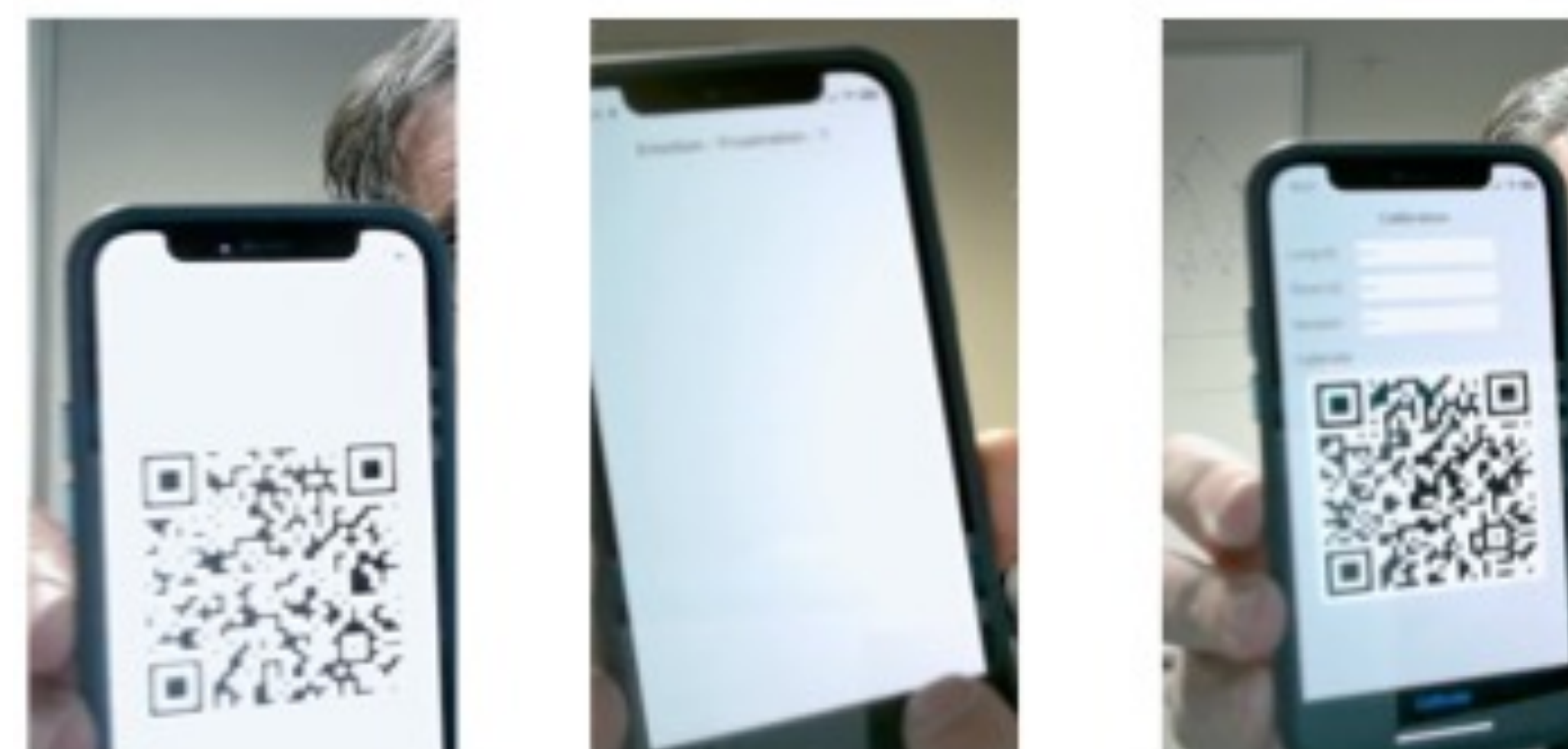


Abstract

- Real time synchronization of labels to recorded video frames
 - is feasible
 - uses low-cost hardware
 - allows for audiovisual recording and labeling to happen at the same time
- The recording platform
 - democratizes labeled video collection
 - by its low-cost hardware
 - by simplifying the labeling process
 - is ideal for undergraduate research
 - usability studies
 - novel message encoding
 - audio and visual processing
 - using new sensors

