Interpreting Emotional State from Facial Expressions in Deaf Signers

Matt Dye, RIT/NTID

July 5, 2024

Expressing Emotions in Sign Language
ExEmSiLa Workshop 2024

ViCom

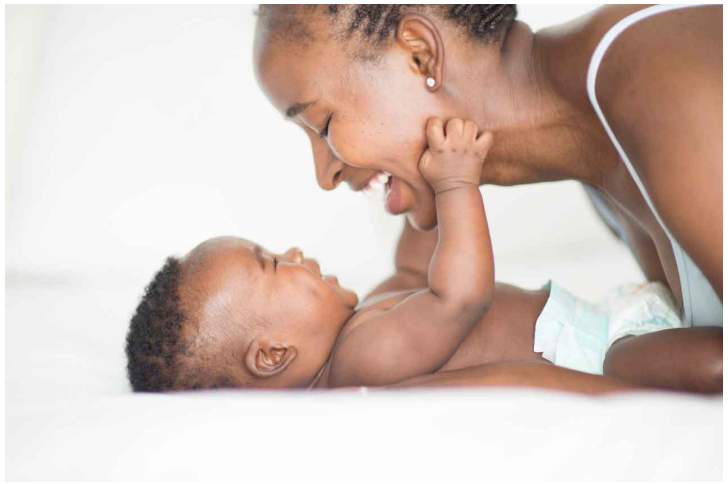
RIT

Talk Outline

- **What's The Problem?**
  - Affective and grammatical expressions from a cognitive science perspective
- **What Do We Know?**
  - What we know about face processing (incl. emotion) in signers
- **What Next?**
  - Suggestions for moving forward

What's The Problem?

Faces and Human Cognition



VS

Recognizing Emotion

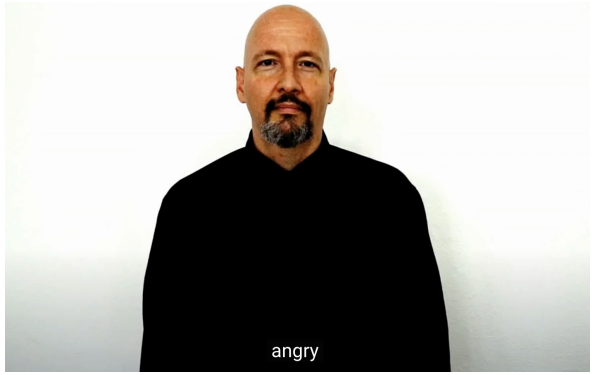
Now For Our Daily Laugh …

SADNESS ANGER FEAR DISGUST HAPPINESS SURPRISE


Pause
I'm going to ask questions, but they're all random topics.
0:00 / 2:59

Expressing or Communicating Emotion?

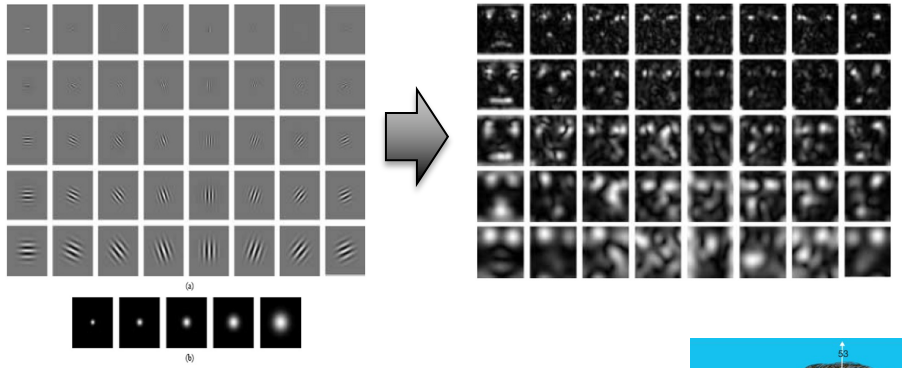
- **How is emotion expressed and understood within a linguistic system, e.g. ASL or DGS?**
  - "Signing Anger"
- **How can emotion be distinguished from information encoded within a linguistic system?**
  - "Angry While Signing"



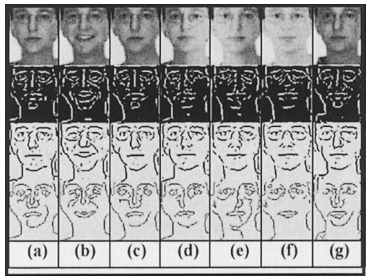
Emotion	Language
Peripheral nervous system	Central nervous system
Physiological	Motor
Difficult to control	Under control

