Qualitative differences between professional forensic face examiners and untrained people in person recognition revealed by item analysis

Ying Hu[1], Kelsey M. Jackson[1], Amy Yates[2], David White[3], P. Jonathon Phillips[2] & Alice J. O'Toole[1]

[1]*The University of Texas at Dallas, USA;* [2]*National Institute of Standards & Technology, USA;* [3]*University of New South Wales, Australia*

Introduction

Goal

- understand face recognition by forensic face examiners at a qualitative/strategic level

Background

- professional forensic face examiners surpass untrained participant groups on challenging face identity-matching tasks (White et al., 2015)
- quantitative performance comparisons are informative, but incomplete

Approach

- use data from 2 tests of professional forensic examiners and controls (White et al., 2015)
- quantify performance between groups by looking at patterns of errors

Base Study (White et al., 2015)

Participants

- **examiners**: international group of professional forensic face examiners (*n* = 27)
- **FISWG controls**: Facial Identification Scientific Working Group (FISWG) attendees, face recognition policy, but not trained (*n* = 14)
- **student controls**: untrained undergrads, typical sample in literature (*n* = 32)

Identity Matching Paradigm

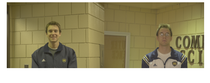
1.) sure same person; 2.) think same; 3.) don't know; 4.) think different; 5.) sure different people

Identity Matching Tests

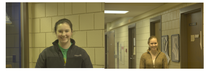
Person Identification Challenge Test (PICT)

- reliance on face versus body for identification
- screened by computer algorithm to represent a worst-case scenario

Matched identity (low similarity)



Non-matched identity (high similarity)



whole face body

- Image pair type
  - face- informative = whole (info: face > body)
  - body-informative = whole (info: body > face)

(Rice et al., 2013)

Expertise in Facial Comparison Test (EFCT)

- strategic differences with same- versus different- identity trials
- selected to be challenging for computers & humans

Matched identity

Non-matched identity

(Phillips et al., 2012)

Methods

Item Winner Analysis

- formulate the analysis of each item as a competition between the three groups

define ideal response: same-identity: "1" (sure same) different-identity: "5" (sure different) ⇨ challenge level score = absolute (actual response - ideal response) ⇨ item winner = group with the lowest challenge level score

Fusion Analysis

- examine the "wisdom-of-the-crowd" effect as a function of varying sample size

1. randomly sample *n* participants from within each group
2. average item responses for each sampled participant
3. compute mean challenge level score for groups
4. assign the winners
5. repeat steps 1-4 100 times and calculate the percentage of wins for each group

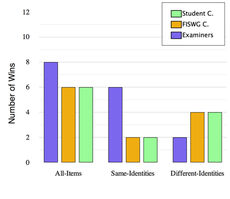
Chi-Square Analysis

- test the difference between the distribution of winners for image pairs in different conditions (PICT ⟷ FI vs. BI; EFCT ⟷ same- vs. different- identity)

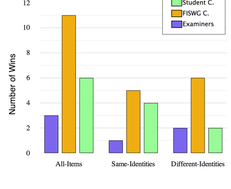
Results

PICT: Item Winner

Face: Tally of Item Wins



Body: Tally of Item Wins



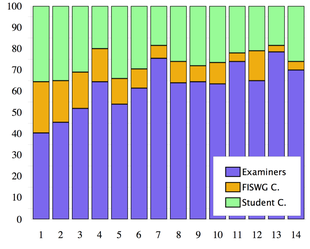
- Do examiners and untrained students perform differently on *face-informative* and *body-informative* stimuli?

⇨ YES (All-Items; Same-Identities)
Face-informative: examiners > controls
Body-informative: controls > examiners

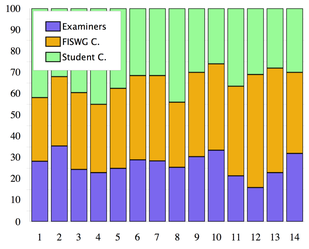
PICT: Fusion

- Is distribution of wins different for *face- vs. body-informative* stimuli across groups for *all* items? YES: $p\ (n = 14) < .01$

