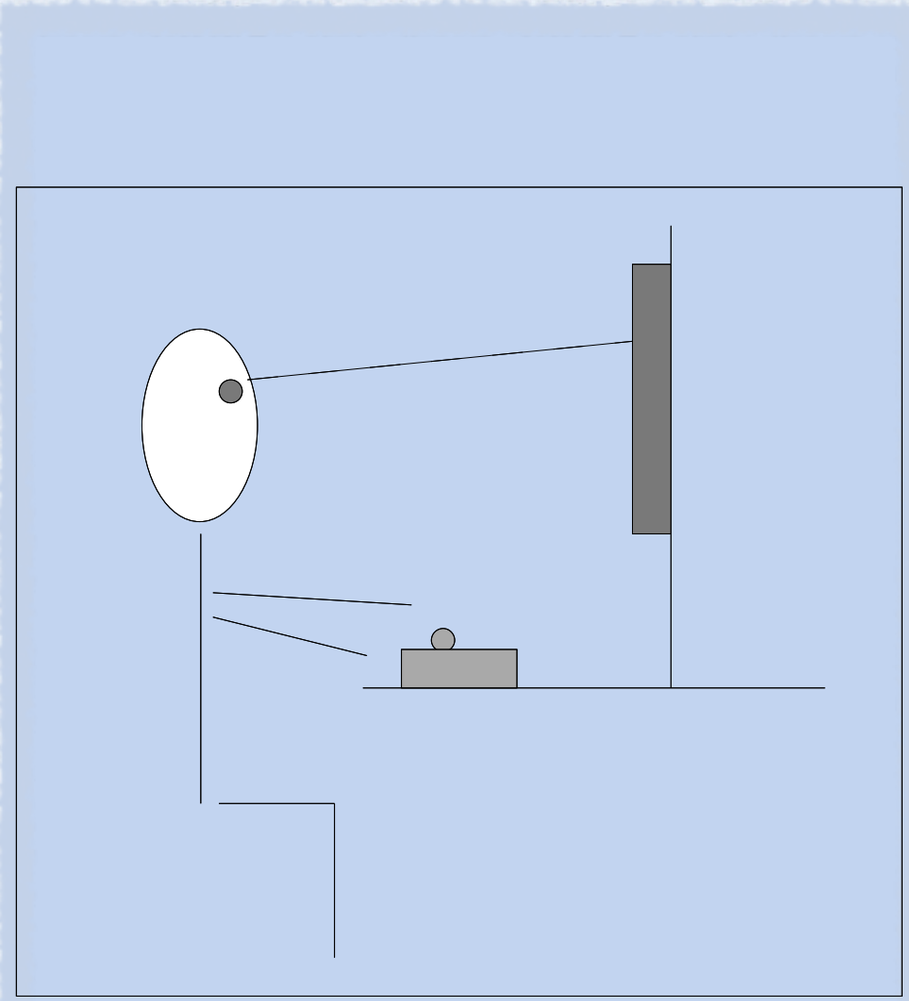
- Is robot gaze that reflects this pattern beneficial for human interaction partners?
- If so, what happens if this behaviour is inconsistent?

Methodology

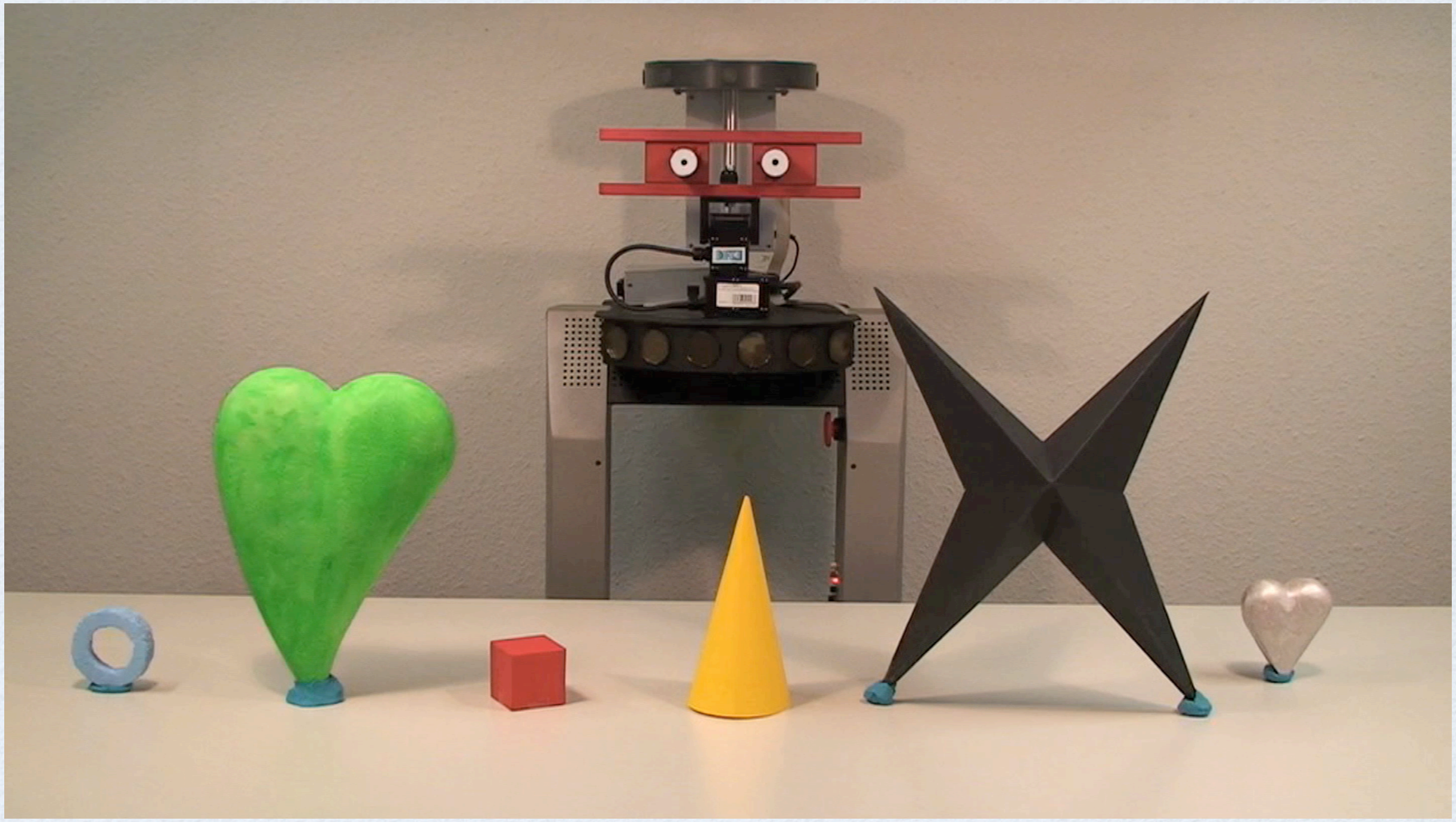
- Paradigm for measuring contribution of cognitively motivated robot gaze to HRI
  - Simulate *production* patterns from HHI
  - Apply to HRI
  - Monitor and evaluate participants' behaviour using precise methods from HHI