Brain sees just one face

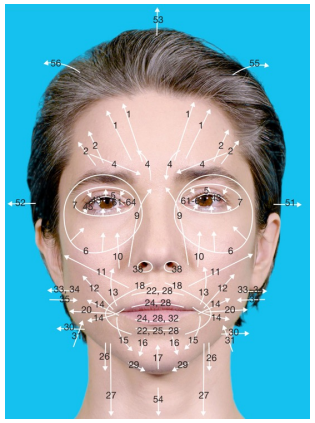
Representation



Liu & Wechsler, 2003



Pujol & Wechsler, 2002

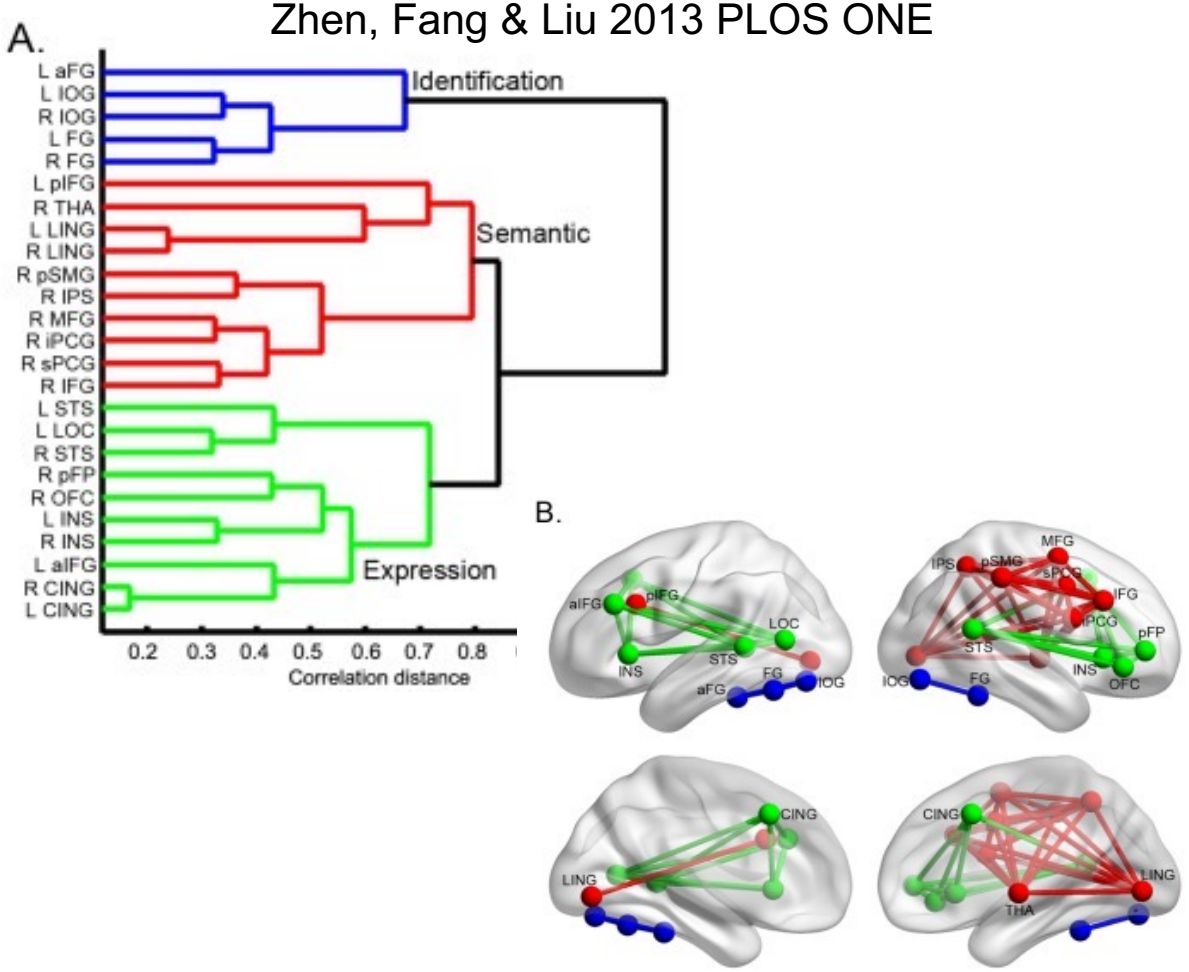


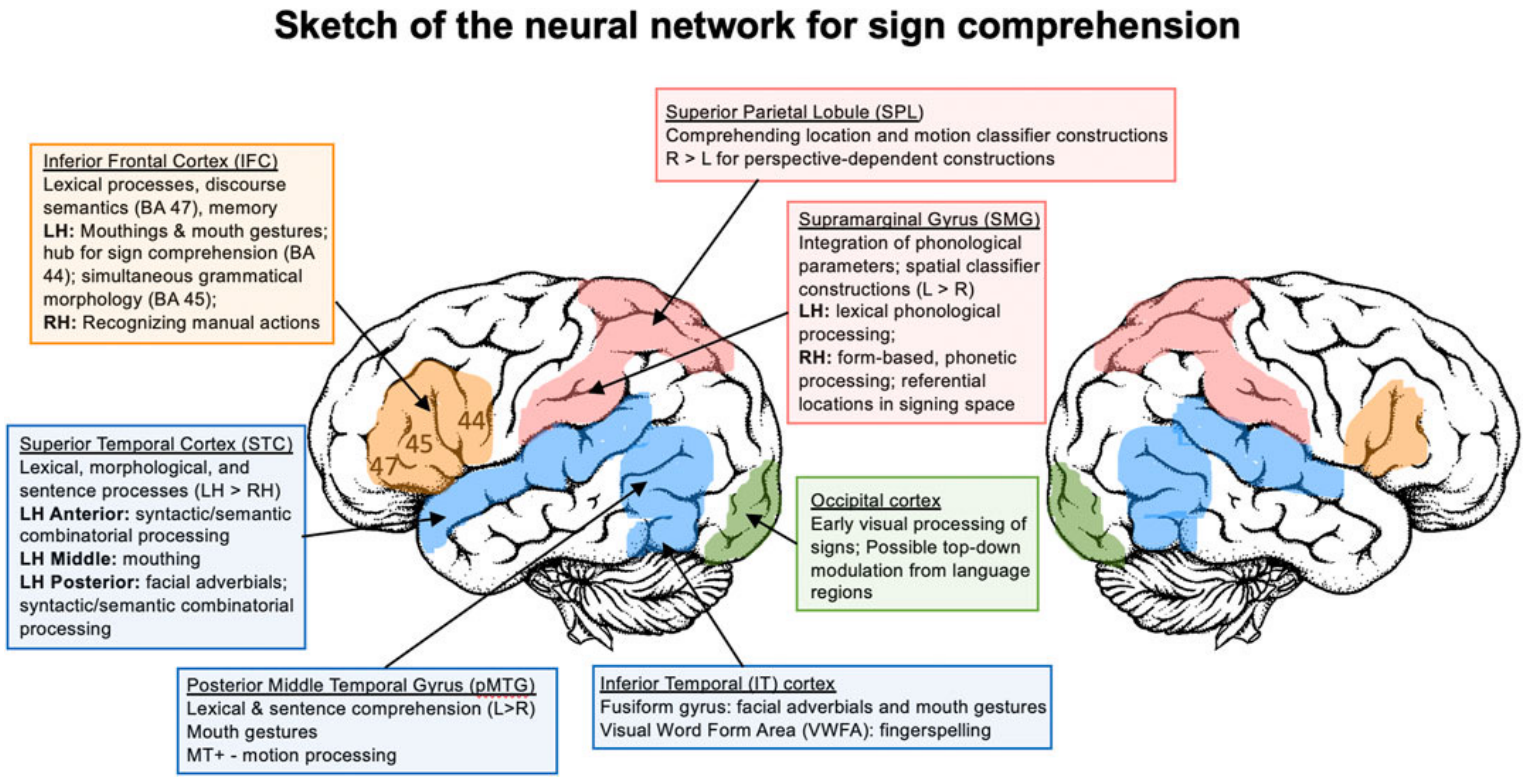
Ekman & Friesen, 1978

- How does the brain go from a visual representation of the face to a decision about the expressed emotional state?
  - Based solely on a visual representation?
  - Does the brain compute a motor representation?
  - Assuming the face is just one of many cues to emotional state, how is it weighted?

- **Computational/brain networks**
  - Primary visual cortex – low level visual features
  - Fusiform face area – face ‘expertise’
  - Inferio-temporal cortex – naming and conceptual knowledge
  - Prefrontal cortex – decision making
  - Medial temporal lobe - movement
  - Primary motor cortex – motor planning

Zhen, Fang & Liu 2013 PLOS ONE





Emmorey 2021 Front Commun

- Mouthing
  - Inferior frontal
  - Middle superior temporal
- Mouth gestures
  - Inferior frontal
  - Middle temporal gyrus
- Facial adverbials
  - Posterior superior temporal
  - Fusiform gyrus
- Motion processing
  - Middle temporal +

Multiple Processing Routes

- Processing of facial expressions processed in LEFT AND RIGHT hemispheres
- Most face processing related to sign languages is in the LEFT hemisphere
- BUT … minimal overlap with facial expression areas
- Perhaps the brain processes language features and emotion features in separate pathways????

What Do We Know?

Methodologies for Studying Face Processing

- **Face recognition**
  - Delayed match-to-sample
  - Face inversion
  - Eye fixations during face viewing

Delayed Match-to-Sample

- Target and foil faces presented from different perspective
- Makes simple visual matching harder
- Observer must derive a viewpoint-independent face representation

