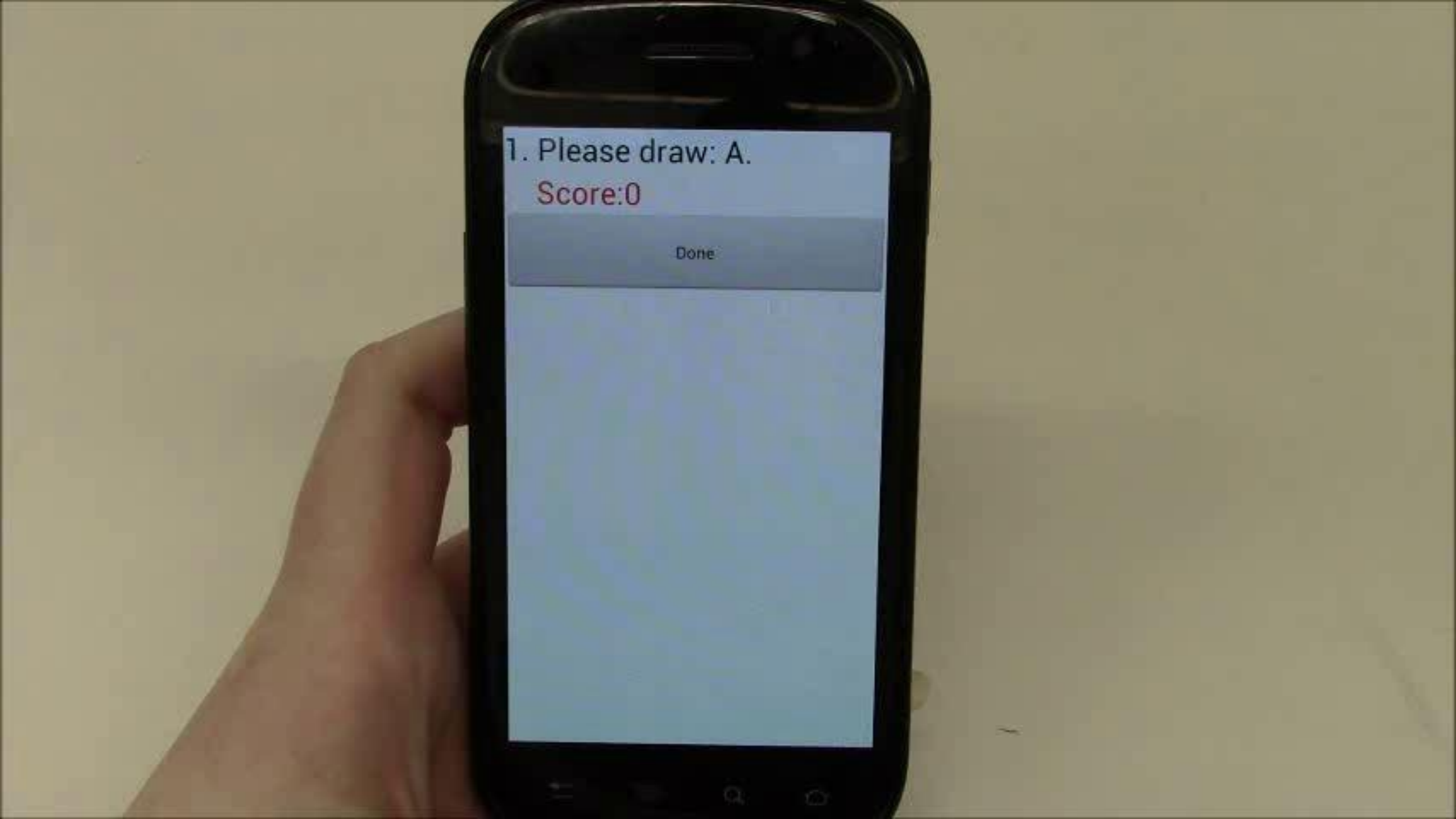
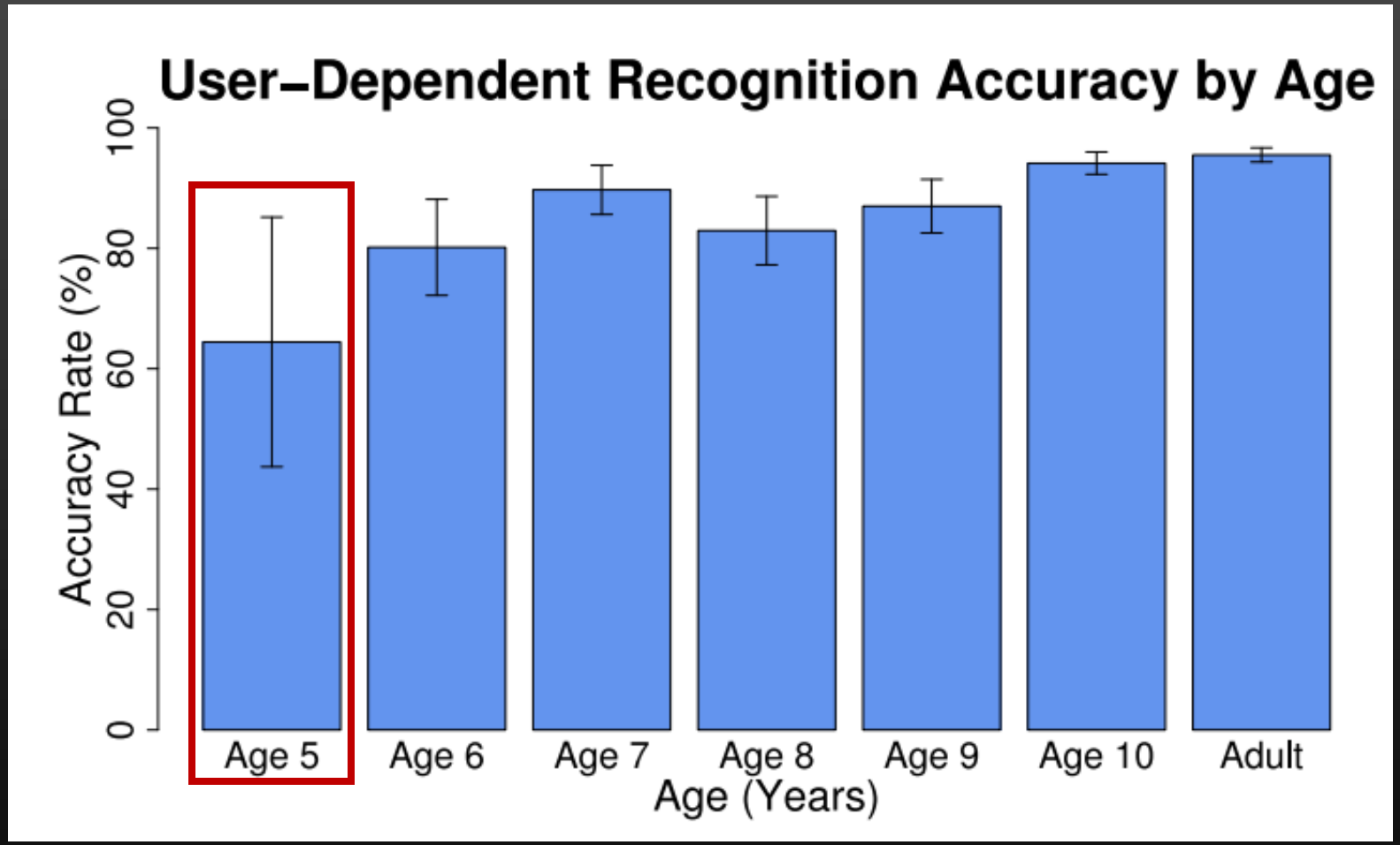
Comparing Human and Machine Recognition of Children's Touchscreen Stroke Gestures