DESIGN

- Record video clip of robot producing utterances and gaze about scene
- Video is played back to participants
- Participants are being eye-tracked while observing robot
- Task: after each clip participant determines truth of the robot's statement and presses a button



"The cone is taller than the heart that is silver."

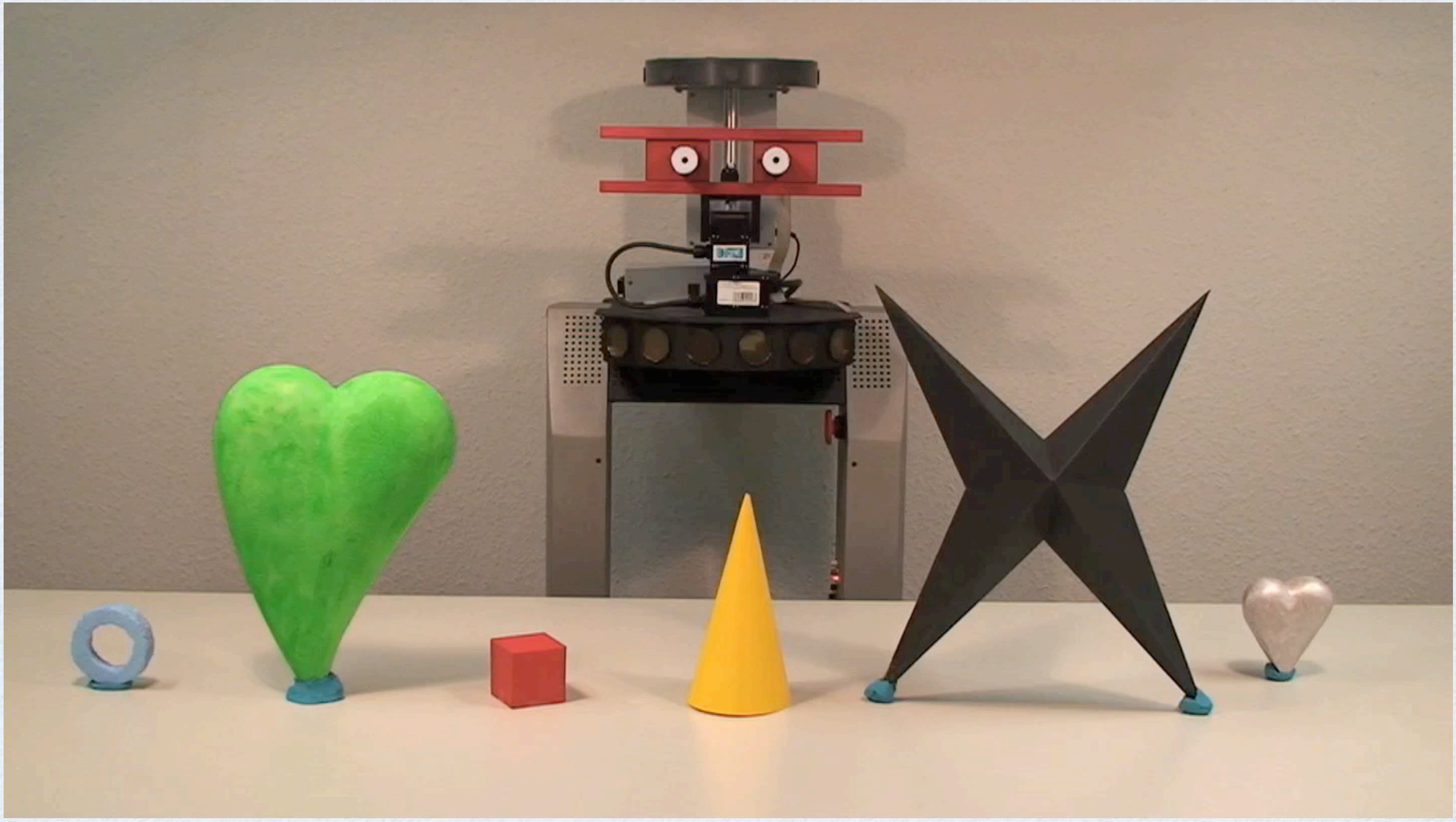


"The cone is taller than the heart that is silver."