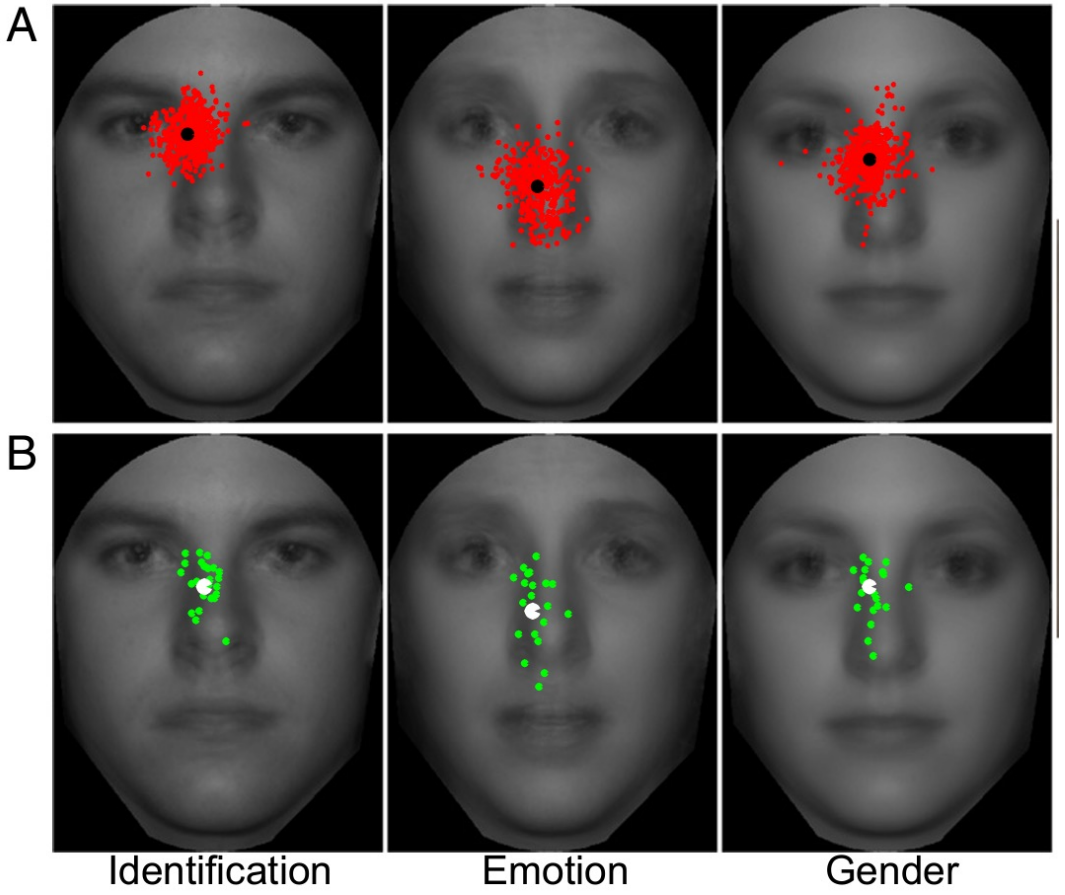
Face Inversion Effect

- Inverting faces makes them harder to recognize as faces
- Same not true for other complex objects
- Taken as evidence for face processing being holistic/configural
- Upside down faces do not as easily activate face representations in FFA (fusiform face area)

FACE or CAR?

Eye Fixations

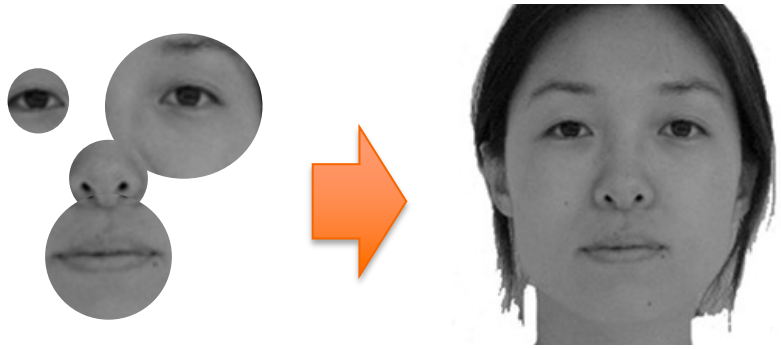
- Eye fixations vary depending upon task
- Predicted by
  - Location of relevant information in the face
  - Acuity of the human visual system – highest at fixation, then drops off quickly



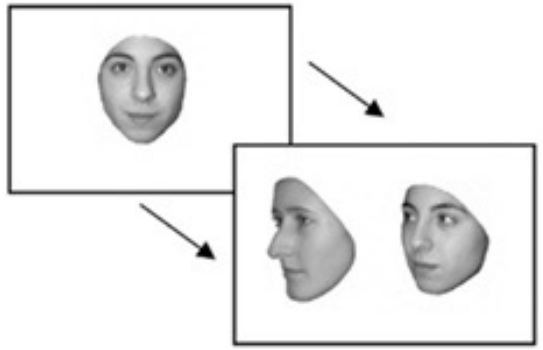
Peterson & Eckstein 2012 PNAS

Eye Movements – Expanding Spotlight

Fixation #1 (short) Fixation #2 (long) Fixation #3 (short) Fixation #4 (medium)

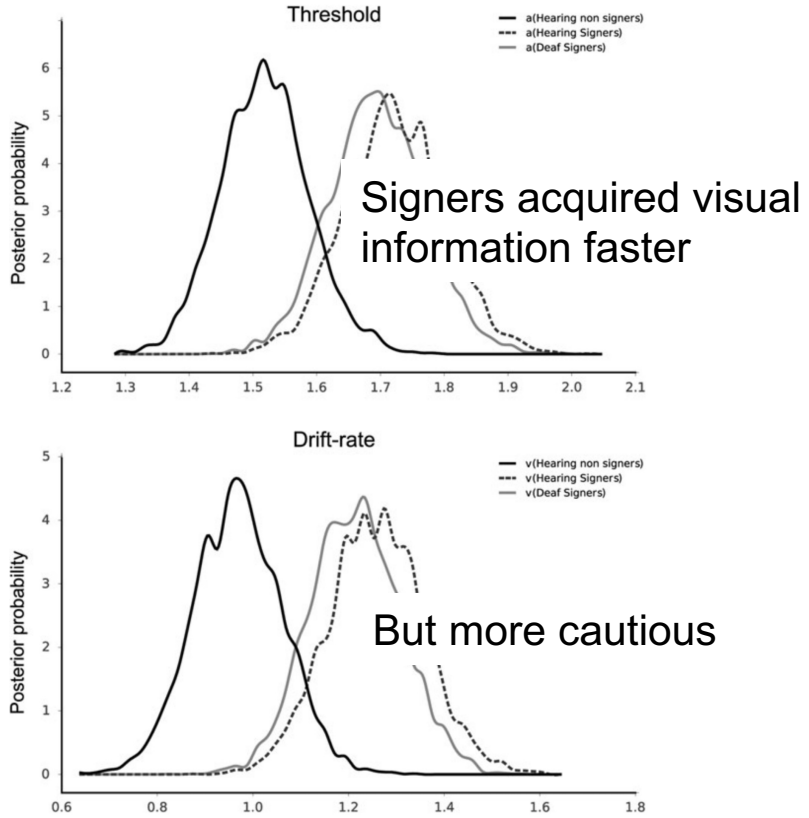


- Information revealed during fixation
- Eye movement removes information and reveals at new location
- Information integrated to perform task