Alex Shaw

Jaime Ruiz

Lisa Anthony

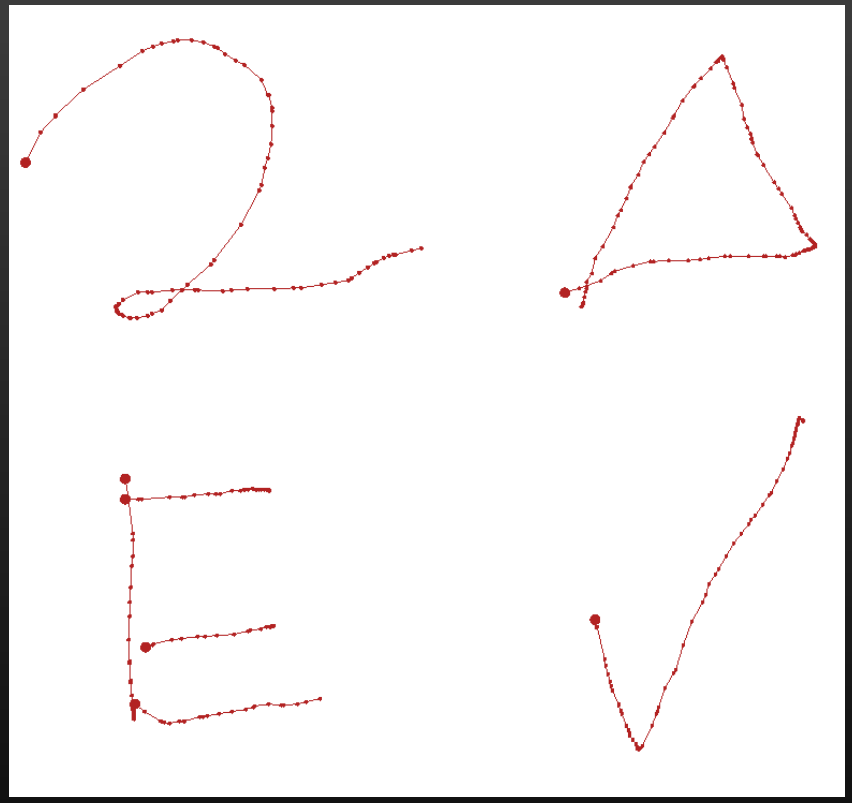

1. Please draw: A.
Score:0
Done



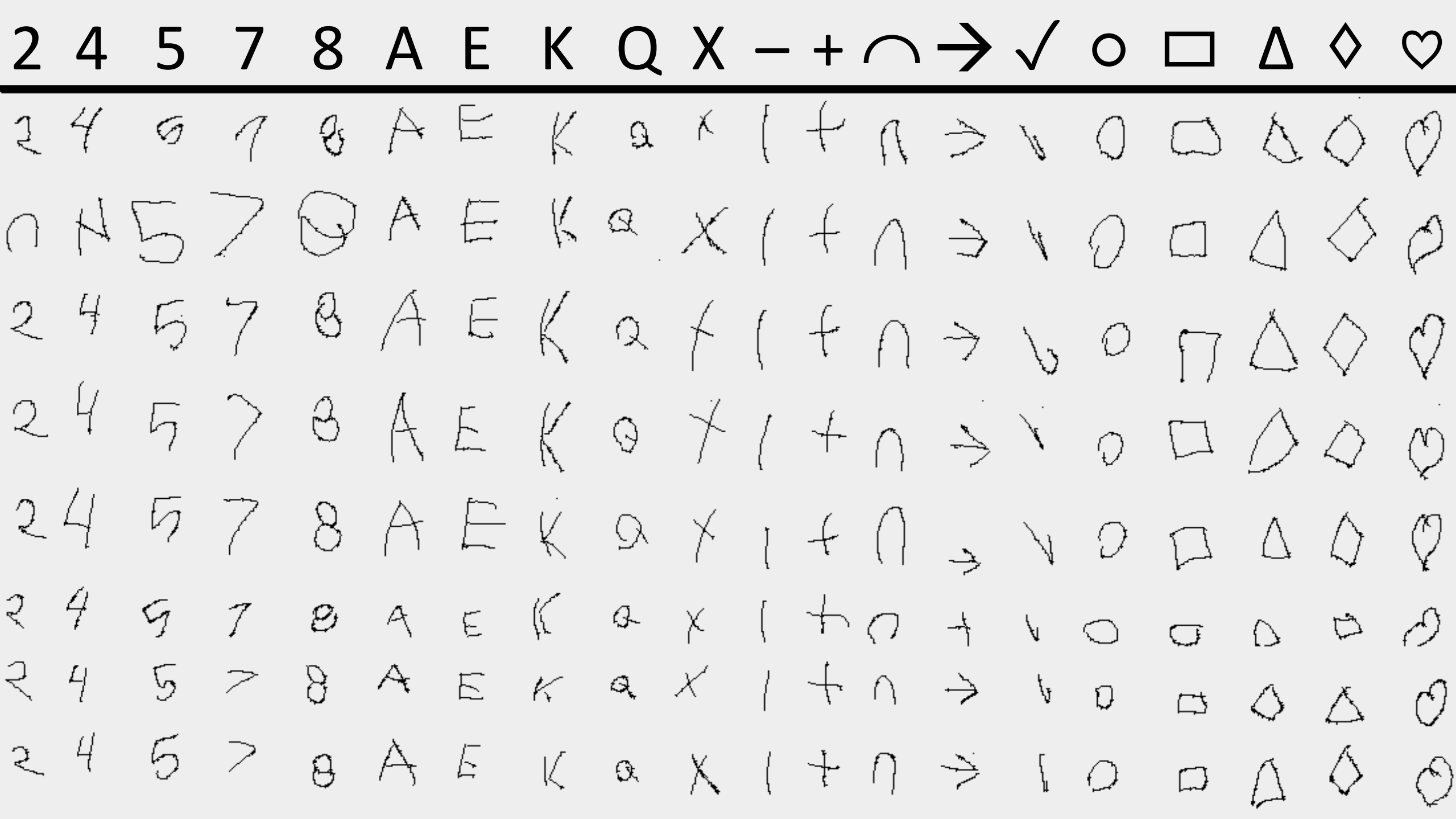
Recognition rates for 5- to 10-year-old children's gestures (Woodward et al., CHI '16)

Definition: Gesture

- Gesture – a series of one or more strokes to create a letter, number, symbol, or shape on a touchscreen