DESIGN

- 2 Factors:
  - statement truth (true, false)
  - gaze congruency (congruent, incongruent, no gaze)
    - = 6 conditions
- 24 items, 48 fillers
- 48 participants

"The cone is taller than the heart that is silver."



"The cone is taller than the heart that is silver."

ANALYSIS

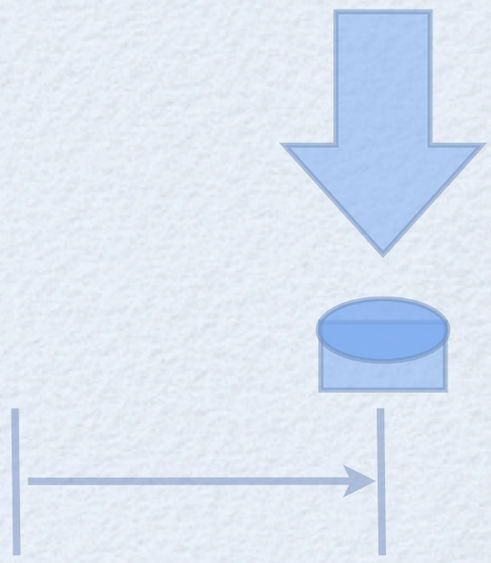
ANALYSIS

- Recording response times

ANALYSIS

- Recording response times

The cone is taller than the heart that is silver.



ANALYSIS

- Recording response times

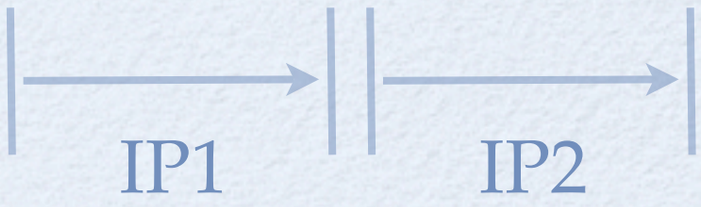
ANALYSIS

- Recording response times
- Segmentation into interest areas (IAs) and interest periods (IPs)

ANALYSIS

- Recording response times
- Segmentation into interest areas (IAs) and interest periods (IPs)

The cone is taller than the heart that is silver.



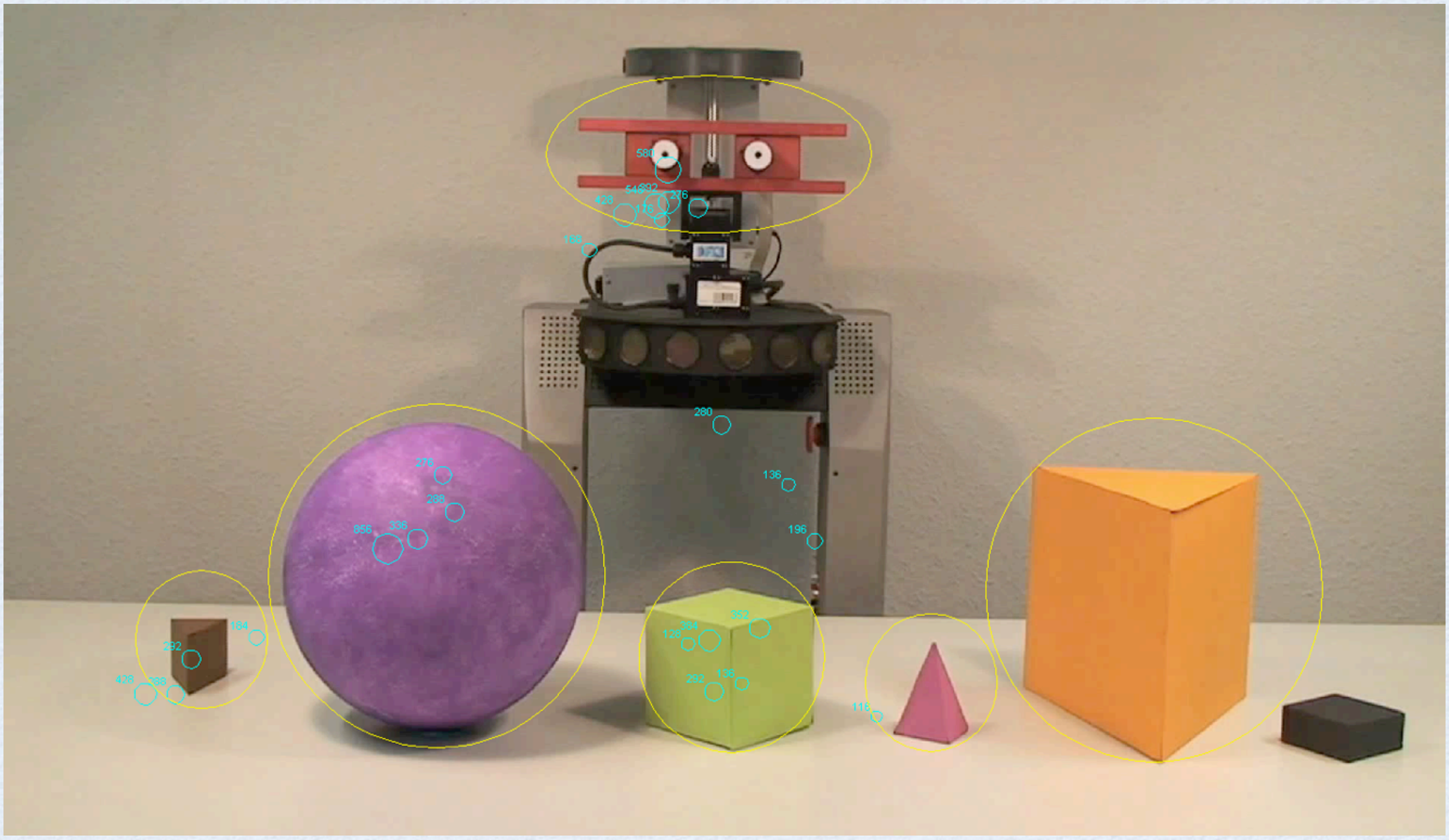
ANALYSIS

- Recording response times
- Segmentation into interest areas (IAs) and interest periods (IPs)