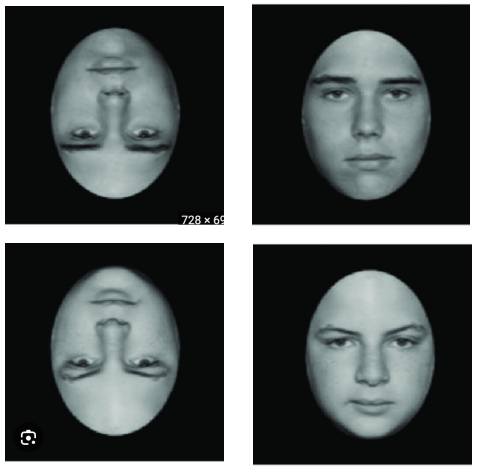


From de Heering et al 2012 Front Psychol

- Deaf and hearing signers vs. hearing nonsigners
- Signers' face recognition was slower but more accurate than hearing nonsigners
- Signers acquired visual information from face more rapidly
- But were more cautious about making a decision



Stoll et al 2017 JDSDE



Lao et al 2017 JoV

- Compared deaf signers and hearing nonsigners
- Deaf signers did not show the anticipated face inversion effect
- Evidence for a lack of configural/holistic face representation?

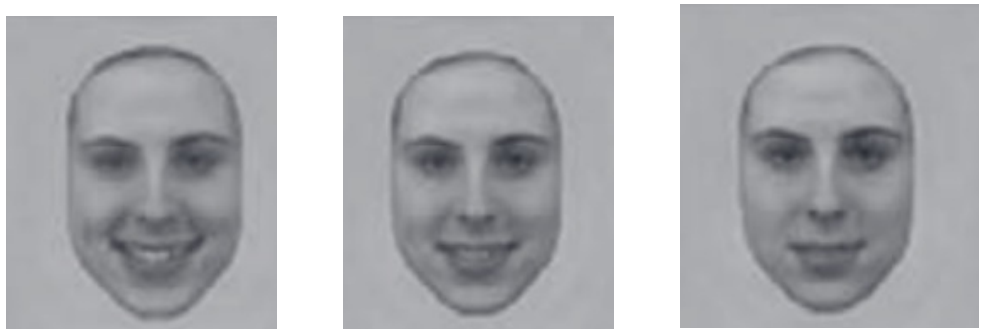
- Expanding spotlight task with deaf signers and hearing nonsigners
- Deaf signers acquired more information with each fixation
- Strategy that minimized eye movements and maximized information uptake

Lao et al 2017 JoV

Methodology for Studying Emotion Recognition

- **Emotion recognition**
  - Stimulus intensity (degree of emotion)
  - Stimulus signal (signal:noise ratio)