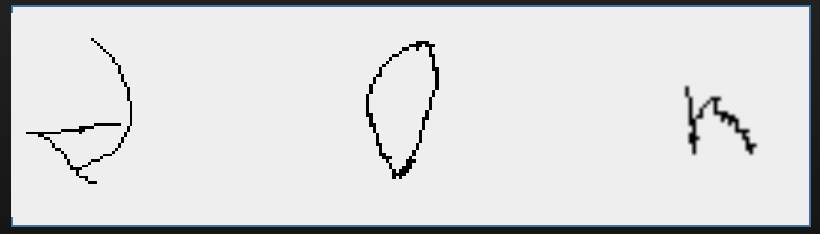


Four examples of different touchscreen gestures

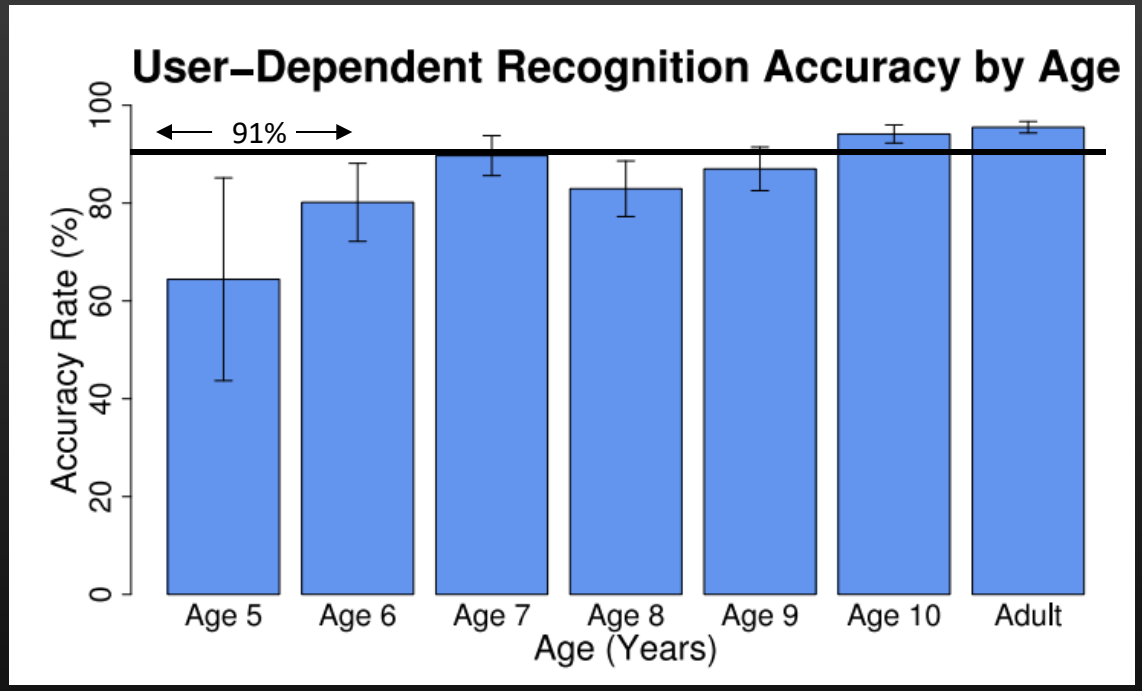

2 4 5 7 8 A E K Q X − + ∩ → ✓ ○ □ Δ ◊ ♡

How good is good enough?

- In handwriting, ~91% (Read et. al., IDC '03)
  - Is this practical?
  - If not, what is?