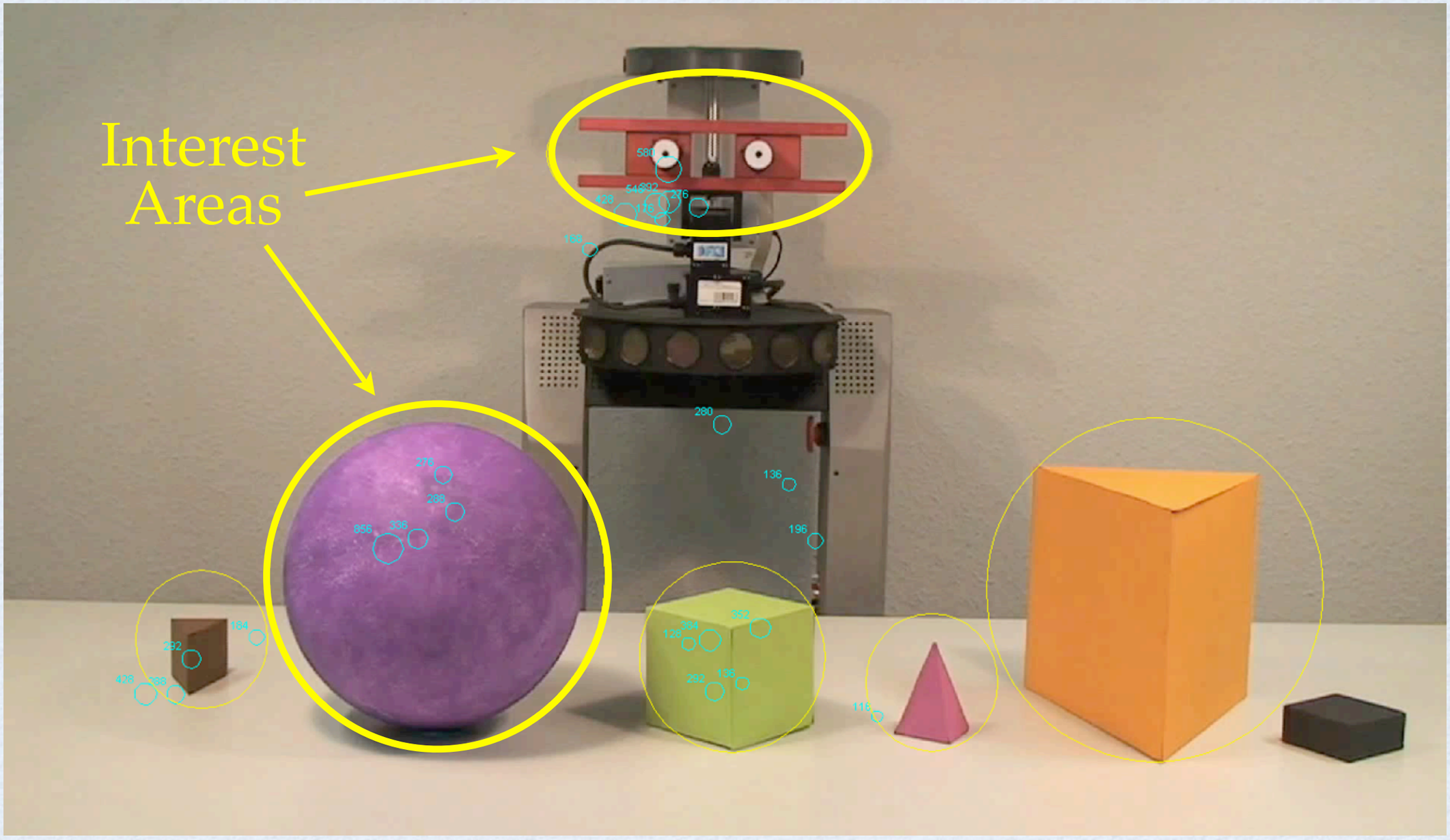
ANALYSIS

- Recording response times
- Segmentation into interest areas (IAs) and interest periods (IPs)
- Fixations per IA & IP

