Variable use of the face and body in person identification

Allyson Rice[1], Jonathon Phillips[2], & Alice O' Toole[1]
[1]The University of Texas at Dallas
[2]National Institute of Standards and Technology

BACKGROUND

Person recognition from body
- compare face, body, & person (O'Toole et al., 2011)
  - static vs. dynamic presentations
  - identification best for whole person in motion

- body in motion – point-light (Prasad & Shiffrar, 2009)
  - accurate identification of self & friend but not strangers

- face more important than body
  - composite images (Robbins & Coltheart, 2011)
  - for familiar people (Burton et al., 1999)

People use the body, when face fails (Rice et al., In press)
- body accounts for performance with entire image
- eye-movement - attention shifts to the body when "face fails" – adaptive and efficient
- subjective ratings -> people are unaware of this

- human vs. algorithms (O'Toole et al., 2012)
  - face-recognition algorithms ranked *face quality* as *good, moderate, or poor.*
  - good and moderate: algorithms > humans
  - poor: algorithms = humans

GOALS

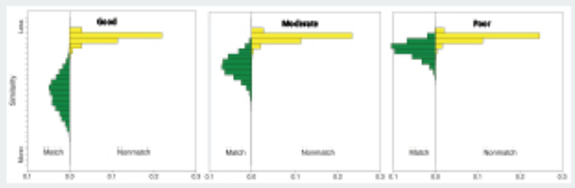
- dissect human performance with face- or body-only
- vary quality of information in face
  - "quality" - defined by face-recognition algorithms

- Algorithms use only the face
  - What are humans using? → measure recognition
  - What are they looking at? → measure eye-movements

- subset of images – variable *face quality* (O'Toole et al., 2012)

STIMULI

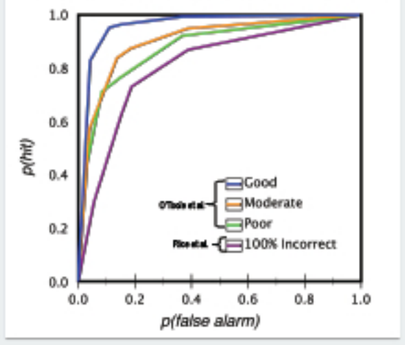
Face Recognition Vendor Test (FRVT 2006) (Phillips et al., 2010)
- International competition
- National Institute of Standards & Technology

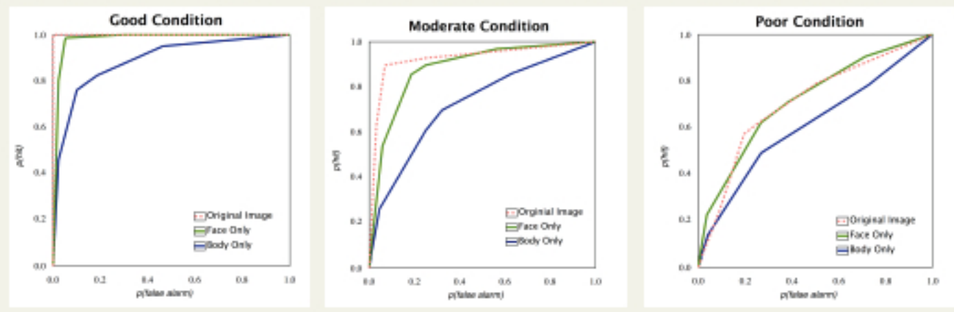


FRVT 2006 (fusion of 3 top performing algorithms)
- algorithms match identity in pairs of images
  - both images uncontrolled illumination
  - (n = 1,177,225)
- similarity score matrix, where $s_{i,j}$
  - computed similarity between i^{th} and j^{th} image
- algorithm fusion
  - fused similarity score S_f over *(k = 1:3)* algorithms
  - $S_f = S_k\ (s_k - \text{median}_k)/\text{MAD}_k$
    - MAD_k *mean absolute deviation for the* k^{th} *algorithm*

Partitioning of scores
- similarity scores for each individual ranked
  - good – top 1/3
  - moderate – middle 1/3
  - poor – bottom 1/3

Whole Person – Identification Performance

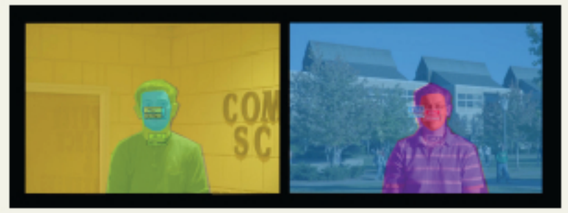




- face > body $F(1, 42) = 58.55$, $p < .0001$
- good > moderate > poor $F(2, 84) = 96.86$. $p < .0001$

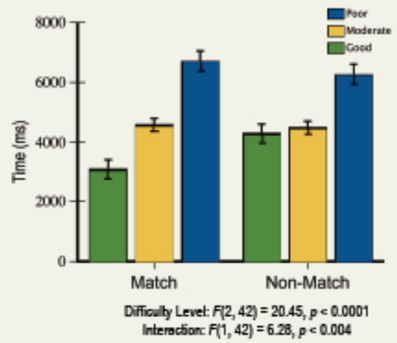
- As difficulty increases, the advantage for the face over body decreases $F(2, 84) = 8.28$, $p < .0006$