Face: All-Items

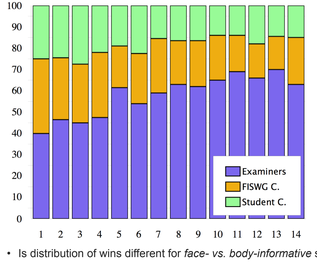


Body: All-Items

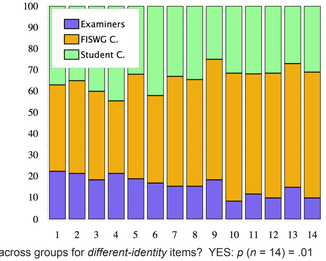


- Is distribution of wins different for *face- vs. body-informative* stimuli across groups for *same-identity* items? YES: $p\ (n = 14) < .01$

Face: Same-Identities

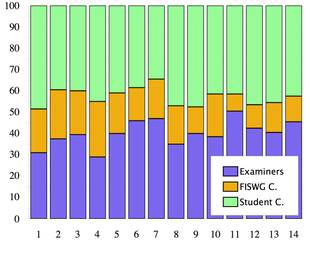


Body: Same-Identities

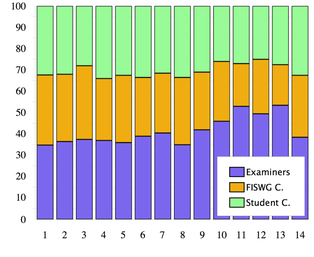


- Is distribution of wins different for *face- vs. body-informative* stimuli across groups for *different-identity* items? YES: $p\ (n = 14) = .01$

Face: Different-Identities

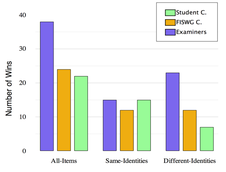


Body: Different-Identities



EFCT: Item Winner

EFCT: Tally of Item Wins



Do examiners perform differently from untrained students on same- vs. different-identity stimuli?

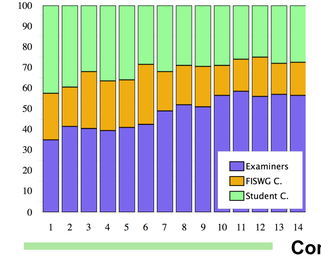
⇩

YES
Same-identity: examiners = controls
Different-Identity: examiners > controls

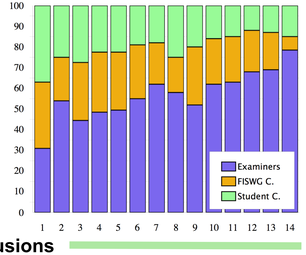
EFCT: Fusion

- Is distribution of wins different for *same-vs. different-identity items across groups*? YES: $p\ (n = 14) < .01$

EFCT: Same-Identities



EFCT: Different-Identities



Conclusions

- examiners surpassed untrained participants when internal face contains better information for identity than external face & body
- accuracy measures for examiners and controls must include **both** 1) same-identity verification and 2) different-identity rejection to understand the role of perceptual skill and response bias in performance

Discussion

- **Possibilities**
  - examiners concentrate on the face alone due to the extensive training
  - examiners believe that internal face is more stable over time than the external face and body
  - examiners process unfamiliar faces in the same way untrained people process familiar faces
  - examiners and untrained people approach the response scales in different ways
    - "retreat to the center of the scale" effect
- **Implications**
  - improve the training process by combining the external features with the internal features

References & Acknowledgement

White, D., Phillips, P. J., Hahn, C. A., Hill, M., & O'Toole, A. J. (2015). Perceptual expertise in forensic facial image comparison. Proceedings of the Royal Society B: Biological Sciences,.282, 1814–1822.
Rice, A., Phillips, P. J., & O'Toole, A. (2013). The role of the face and body in unfamiliar person identification. Applied Cognitive Psychology, 27(6), 761–768. doi:10.1002/acp.2969
Phillips, P. J., Beveridge, J. R., Draper, B. A., Givens, G., O'Toole, A. J., Bolme, D.,…Weimer, S. (2012). The good, the bad, and the ugly face challenge problem. Image and Vision Computing, 30 (3), 177–185. doi:10.1016/j.imavis.2012.01.004
Hu, Y., Jackson, K., Yates, A., White, D., Phillips, P. J., & O'Toole, A. J. (2017). Person recognition: Qualitative differences in how forensic face examiners and untrained people rely on the face versus the body for identification. *Visual Cognition*, 1-15.

Acknowledgement
Funded by the National Institute of Justice (NIJ) and National Institute of Standards and Technology (NIST)