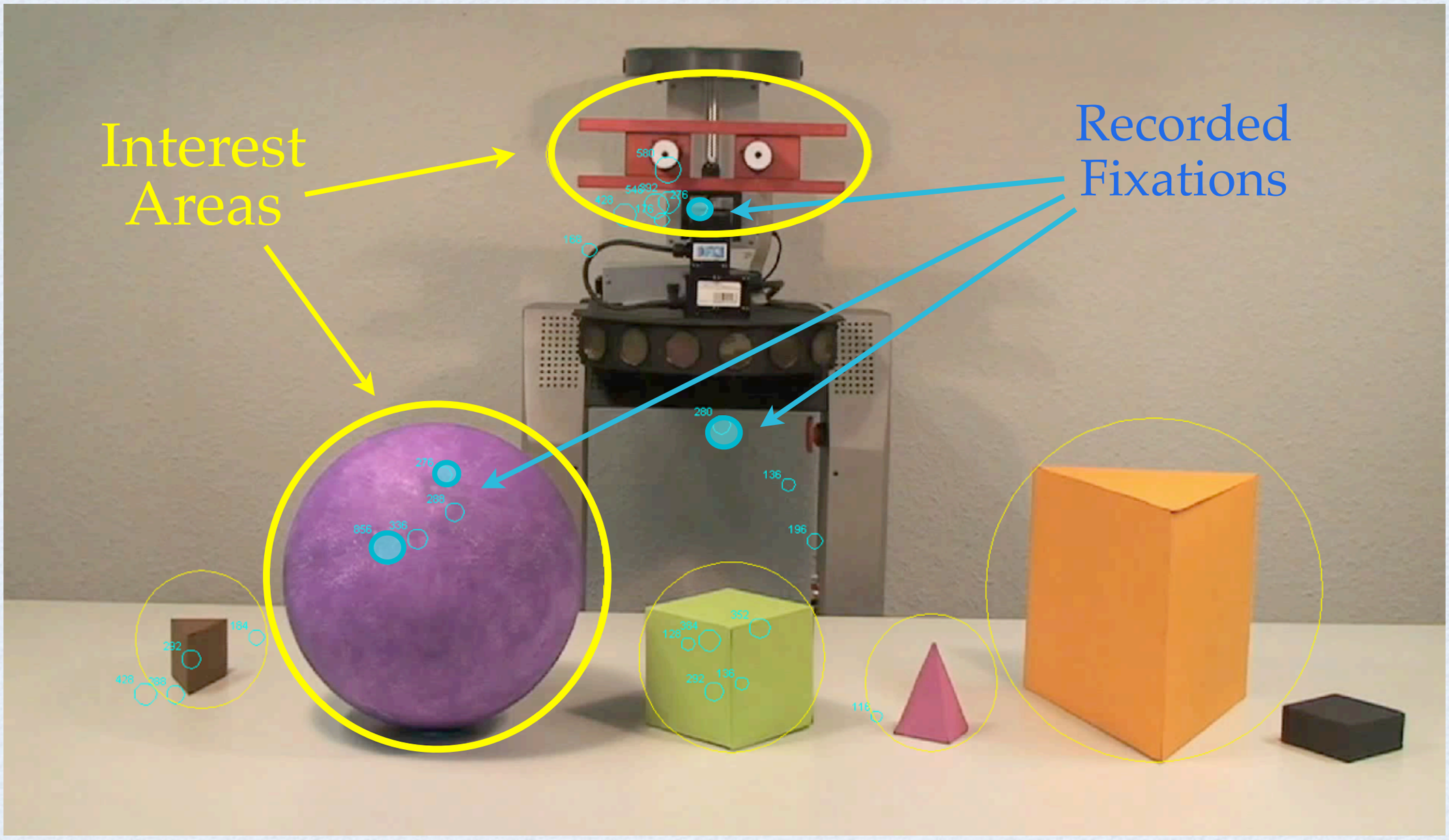
IA, IP & FIXATIONS



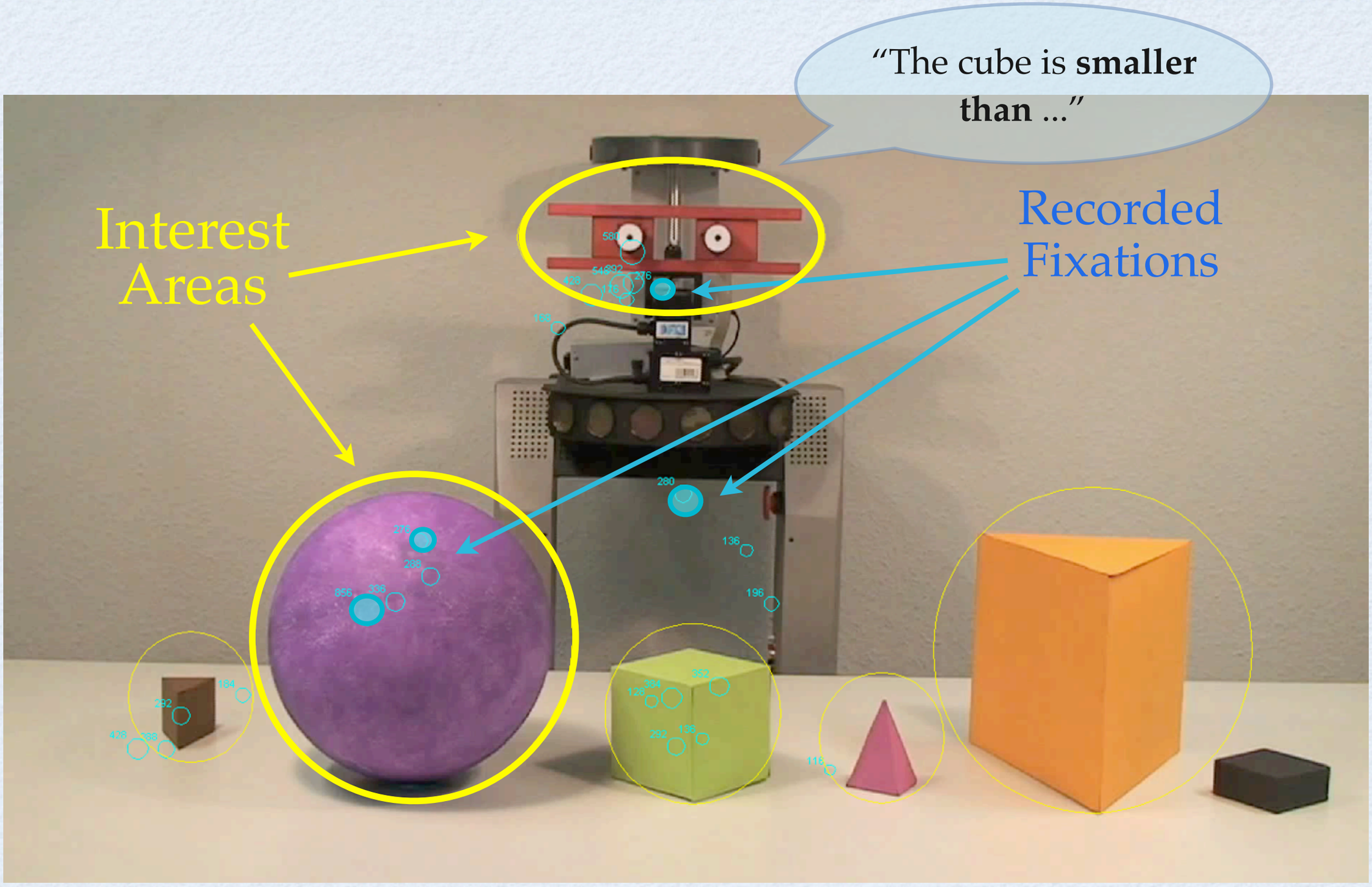
IA, IP & FIXATIONS



IA, IP & FIXATIONS

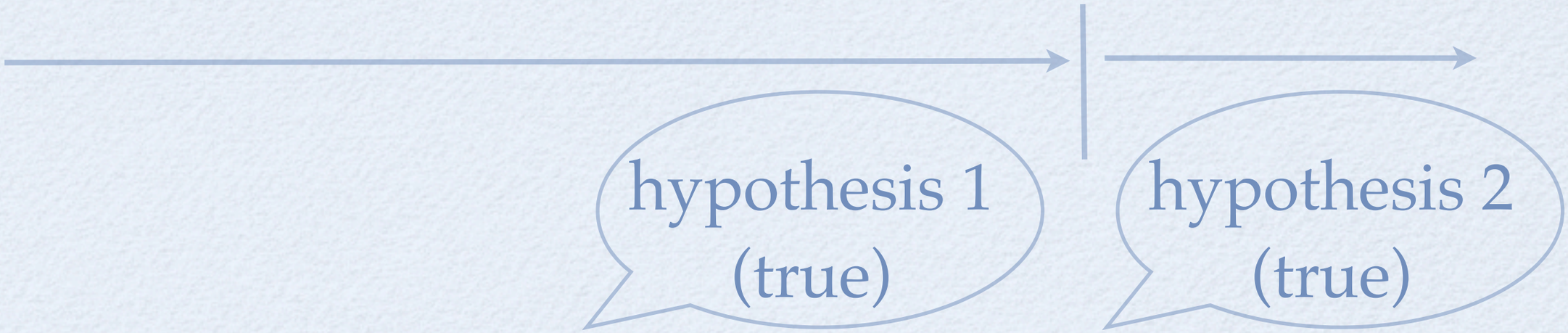


IA, IP & FIXATIONS