Intensity



Signal

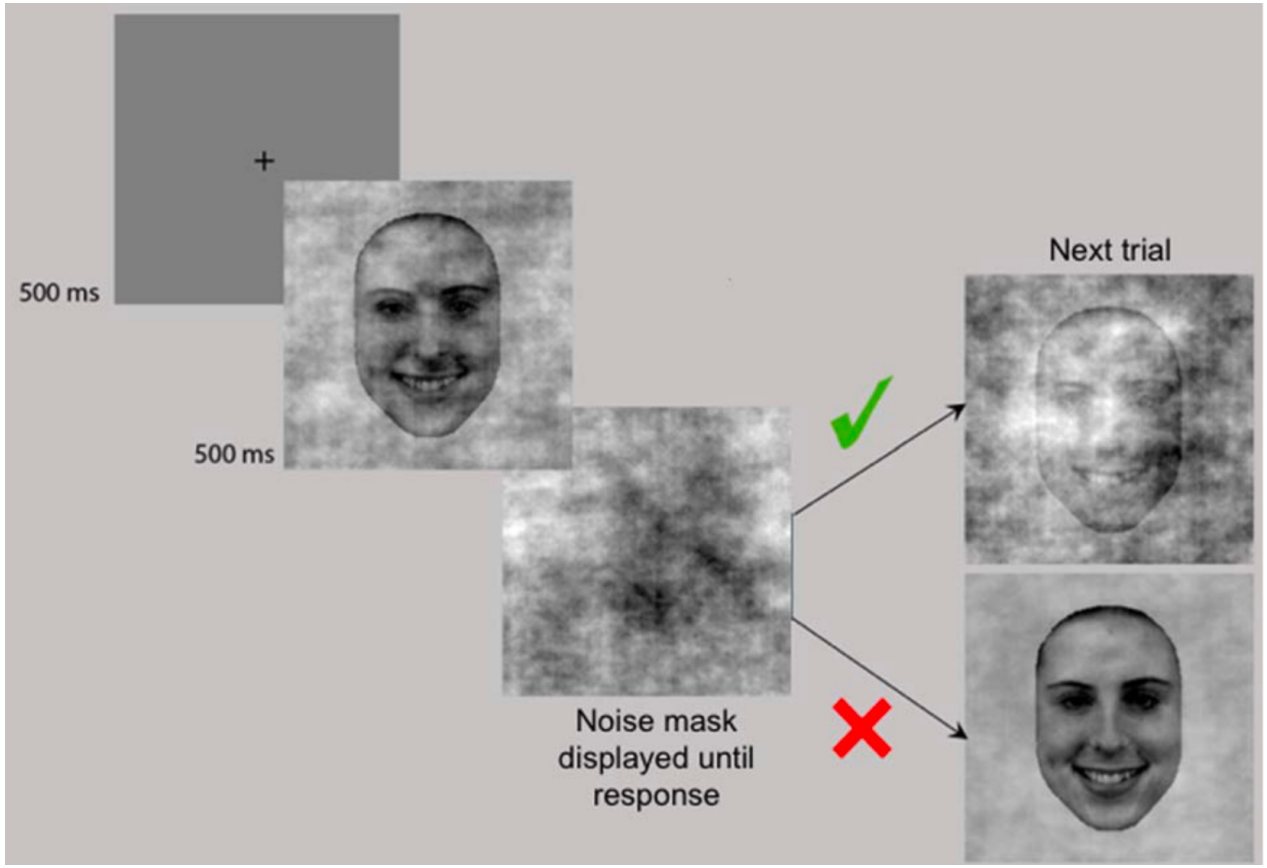
Stoll et al 2019 JDSDE

Adaptive Staircase Procedure

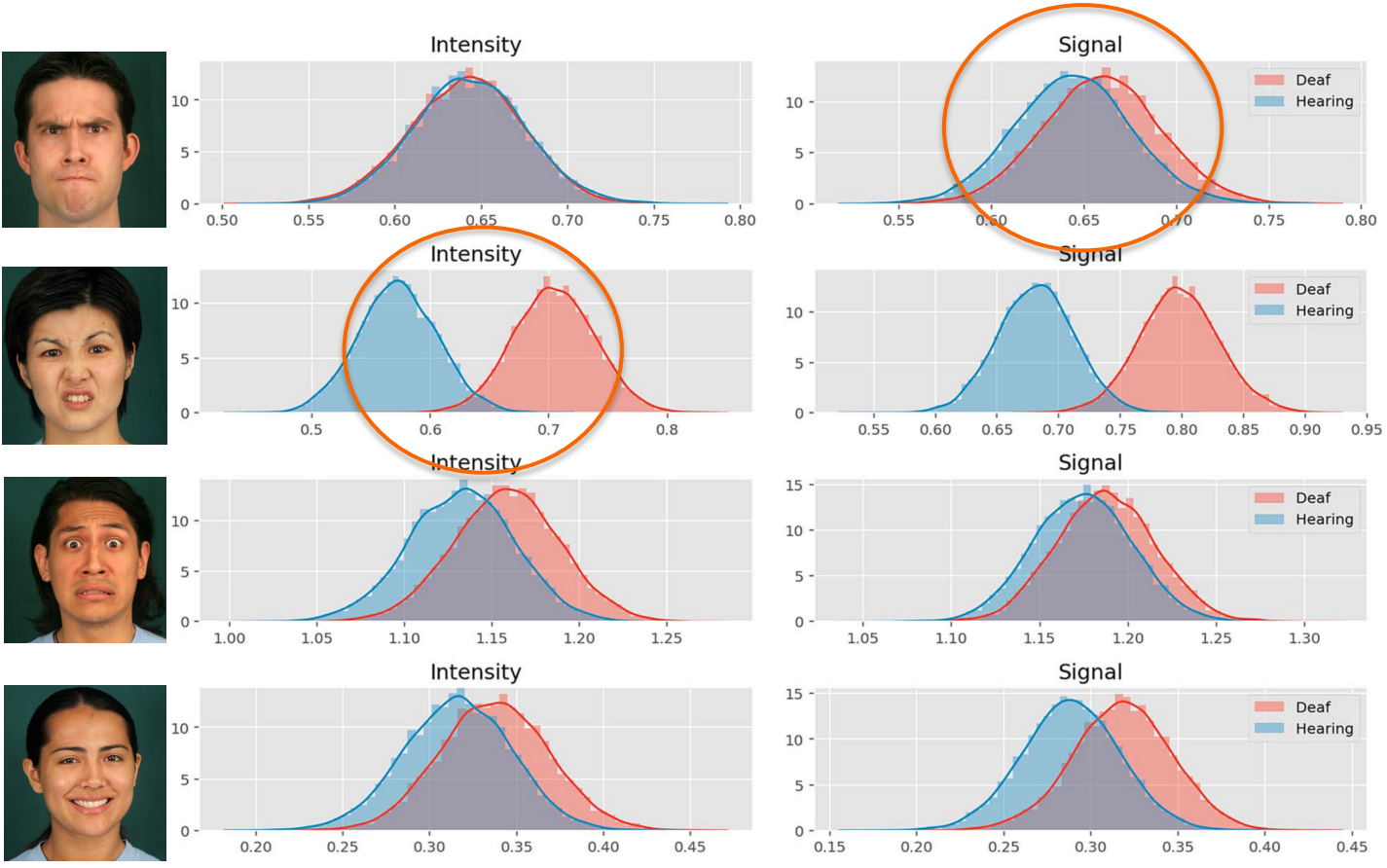
Correct? Make it harder.

