Time Course of Person Recognition in a Naturalistic Environment

Carina A. Hahn[1], Eric Hart[1], Kate Flanagan[1], P. Jonathon Phillips[2], & Alice J. O'Toole[1]
[1]The University of Texas at Dallas, Richardson, TX
[2]National Institute of Standards and Technology, Gaithersburg, MD

Background

Person recognition work has concentrated mostly on the face
- In natural viewing environments, we see *whole people in motion*

Goal: Examine time course of person recognition in natural viewing environments

Part of 3 part project:
1) *How do we recognize people in natural environments?*
2) Computer face recognition technology
   - Recent work includes the body (Chen et al., In press)
3) Anatomically distinct neural areas selective for faces and bodies
   - FFA (Kanwisher et al., 1997), OFA (Halgren et al., 1999; Puce et al., 1996; Pitcher et al., 2006), STS (Allison et al., 2000), EBA (Downing et al., 2001), FBA (Peelen & Downing, 2005a)

Information that supports person recognition: face, body, gait
- How do we combine this information to recognize people in the real world?
- Face & body contributions to recognition (Burton et al., 1999; Rice et al., accepted; Robbins & Coltheart, 2012)
- Gait (Loula et al., 2005)

Present Approach:
Build robust representation of identities -> test recognition *with two tests to examine time course of recognition*
- Possible scenarios:
  1. Judgments and *decisions gradually build* over time
  2: Wait until clear view of face to make decision
  3: Dependent on video/identity

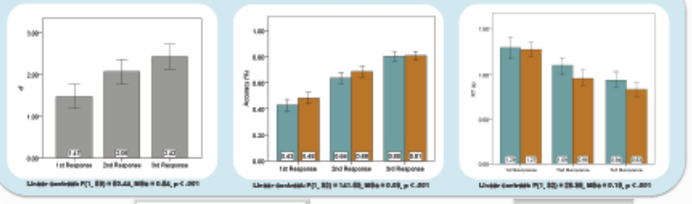
Results

1) Prompted Responses

How do responses evolve over time?

For all graphs: Unfamiliar, Familiar
All error bars: CI 95%

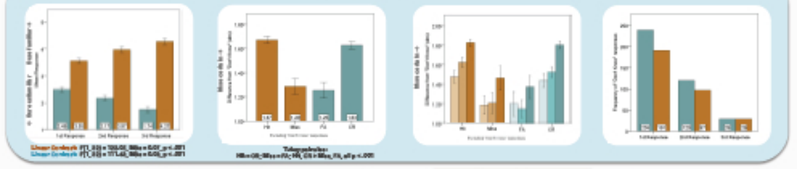
Accuracy & Speed



More accurate over time

Faster decisions at each response

Confidence



More confident over time

Higher confidence when correct...

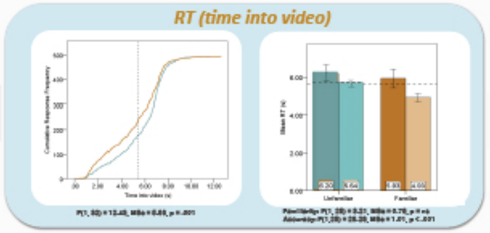
Confidence increases over time, even when incorrect

Fewer "Don't Know" responses over time

2) Free Response

How are decisions made freely?

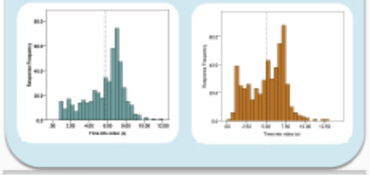
RT (time into video)



Familiar decision made faster (M = 6.68s) made faster than unfamiliar (M = 6.77s)

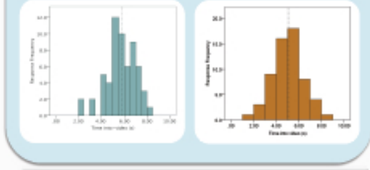
Correct decisions (right here) made faster than incorrect (left here)

RT Distributions: By Trial (accurate only)



Unfamiliar decision: Slower, wait longer to decide
Familiar decisions: Either fast, or wait– not much intermediate

RT Distributions: By Video (accurate only)

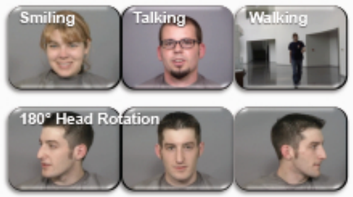


Unfamiliar identities: Fewer recognized early
Familiar identities: Normally distributed recognition time

Method

Training

- Familiarized with 30 identities
  - 4 motion-based actions
  - 120 videos total



Recognition Testing

Tested on 60 identities (counterbalanced) – 30 familiarized/ 30 novel identities – Videos from different day than training

1) Prompted Response

At each third, indicated if identity was unfamiliar to familiar on scale of 1 to 5 (sure unfamiliar to sure familiar)

2) Free Response

- Videos play continuously
- Participants responded "as soon as they felt confident"
- No explicit confidence rating
- Binary choice: Familiar or Unfamiliar

Conclusions

Familiar:
- Faster *decisions*, on average
  - 2 types of recognition (fast and slow)
- Recognition of *individual identities* normally distributed
- Accuracy and confidence build

Unfamiliar:
- Wait longer to make decisions
- Skewed distribution: few decisions made quickly

Confidence and accuracy:
- Equally confident rejecting and recognizing
- More confident/faster when correct even at earliest time point
- No speed accuracy tradeoff
  - High resolution view of face not always required

Open questions:
- Understanding performance improvement as changes in quality of information and/or information accumulation
- What is relative contribution at each time point of face, body, & gait?

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: The role of the STS region. Trends in Cognitive Science, 4(7), 267–278.
Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video. Psychological Science, 10, 243–248.
Chen, Y., Patel, D. M., Phillips, P. J., & Cherniaga, R. (accepted). Dictionary-based face and person recognition from unconstrained video. IEEE Transactions on Pattern Analysis and Machine Intelligence.
Downing, P.E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. Science, 293 (5539), 2470–2473.
Duchaine, B., Pitcher, D., & Walsh, V. (2011). The role of the occipital face area in the cortical face perception network. Experimental Brain Research, 209(4), 481–493.
Halgren, E., Dale, A.M., Sereno, M. I., Tootell, R. B. H., Marinkovic, K., & Rosen, B. R. (1999). Location of human face-selective cortex with respect to retinotopic areas. Human Brain Mapping, 7, 29–37.
Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movements. Journal of Experimental Psychology: Human Perception and Performance, 31, 210–220.
Powell, M.S., & Cowting, P.E. (2006a). Selectivity for the human body in the fusiform gyrus. Journal of Neurophysiology, 95 (1), 550–555.
Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. J Neurosci 16, 5205–5215.
Rice, A., Phillips, P. J., & O'Toole, A. J. (accepted). The role of the face and body in person identification. Psychological Science.
Robbins, R. A., & Coltheart, M. (2012). Left = right holistic integration of human bodies. The Quarterly Journal of Experimental Psychology, 65(10), 1950–1954.