4 Rectangle K



Recognition accuracy with 91% accuracy line representing target from Read et al. (IDC'03)

Recognition Through Crowdsourcing

- Idea: Have humans classify gestures to get target accuracy
- “A paradigm for utilizing human processing power to solve problems that computers cannot yet solve” (Von Ahn dissertation, 2005)

Image by Pixabay under Creative Commons 0
https://pixabay.com/en/people-group-crowd-team-isolated-309093/

Gesture Set and Corpus

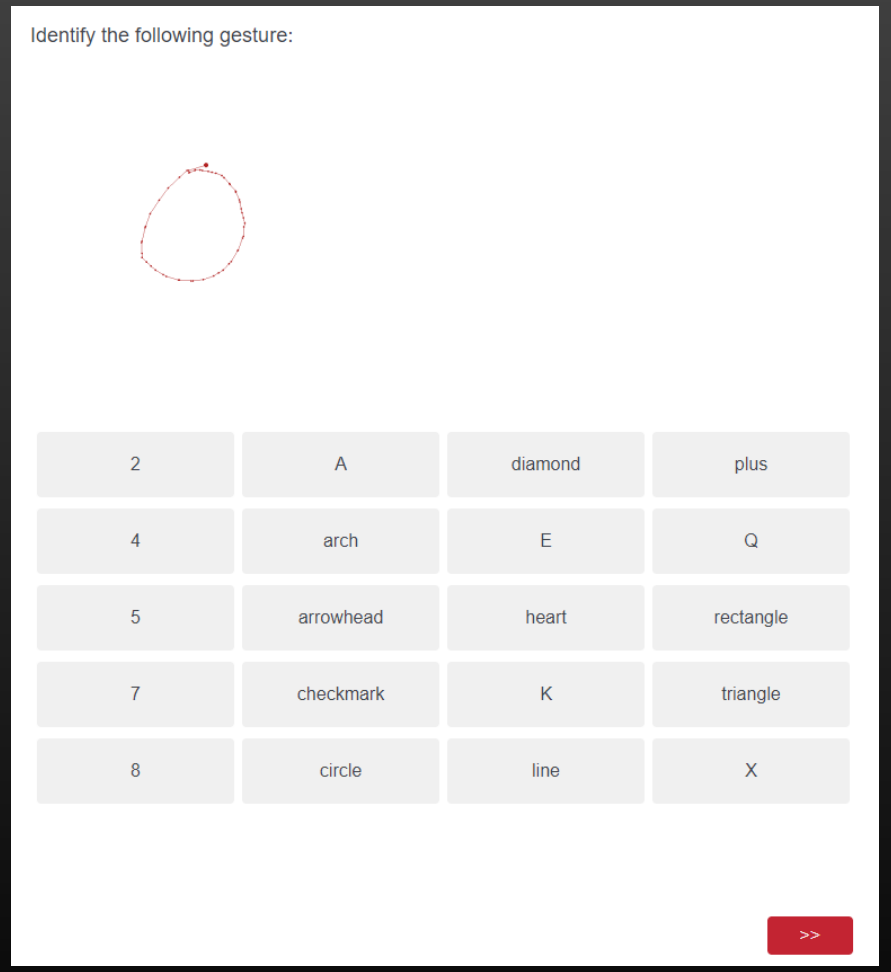
- 20 gestures
  - 4 categories
- Gestures collected by Woodward et al. (CHI ’16)
- 26 writers ages 5 to 10
- 26 writers x 20 gestures x 5 examples = 2,600 gestures

Letters:	A	E	K	Q	X
Numbers:	2	4	5	7	8
Symbols:	—	+	⌒	→	✓
Shapes:	○	▭	△	◇	♡

Gesture types in our experiment
(from Anthony et al., ITS ‘12)

