Who you lookin' at? Perception of gaze direction in group settings depends on naturalness of gaze behavior and clutter

Ruth Rosenholtz, Niall L. Williams, Michael Stengel, Andrew Russell

rrosenholtz@nvidia.com | niallw@umd.edu | mstengel@nvidia.com | arussell@nvidia.com

BACKGROUND & GOALS

Efficient estimation of another human's direction of gaze is important for perception of social cues. Typical online meetings inherently impair our ability to perceive such cues.

Previous work: Individual gaze direction
- Most research: viewer facing the camera. Threshold ~ 2° [3,4]
- Observed biases both towards (+) and away (-) from the subject
- Often unnatural, Photoshopped stimuli
- Todorović [1]: gaze perception = head pose + eyes relative to head
- People rotate eyes, head, and torso in regular ways to look at a target, Sidenmark & Gellersen [2] quantify this
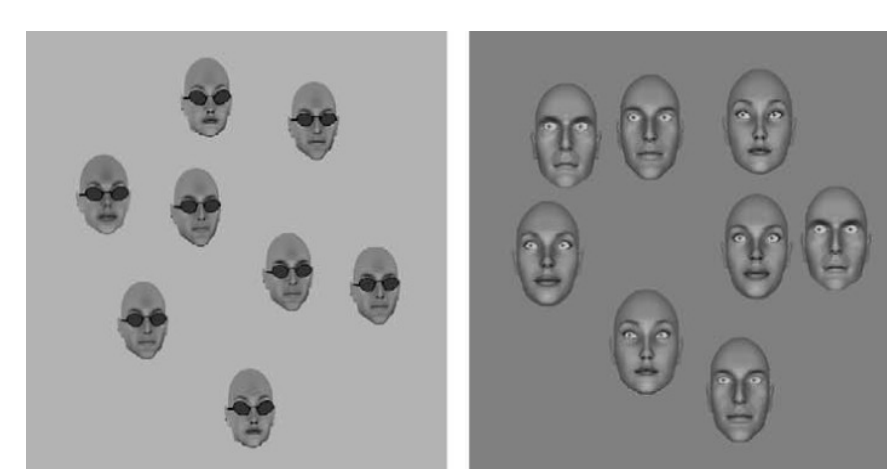
Questions:
- How accurate is gaze perception over a more natural range of viewing directions?
- Does naturalness of pose matter?
- Does a cluttered vs. sparse scene matter?

Previous work: Ensemble gaze perception
- Humans can efficiently extract ensemble properties, including of a set of people
- Mean gaze direction (left/right) in 300-1600 ms [5,6]
- Mean emotion and gender [7]
- Often 2D array of isolated faces w/ no visible/common gaze target, but see [8,9]
- Ensemble estimates discount outliers [10]

Questions:
- Accuracy estimating common gaze target?
- Accuracy detecting an outlier gaze?