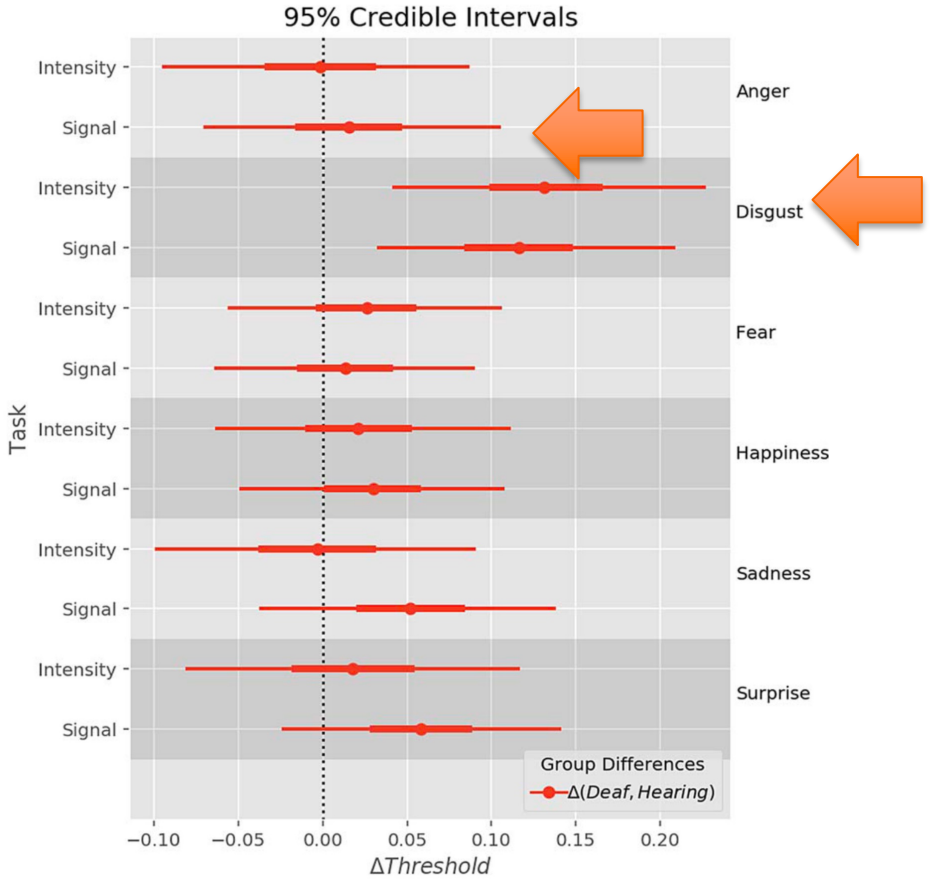
Wrong? Make it easier.

Find the “sweet spot”
AKA threshold



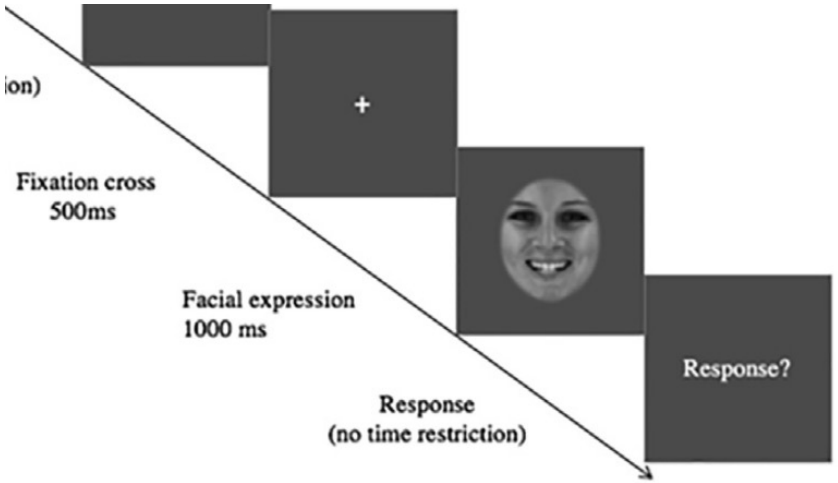
Stoll et al 2019 JDSDE





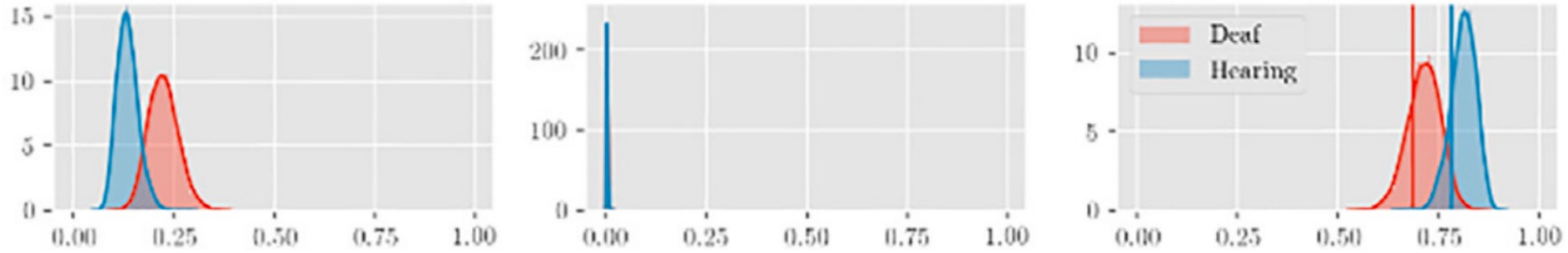
Stoll et al 2019 JDSDE

Rodger et al 2021 Heliyon



- Stimulus duration = 1 second
- Three conditions:
  - static
  - dynamic
  - shuffled

- Surprise often miscategorized as happiness
- Seen for both deaf and hearing observers
- Deaf and hearing performed similarly across static, dynamic (and shuffled) conditions



Signers' Emotional Face Processing

- **The face is special for signers**
  - Face information absorbed more rapidly (more attention?)
  - More cautious in deciding (competing emotional and linguistic signals?)
- **Deaf signers**
  - Different viewing strategy – minimizes eye movements
  - Evidence for individuation of facial features
  - Surprise and Disgust often confused (AU complexity or overlap with grammatical features?)