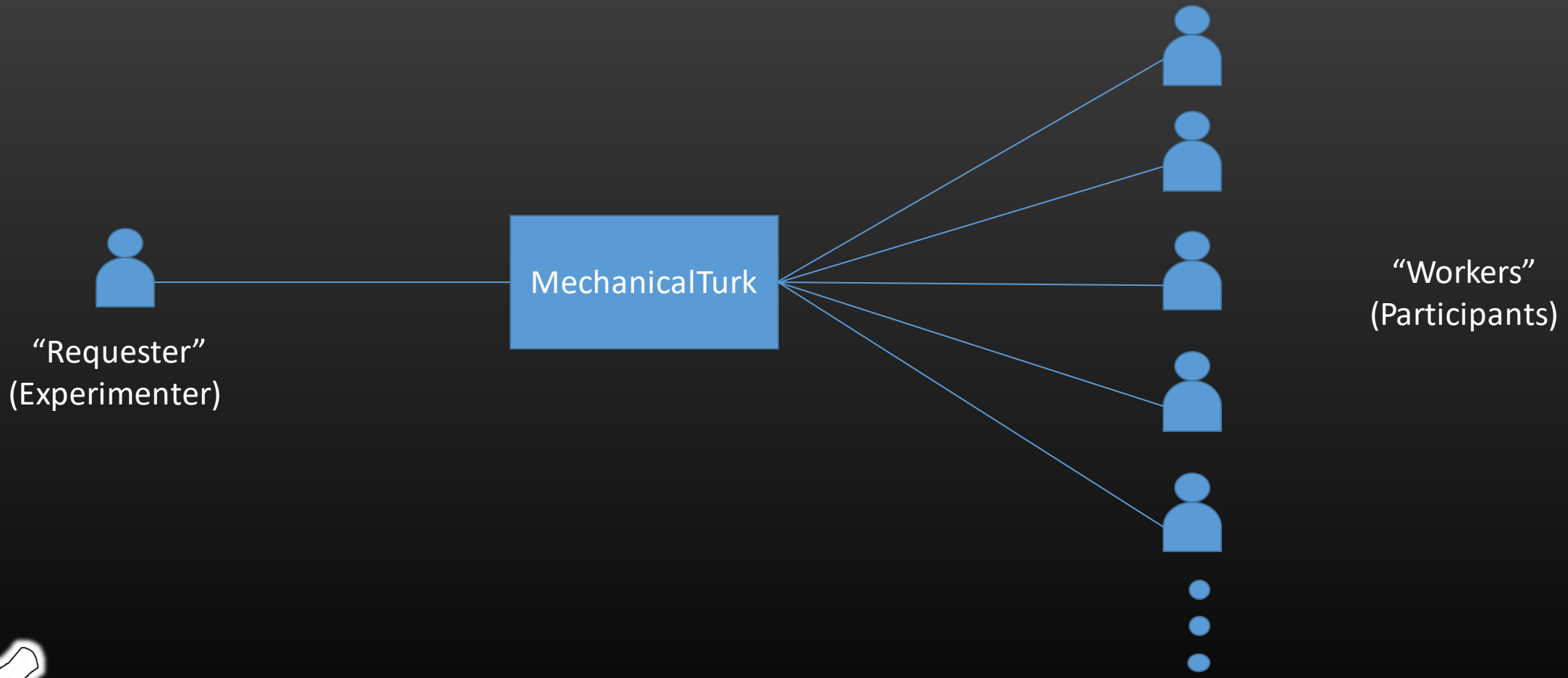
Implementation

- Survey using Qualtrics (online data collection tool)
  - 2,600 questions (1 per gesture)
- 131 participants total
  - Recruited via MechanicalTurk
  - All 18+ and from United States



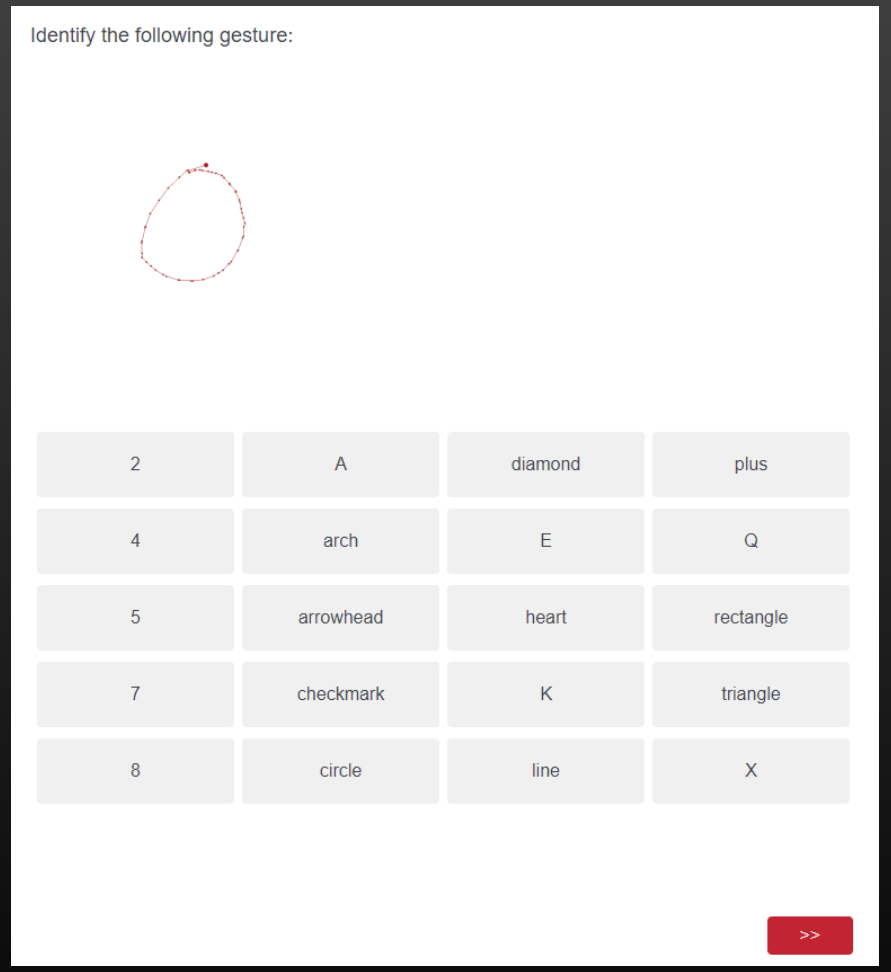
A question from our survey.

Human Recognition



Human Recognition

- Each gesture seen by 3+ participants
  - Marked as correct if over half respond correctly
  - Ties broken with additional participants



A question from our survey.

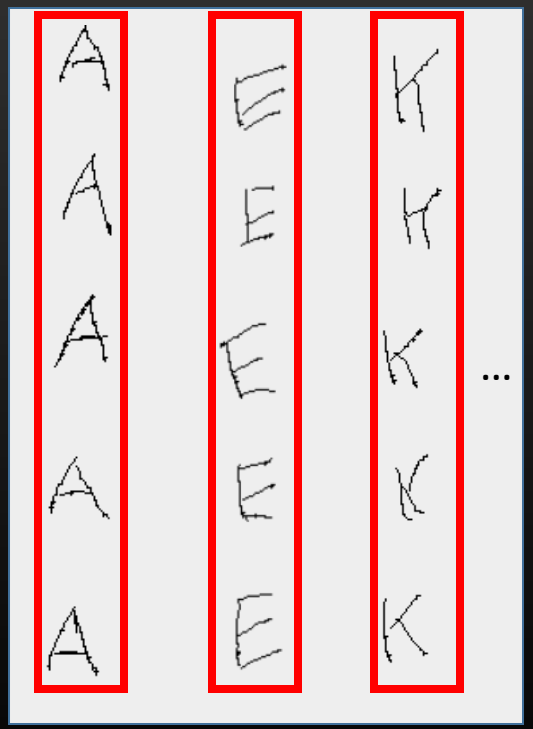
Machine Recognition

$P Recognizer
(Vatavu et al., ICMI '12)