METHODS

EXPERIMENT 1
- Where is the cued avatar looking?
- Natural and unnatural body poses
- Cluttered and sparse scenes
- 500ms display time or until response
- 40° horizontal field of view
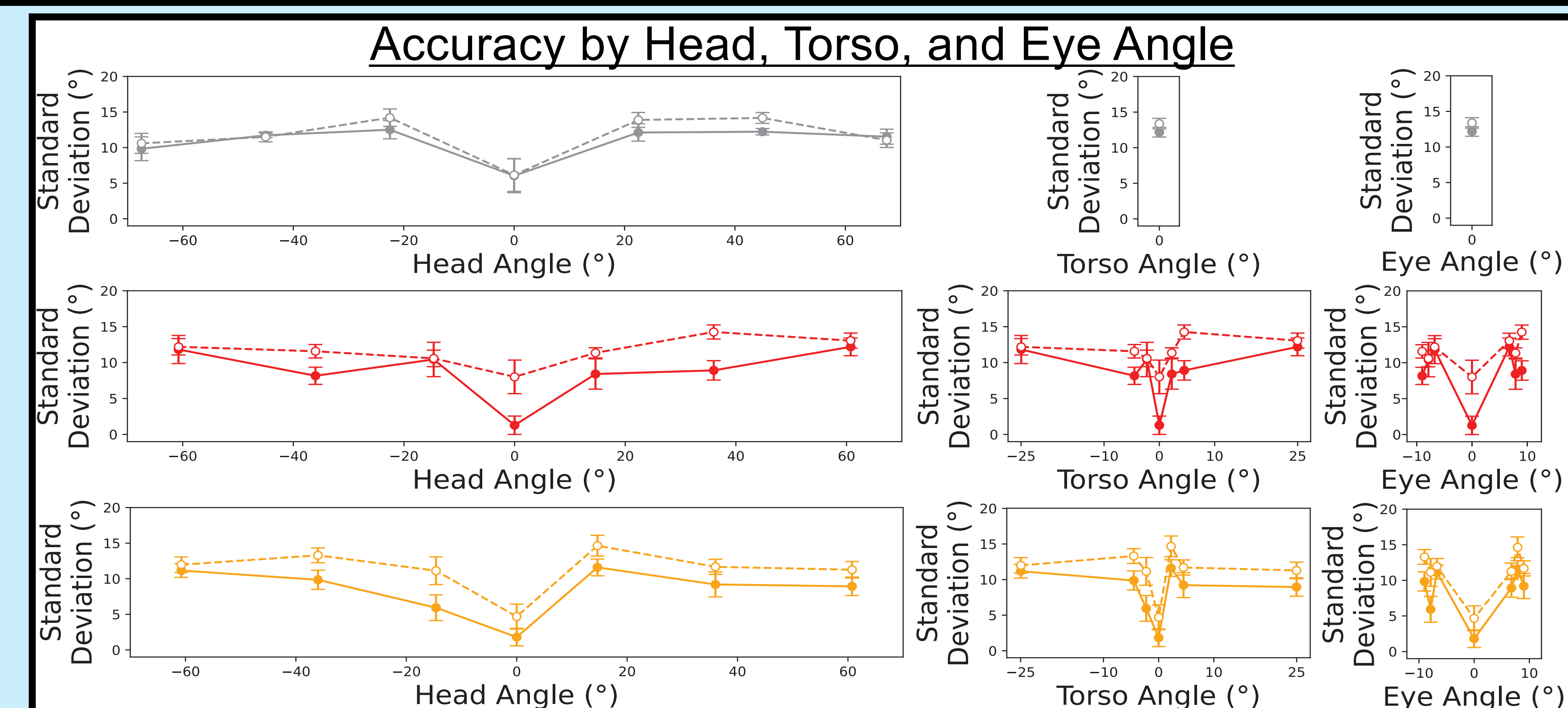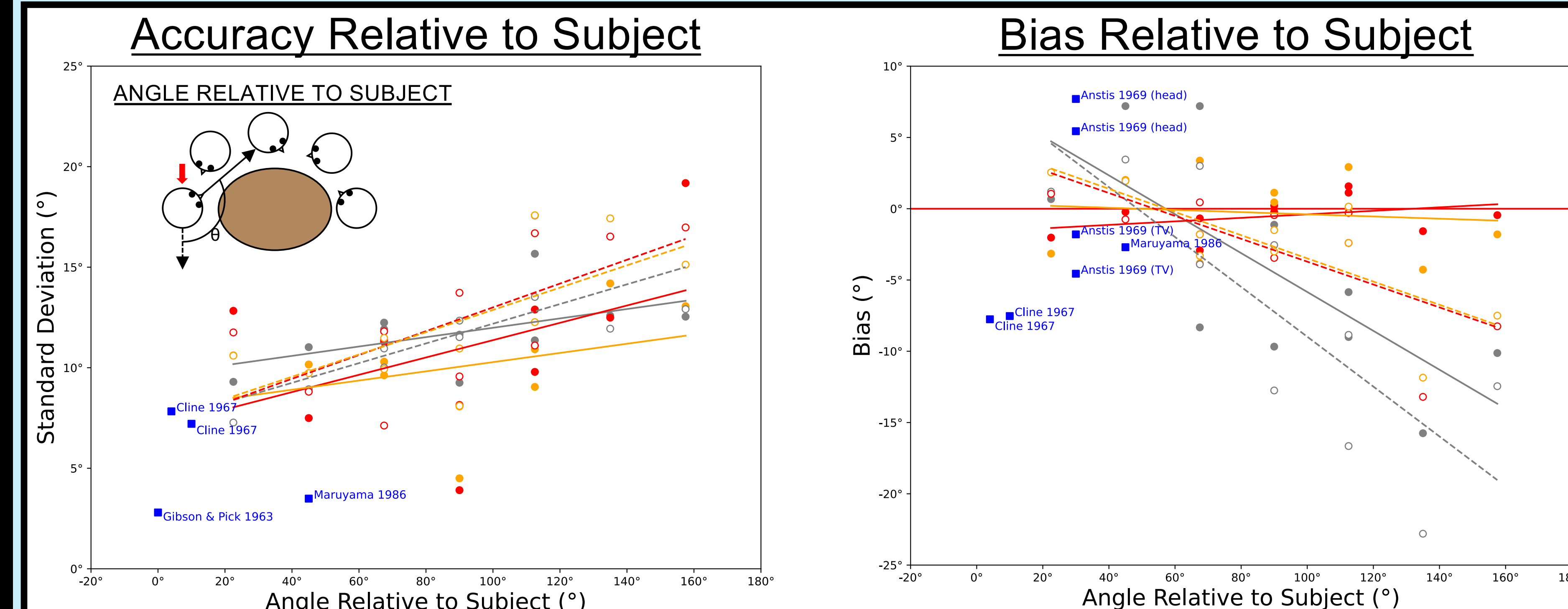- 300 trials for each participant (N = 25)