Conclusions

- **Conclusion 1: Differences seem to be due to visual processing differences - could be driven by both deafness (viewing strategy) and by sign language experience (increased attention to face)**

Conclusions

- **Conclusion 2: Face representations may be less holistic due to emphasis on features needed for grammatical purposes (need more data from hearing signers)**

Conclusions

- **Conclusion 3: Current paradigms don’t allow us to determine whether confusability is due to representational issues or due to labelling – studies use English labels**

What Next?

Context

Body Posture

Prosody/Tone

Physiological Changes

AU	Importance
head pos: back	0.101988
head pos: tilt front	0.089116
head pos: turn right	0.067589
head pos: tilt left	0.064259
head pos: tilt right	0.063508
nose wrinkle/tensed	0.063151
head mvmt: nod	0.062214
head pos: turn left	0.060699
eye gaze to addressee	0.053793
eye brows lowered	0.039102

Kezar & Zhou 2021 ACM

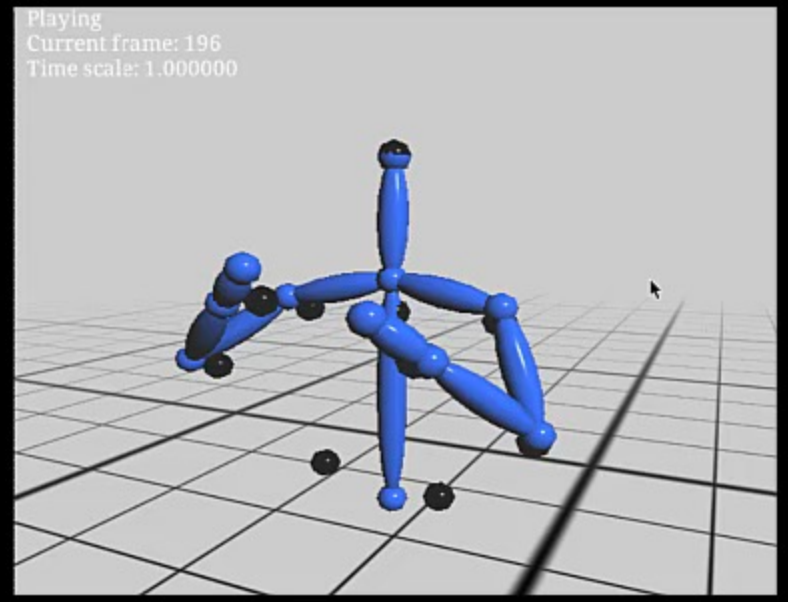
- **Used machine learning to train a classifier to detect and label emotional states (based on FACS) from an ASL corpus**
  - Classification accuracy was high (79%)
  - Body movement more important than facial features
  - Cognitive offloading – signers use body movements to convey emotion because face is co-opted by grammatical system?

Future Research

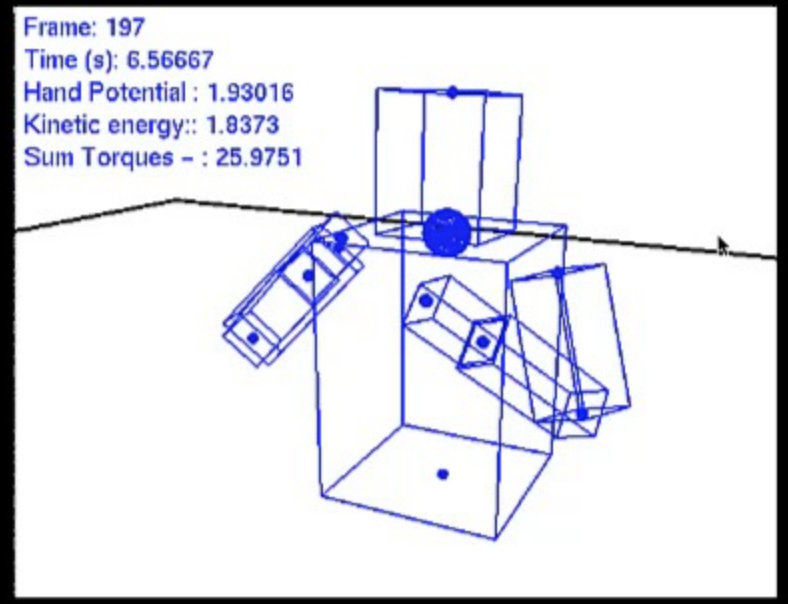
- **Should take a multimodal approach**
  - Facial expressions, body shifts, semantic context
- **Based on corpora of naturalistic signing**
  - This is the “learning signal”
- **Take advantage of AI approaches to automatically label and categorize features within the sign space …**

Results: RGB Video with skeletal model

Original Video with 3D Body Model Super-imposed



3D Points and 3D Body Fit



Physically-based Model

Chloé Stoll

Roberto Caldara

Olivier Pascalis

Junpeng “Charles” Lao

Helen Rodgers

Thank You