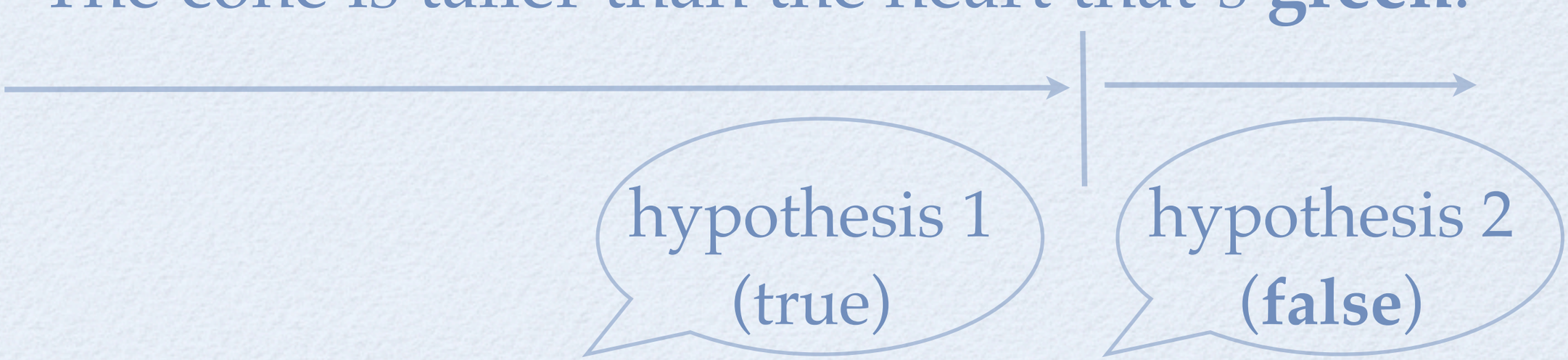
PREDICTIONS

"The cone is taller than the heart that's silver."



PREDICTIONS

"The cone is taller than the heart that's **green**."



PREDICTIONS

PREDICTIONS

- Response Times:

PREDICTIONS

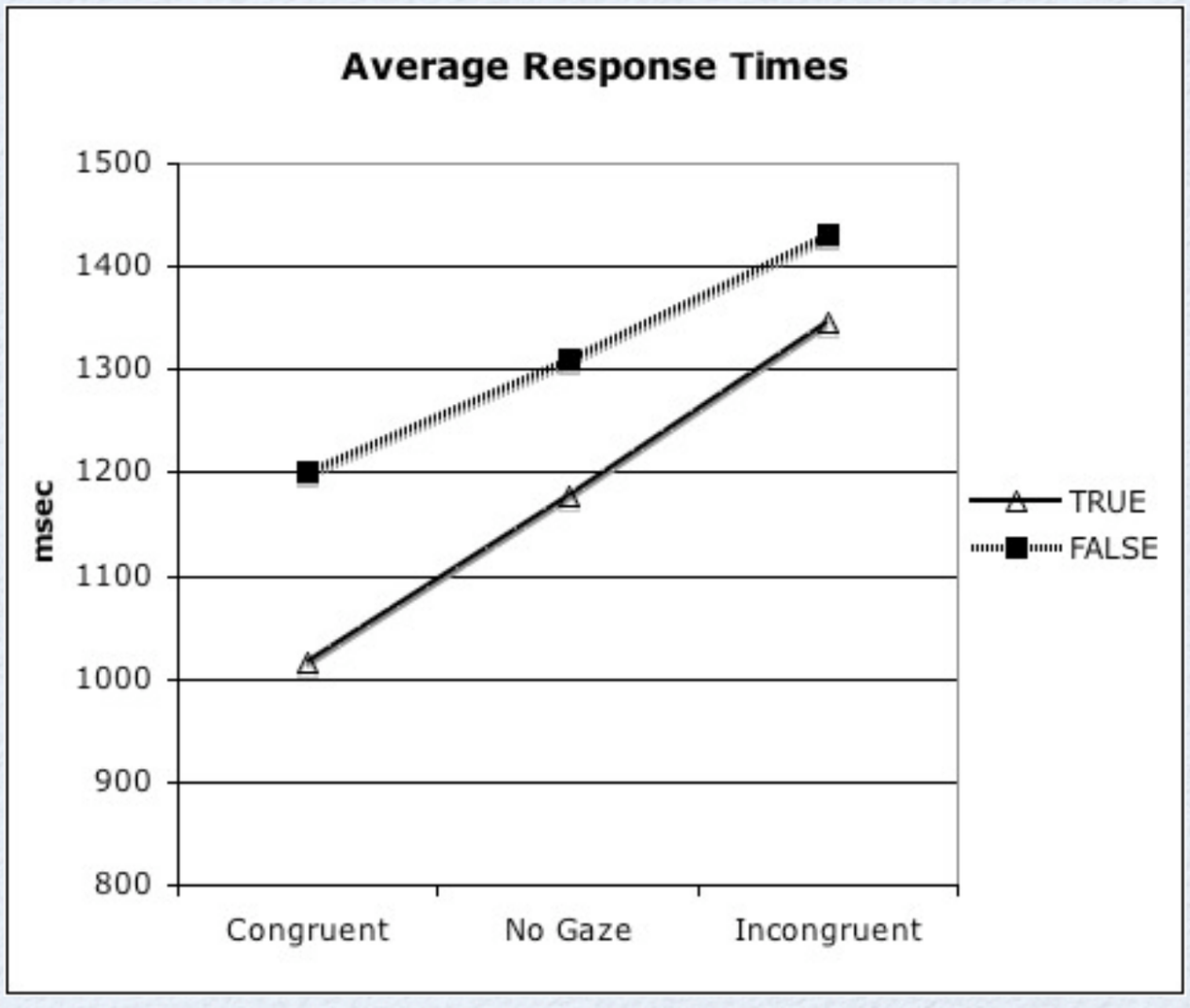
- Response Times:
  - true < false

PREDICTIONS

- Response Times:
  - true < false
  - congruent < no gaze < incongruent

RESULTS - EXPERIMENT 2

- Response Times:
  - true statement = shorter RT
  - congruent < no gaze < incongruent
  - tc = fastest