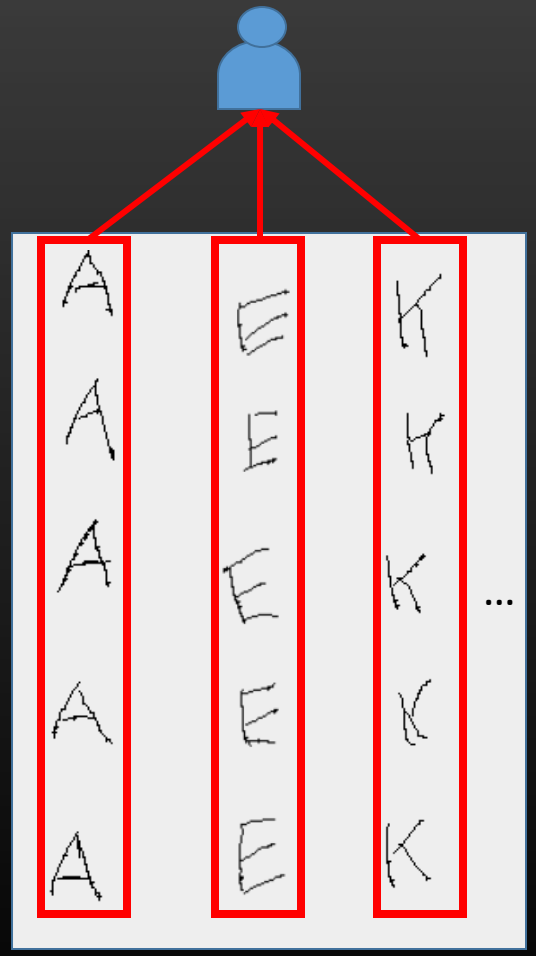
Results

Computing Recognition

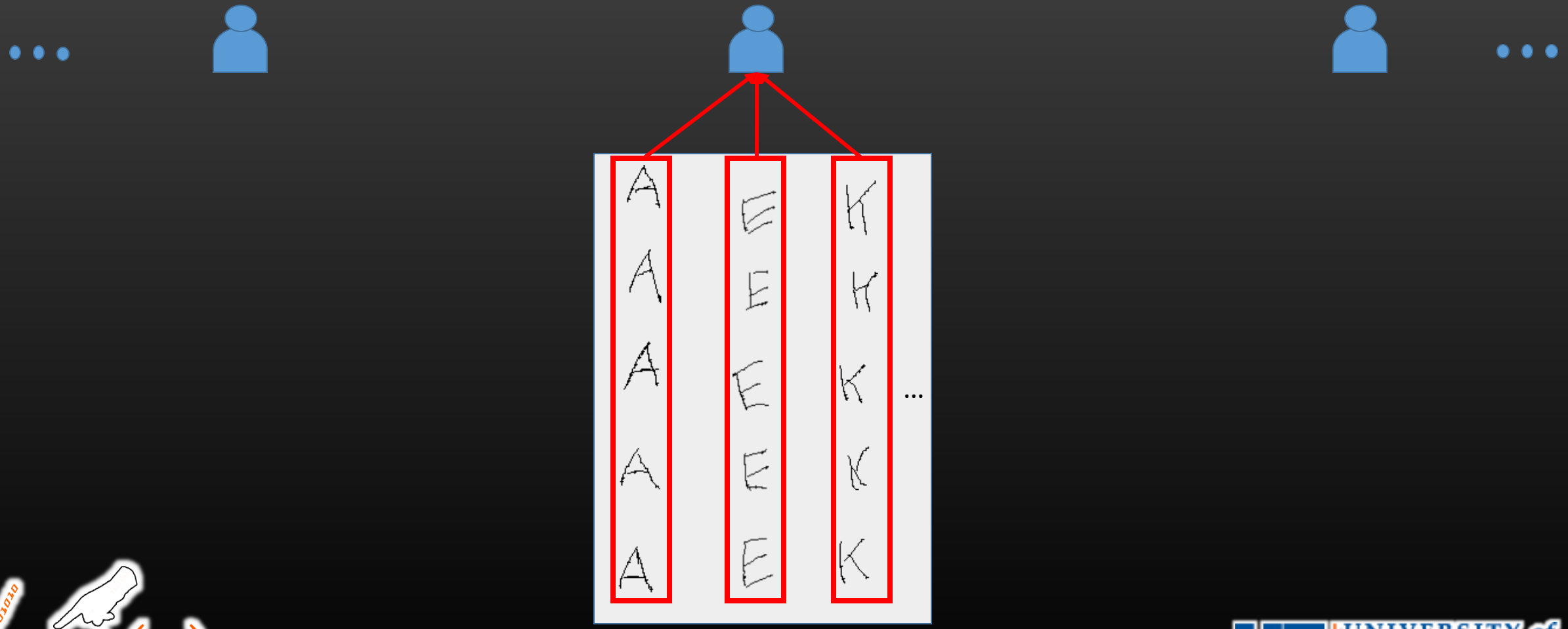
Computing Recognition



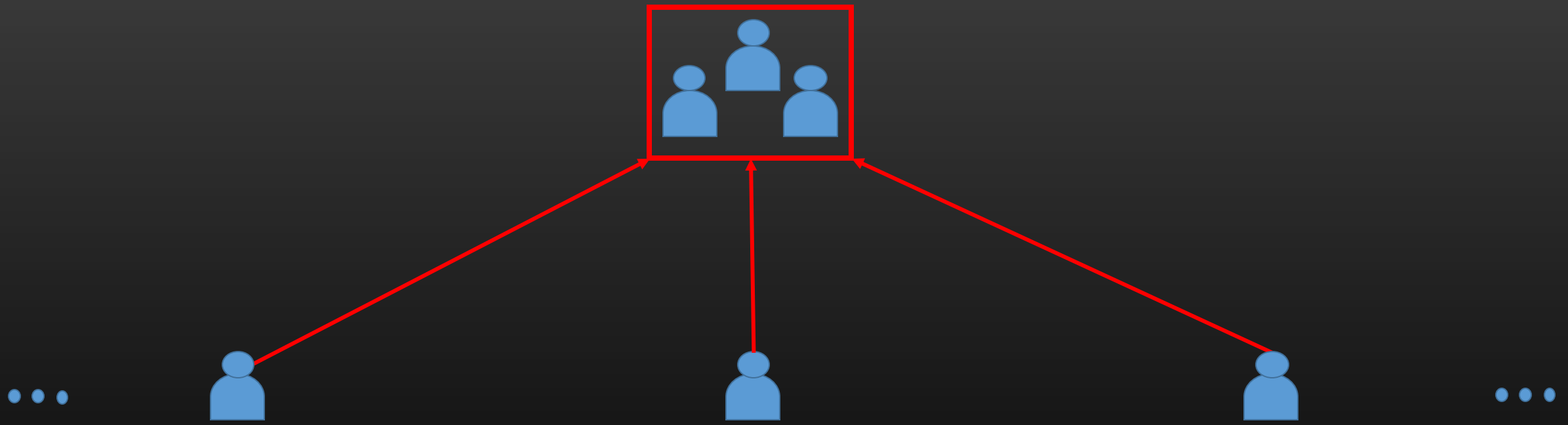
Computing Recognition



Computing Recognition



Computing Recognition



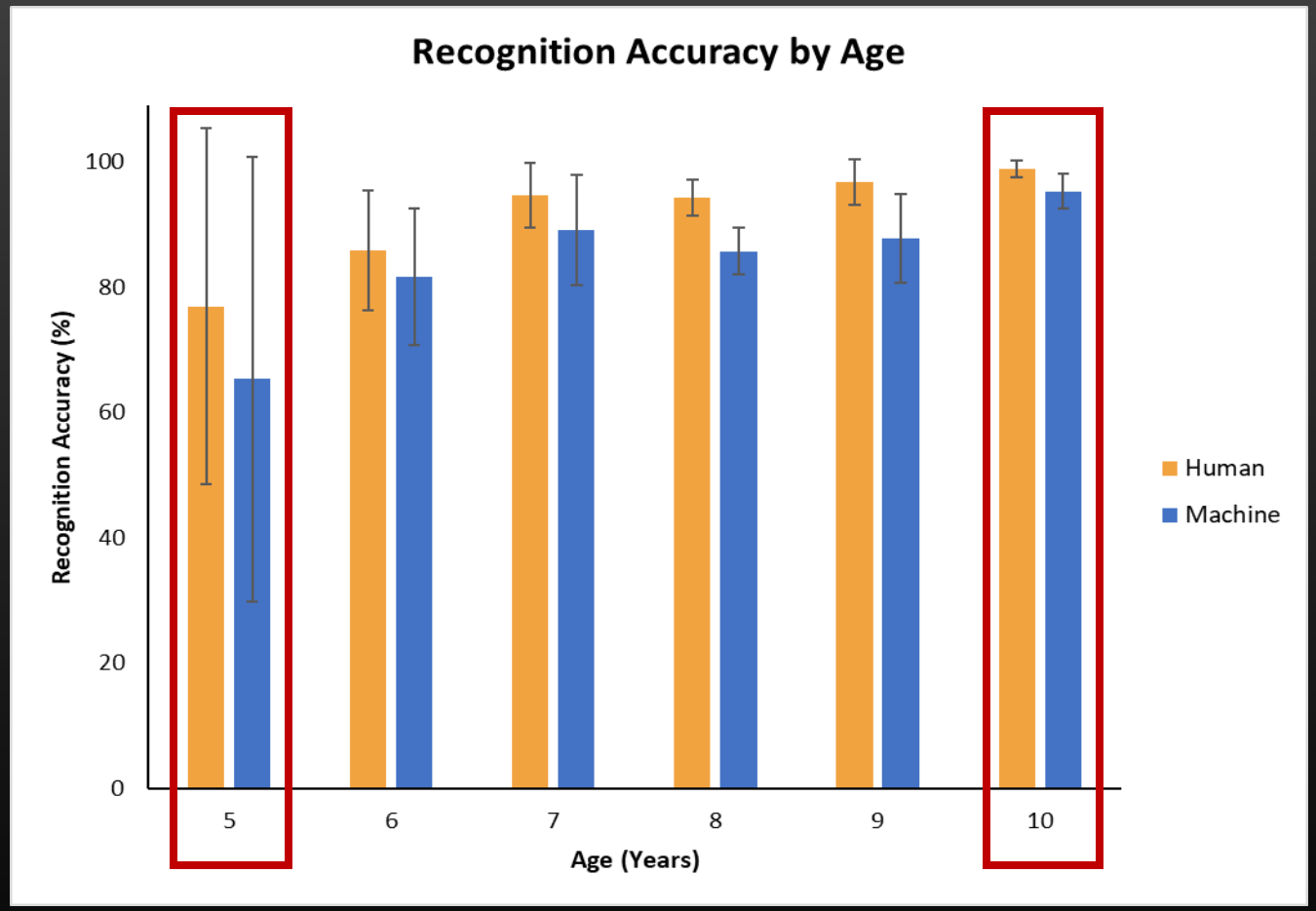
Results

90.60%

84.14%

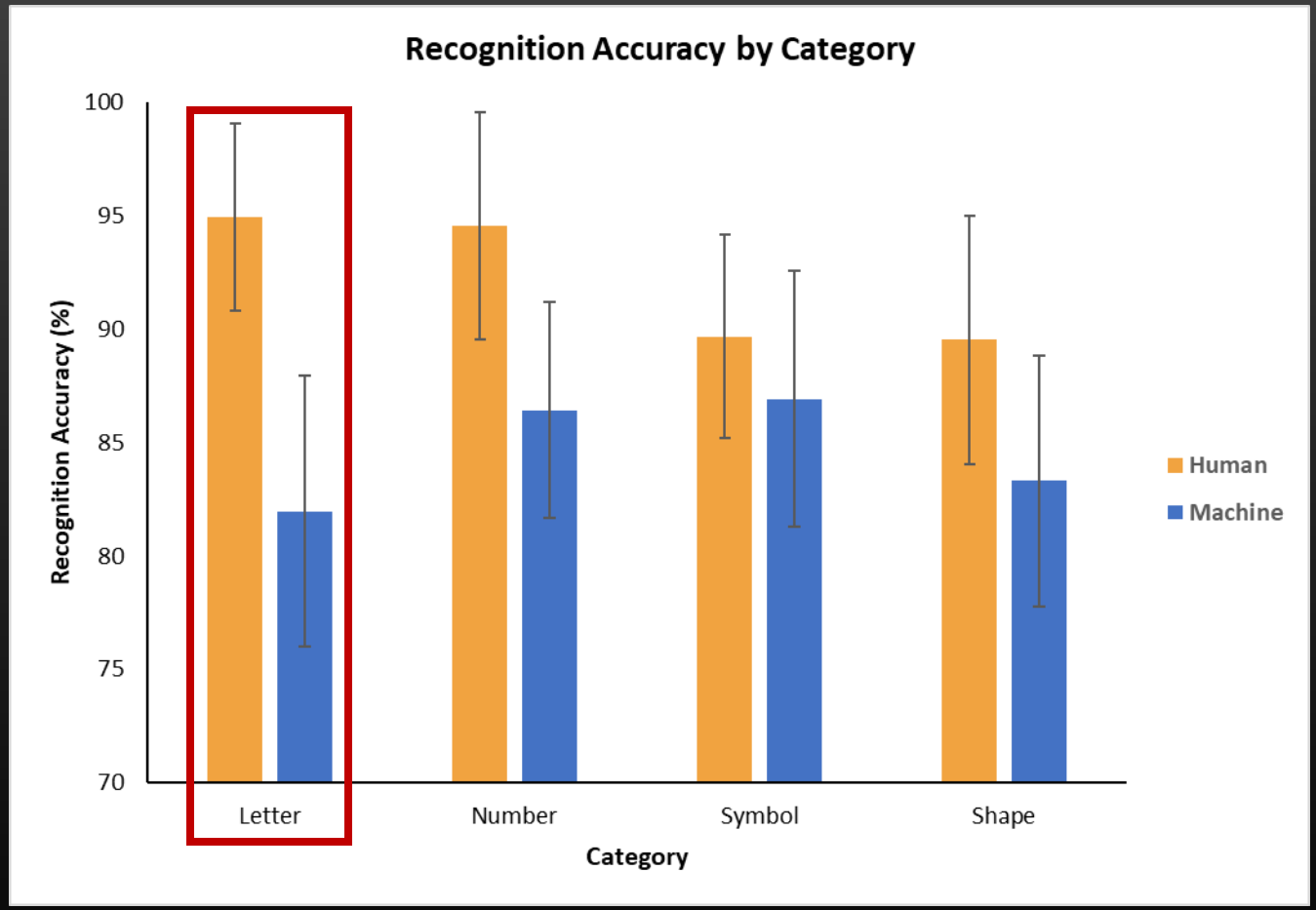
Results

- Human recognition significantly better than machine recognition ($p < 0.05$)
  - Tukey post-hoc test ($p < 0.05$):
    - 5-year-olds
    - 8-year-olds
    - 9-year-olds



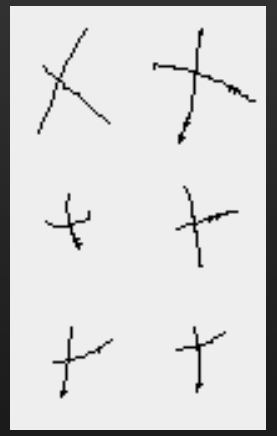