Experiment 1

Device Recording Procedure



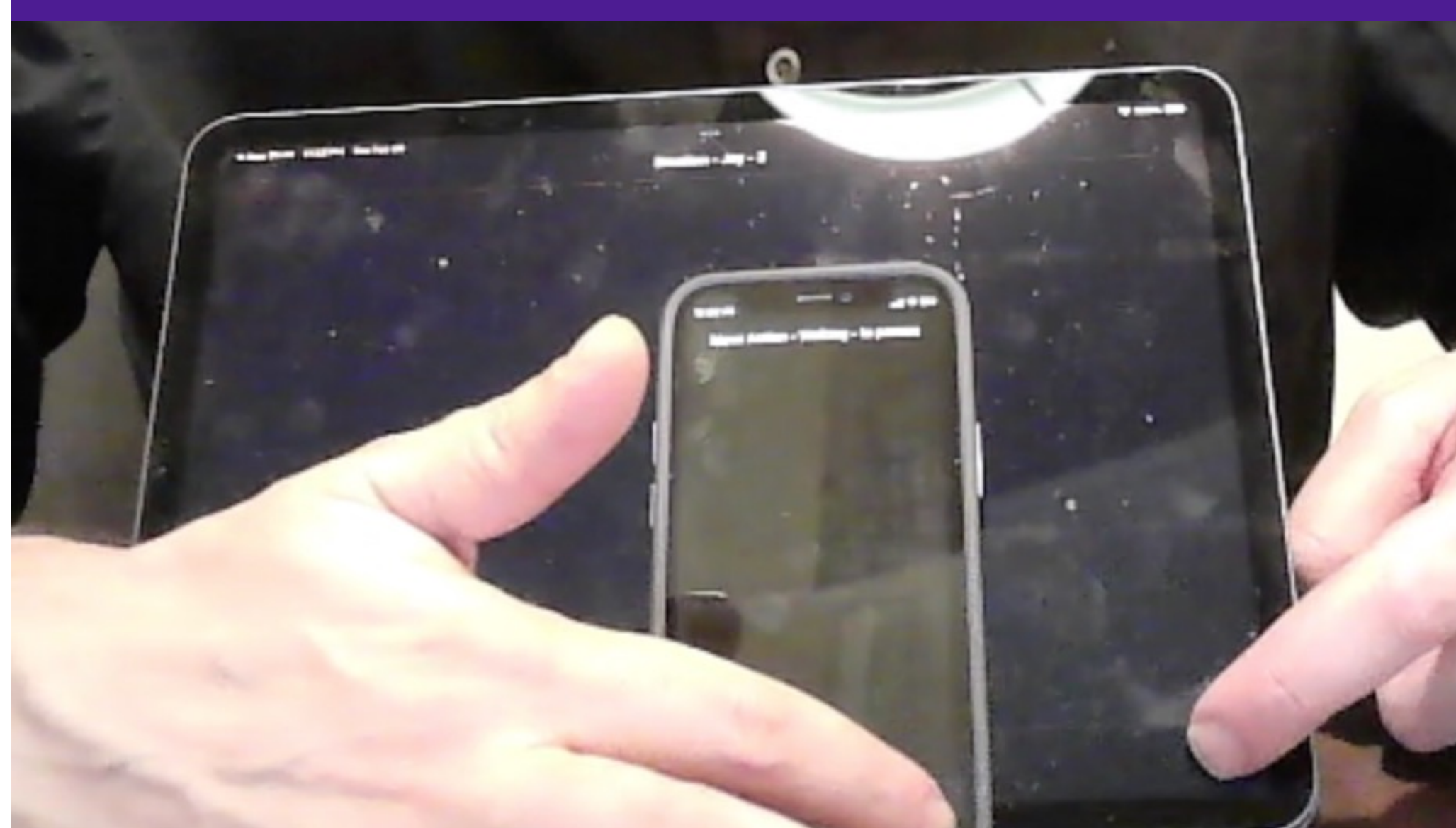
Begin Calibration
(> 1 sec)