RESULTS - EXPERIMENT 2

RESULTS - EXPERIMENT 2

- Fixations

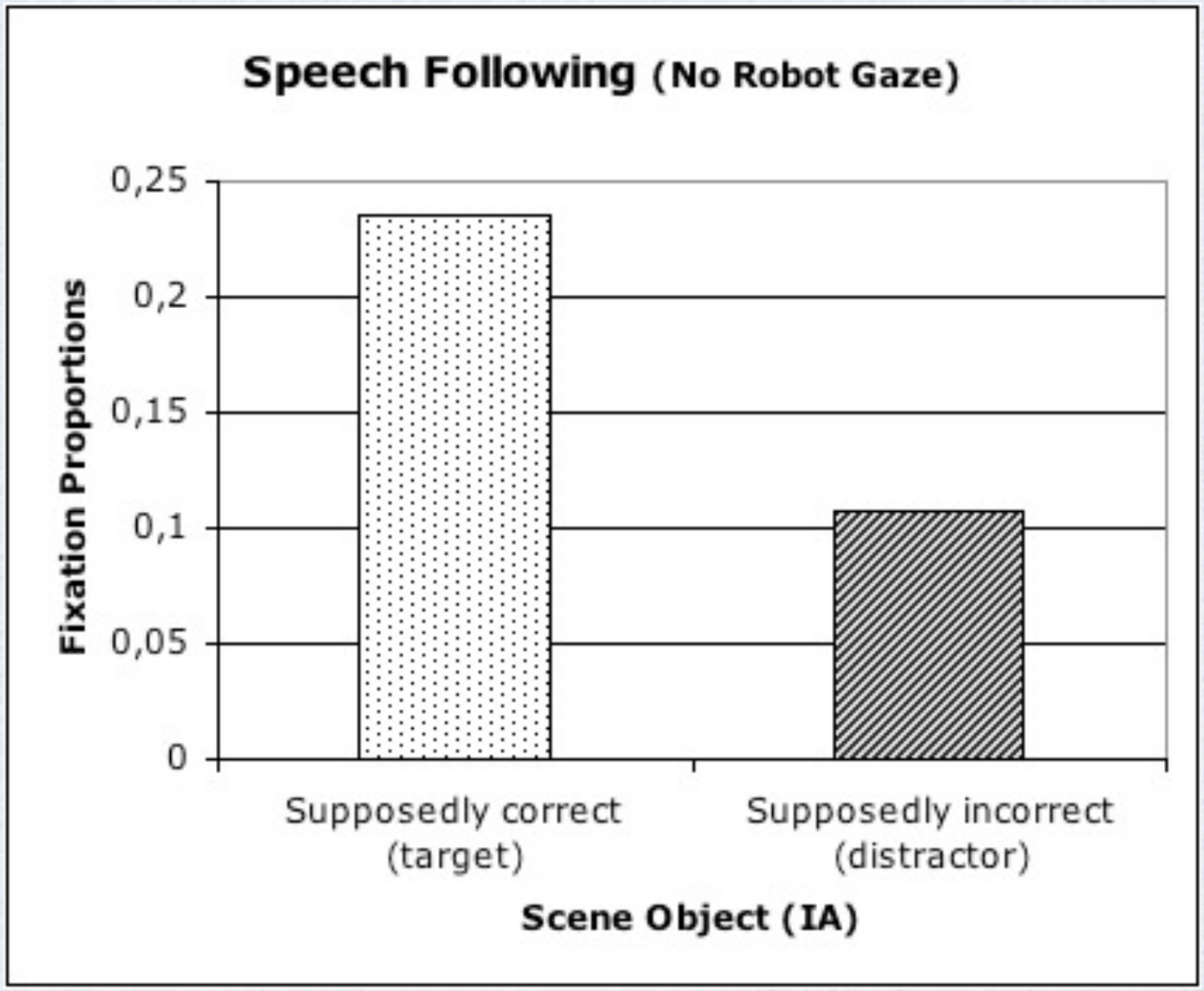
RESULTS - EXPERIMENT 2

- Fixations

- Utterance-driven gaze

RESULTS - EXPERIMENT 2

- Fixations
- Utterance-driven gaze



RESULTS - EXPERIMENT 2

- Fixations

- Utterance-driven gaze

RESULTS - EXPERIMENT 2

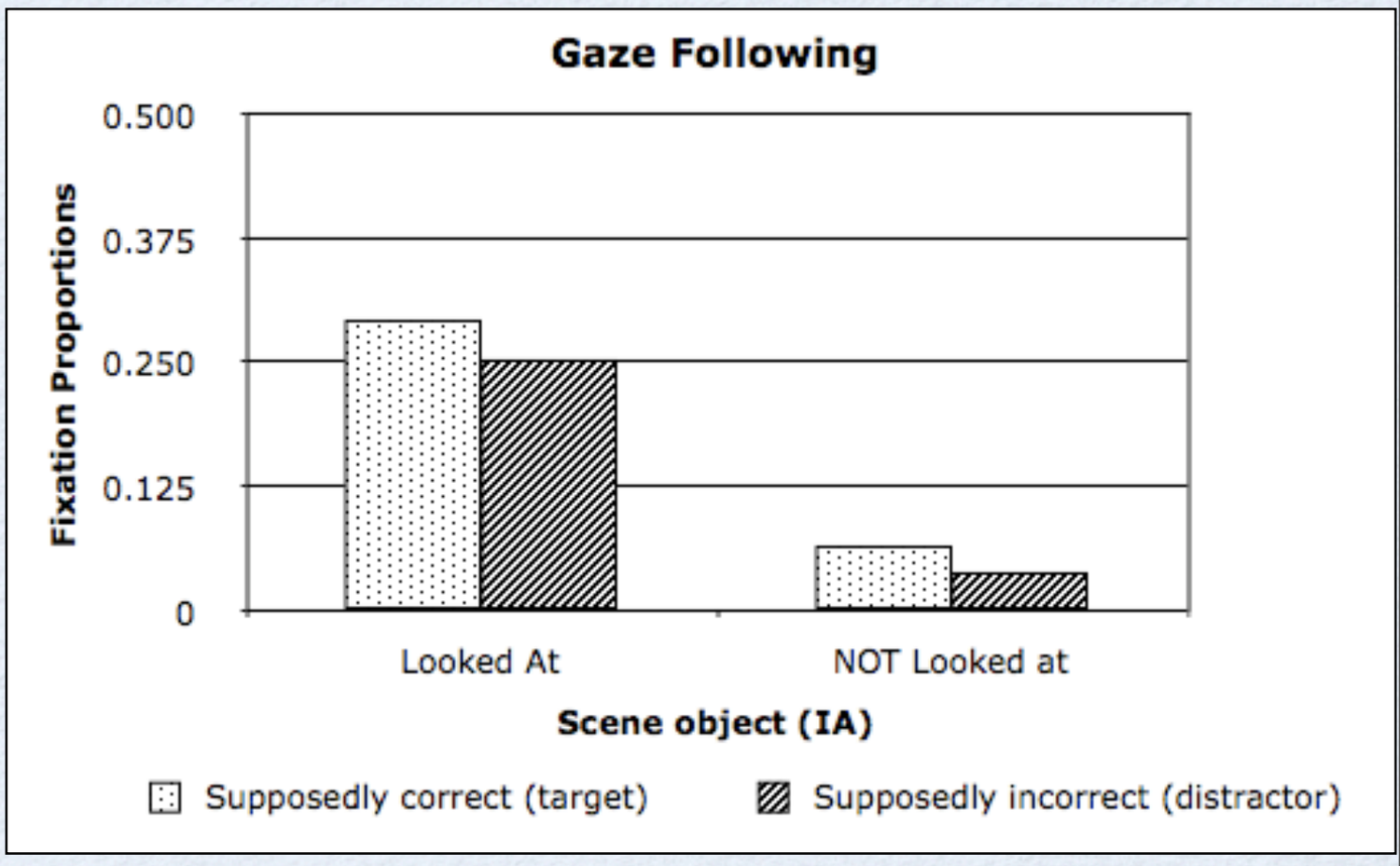
- Fixations

- Utterance-driven gaze

- Joint attention

RESULTS - EXPERIMENT 2

- Fixations
  - Utterance-driven gaze
  - Joint attention



CONCLUSIONS

- HRI behaviour is consistent with behaviour observed in HHI
  - Utterance-mediated gaze & joint attention
- Trust in robot capabilities, belief that robot gaze expresses visual attention of robot
- Robot gaze can be useful and can be irritating
- Robot should use its gaze with caution!