Results

- Significant main effect of category on recognizer ($p < 0.05$)
- Significant interaction between category and recognizer type ($p < 0.05$)
- Tukey post-hoc test ($p < 0.05$)
  - Letter

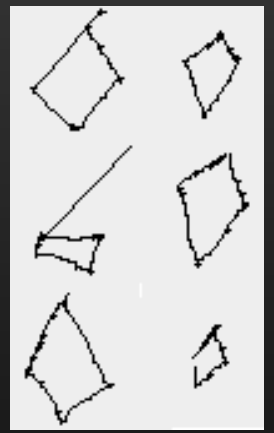


Commonly Confused Pairs (Human)

Similar Gesture Types



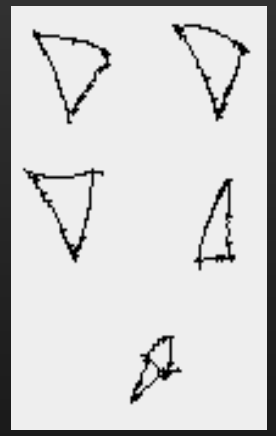
"Plus" recognized as "X" (23%)



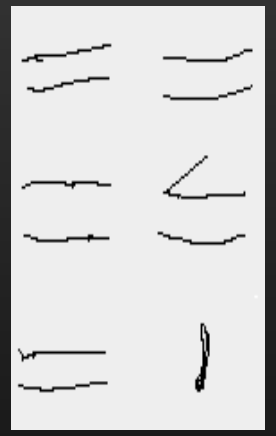
"Diamond" recognized as "rectangle" (12%)

Articulation/Development

"Diamond" recognized as "circle" (5%)



"Diamond" recognized as "triangle" (5%)



"Rectangle" recognized as "line" (6%)

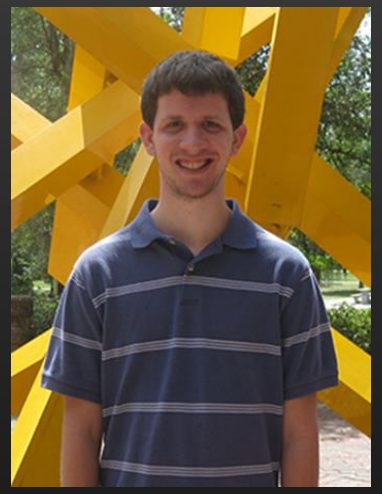
"2" recognized as "5" (6%)

Conclusion & Future Work

- Human accuracy (90.6%) significantly better than machine (84.14%)
- Significant effect of category on accuracy
- Future work
  - Bridging the gap between machine and human recognition
  - Comparing other types of recognizers for children's gestures

Image by Pixabay under Creative Commons 0 https://pixabay.com/en/people-white-phone-holding-kid-315907/

Thank You!



Alex Shaw
alexshaw@ufl.edu

Jaime Ruiz
jaime.ruiz@ufl.edu

Lisa Anthony
lanthony@ufl.edu