EXPERIMENT 2
- Which common target object are most avatars looking at? Who is the outlier?
- 3, 4, or 5 avatars in the scene
- 1600 ms display time
- 40° horizontal field of view
- 540 trials for each participant (N = 9)

RESULTS: INDIVIDUAL GAZE PERCEPTION

Gaze Direction Bias

Legend

Accuracy and Bias by Viewer-Target Pair

RESULTS: GAZE PERCEPTION BY BODY POSE

Accuracy Relative to Subject

Bias Relative to Subject

Accuracy by Head, Torso, and Eye Angle

RESULTS: ENSEMBLE GAZE PERCEPTION

Common Target Accuracy Depends on Target Offset

Offset by 1 Linear regression: r: 0.73930 p-value: 0.00001

Offset by 2 Linear regression: r: 0.36531 p-value: 0.06097

— Offset by 1
-- Offset by 2

Outlier Detection Accuracy Depends on Target Offset

Offset by 1 Linear regression: r: -0.48363 p-value: 0.01060

Offset by 2 Linear regression: r: -0.19678 p-value: 0.32522

— Offset by 1
-- Offset by 2

DISCUSSION

Experiment 1: Individual Gaze Perception
- Judgments are usually biased away from the camera, especially for unnatural avatar poses
- Rendering avatars with natural body poses improves subjects' gaze perception accuracy
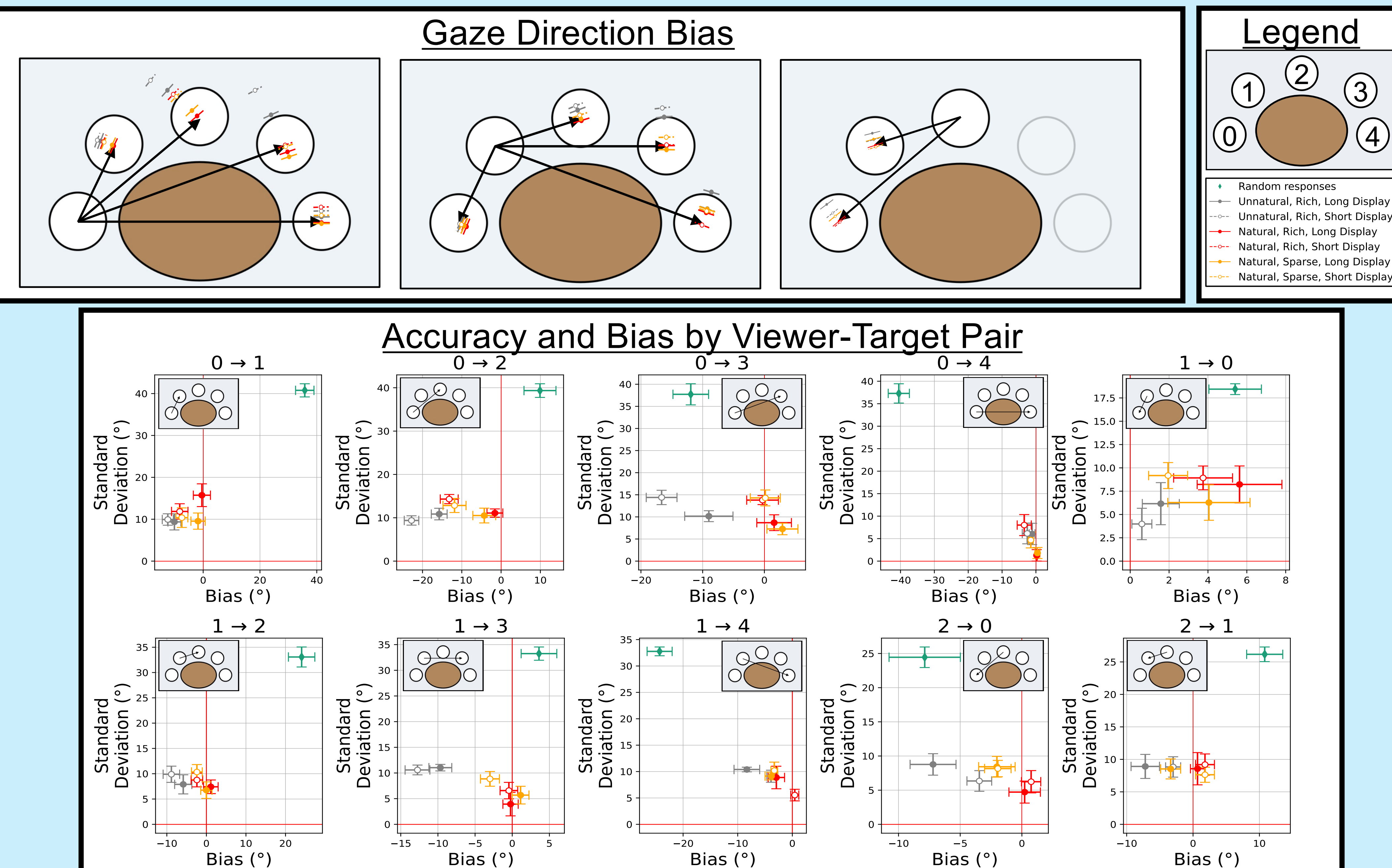
Experiment 1: Gaze Perception By Body Pose
- Accuracy decreases modestly as a function of avatar orientation relative to the subject
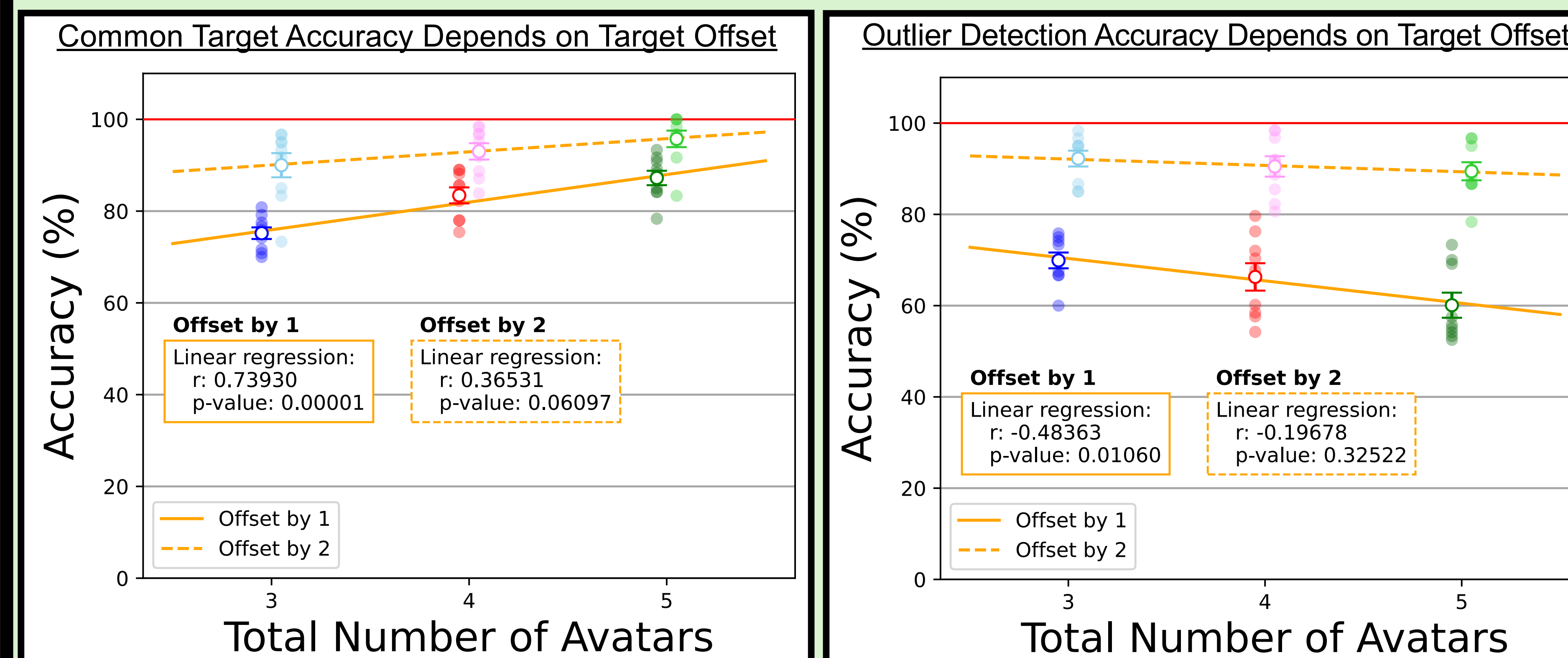
Experiment 2: Ensemble Gaze Perception
- Subjects are quite good at quickly judging common gaze target in rich, natural scenes
- Common target judgment performance improves as set size increases
- Outlier judgment performance decreases as set size increases
- Comparable to previously reported gaze-judgment performance (~5° − 15° threshold [4,5,8] vs. 8.6-17.2° target spacing)
- But comparison is not straight-forward due to lack of gaze-variability, more complex stimuli, and dual-task paradigm