EYE TRACKING



Face Bias = $\frac{\textit{gaze time to internal features}}{\textit{total time to person}}$

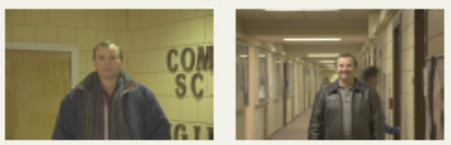
Overall Duration



Difficulty Level: $F(2, 42) = 20.45$, $p < 0.0001$
Interaction: $F(1, 42) = 6.28$, $p < 0.004$

EXPERIMENTS

Sample Trial



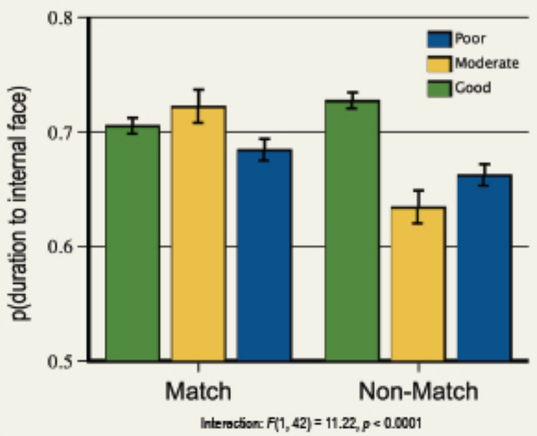
1) Sure they are the same
2) Think they are the same
3) Don't Know
4) Think they are different
5) Sure they are different

Matching Identity (60 identity pairs, 20 good, 20 moderate, and 20 poor)
- Exp 1 – face only (n = 22)
- Exp 2 – body only (n = 22)
- Exp 3 – whole person, eye tracking (n = 22)

Face Only Image (Exp. 1)

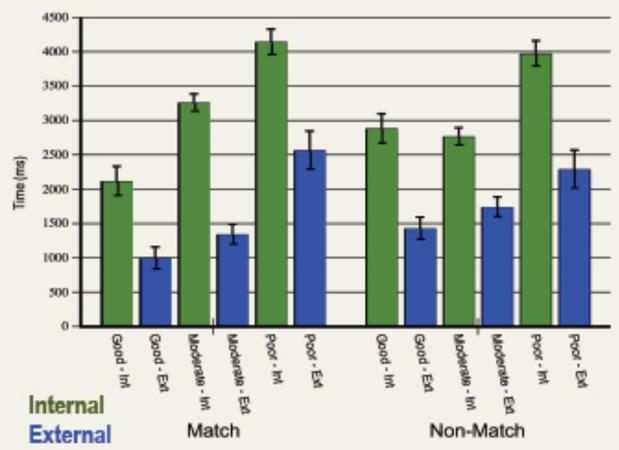
Body Only Image (Exp. 2)

Internal Face Gaze Duration Bias



Interaction: $F(1, 42) = 11.22$, $p < 0.0001$

Gaze Duration to Internal vs to External



Total Gaze Duration to Internal Face

Difficulty Level:
- $F(2, 42) = 15.86$, $p < 0.0001$

Interaction:
- $F(1, 42) = 10.33$, $p < 0.0002$

Total Gaze Duration to External Face & Body

Difficulty Level:
- $F(2, 42) = 6.42$, $p < 0.004$

Match Status:
- $F(1, 42) = 4.92$, $p < 0.04$

RESULTS SUMMARY

- **Accuracy**
  - overall, face > body, but decreases with increasing difficulty
  - algorithm face quality stratifications predict human performance
    i.e. good > moderate > poor
- **Eye tracker**
  - *match*: less face bias for poor condition, $F(2, 42) = 3.19$, $p < 0.05$
  - *non-match*: less face bias for moderate and poor conditions, $F(2, 42) = 13.35$, $p < 0.0001$

CONCLUSION

- algorithm stratification affected:
  - human accuracy
  - time spent on task
  - *face bias* as measured by eye movements
    - strong, but affected by quality of information in the face
    - eye movements do not shift optimally.

References
- Burton, A.M., Wilson, S., Cowan, M. & Bruce, V. (1999). Face recognition in poor-quality video. *Psychological Sciences, 10.* 243-248
- O' Toole, A., Phillips, J.P., Weimer, S., Roark, D.A., Ayyad, J., Barwick, R. & Dunlop, J. (2011) Recognizing people from dynamic and static face and bodies: Dissecting identity with a fusion approach. *Vision Research. 51.* 74-83
- O'Toole, A., An, X., Dunlop, J., Natu, V., & Phillips, J.P. (2012) Comparing face recognition algorithms to humans on challenging tasks. *ACM Transactions on Applied Perception, 9.*
- Prasad, S. & Shiffrar, M. (2009). Viewpoint and the recognition of people from their movements. *Journal of Experimental Psychology: Human Perception and Performance. 35.* 39-49
- Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. (in press). Unaware person recognition from the body when face identification fails. *Psychological Science.*
- Robbins, R & Coltheart, M. (in press). Head, bodies and holistic processing in person recognition. *Journal of Vision.*

Acknowledgements
Work funded by the Technical Support Working Group of the DOD