REFERENCES

1. Todorović, Dejan. "Geometrical basis of perception of gaze direction." Vision Research 46.21 (2006)
2. Sidenmark, Ludwig, and Hans Gellersen. "Eye, head and torso coordination during gaze shifts in virtual reality." ACM Transactions on Computer-Human Interaction (TOCHI) 27.1 (2019)
3. Gibson, James J., and Anne D. Pick. "Perception of another person's looking behavior." The American Journal of Psychology 76.3 (1963)
4. Sweeny, Timothy D., and David Whitney. "Perceiving crowd attention: Ensemble perception of a crowd's gaze." Psychological Science (2014)
5. Florey, Joseph, et al. "Spatial limitations in averaging social cues." Scientific Reports 6.1 (2016)
6. Haberman, Jason, and David Whitney. "Rapid extraction of mean emotion and gender from sets of faces." Current Biology 17.17 (2007)
7. Capozzi, Francesca, Andrew P. Bayliss, and Jelena Ristic. "Gaze following in multiagent contexts: Evidence for a quorum-like principle." Psychonomic Bulletin & Review 25 (2018)
8. Florey, Joseph, Steven C. Dakin, and Isabelle Mareschal. "Comparing averaging limits for social cues over space and time." J. Vision 17.9 (2017)
9. Capozzi, Francesca, Andrew P. Bayliss, and Jelena Ristic. "Gaze following in multiagent contexts: Evidence for a quorum-like principle." Psychonomic Bulletin & Review (2018)
10. Corbett, Jennifer E., Igor Utochkin, and Shaul Hochstein. The pervasiveness of ensemble perception: Not just your average review. Cambridge University Press, 2023.
11. Anstis, Stuart M., John W. Mayhew, and Tania Morley. "The perception of where a face or television 'portrait' is looking." The American journal of psychology 82.4 (1969): 474-489.
12. Maruyama, Kinya, and Mitsuo Endo. The effect of face orientation upon apparent direction of gaze. Diss. Tohoku University, 1984.
13. Maruyama, Kinya, Mitsuo Endo, and Kenzo Sakurai. An Experimental Consideration on" Mona Lisa Gaze Effect". Diss. Tohoku University, 1986.