Send Label
5-7 Times

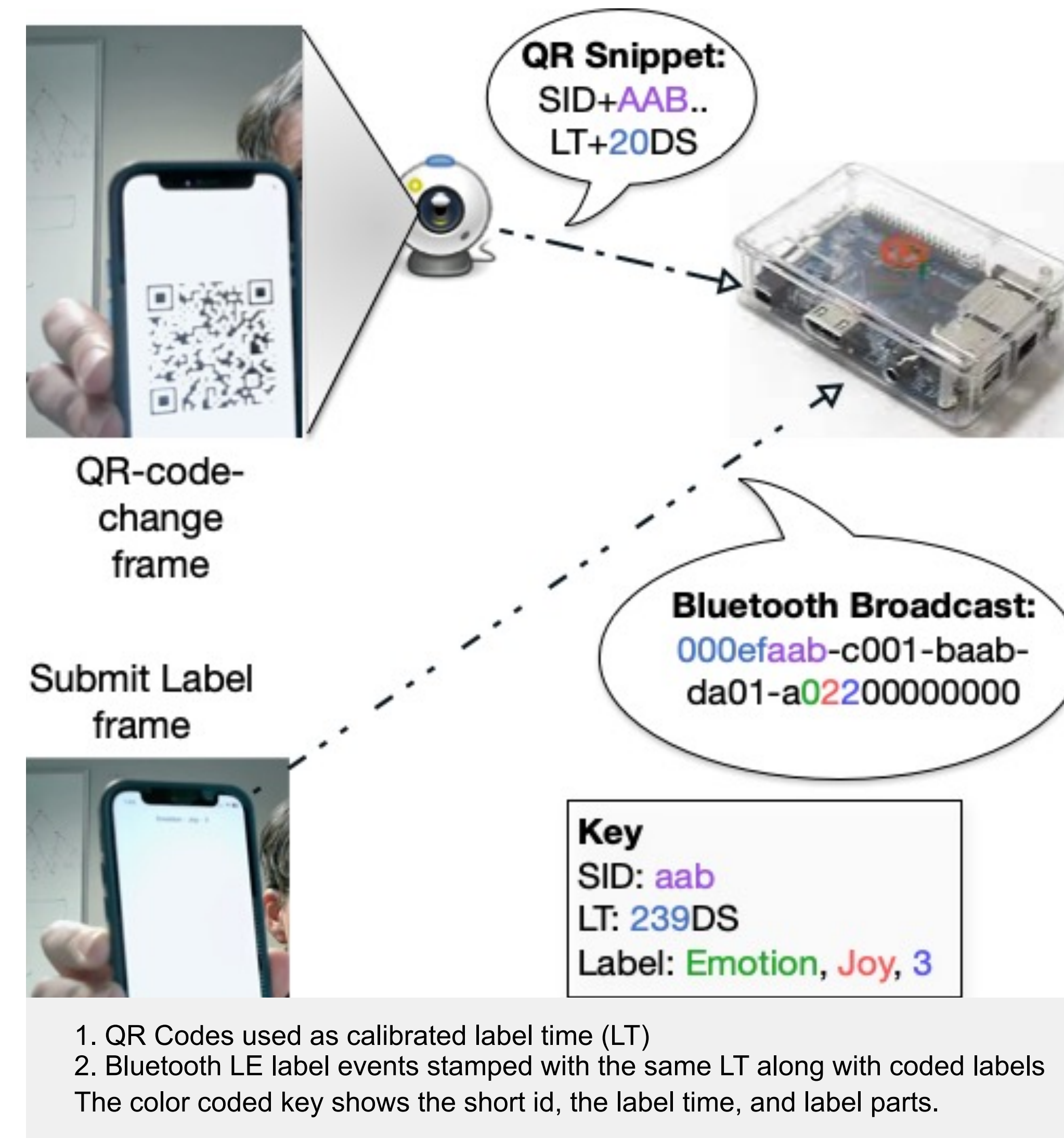
End Calibration
(> 1 sec)

Trials capture multiple label events over video and Bluetooth LE to verify that synchronization is working using an iPhone for labeling.

Experiment 2



Trials capture multiple label events over video and Bluetooth LE to verify that synchronization is working using an iPhone and an iPad for labeling.

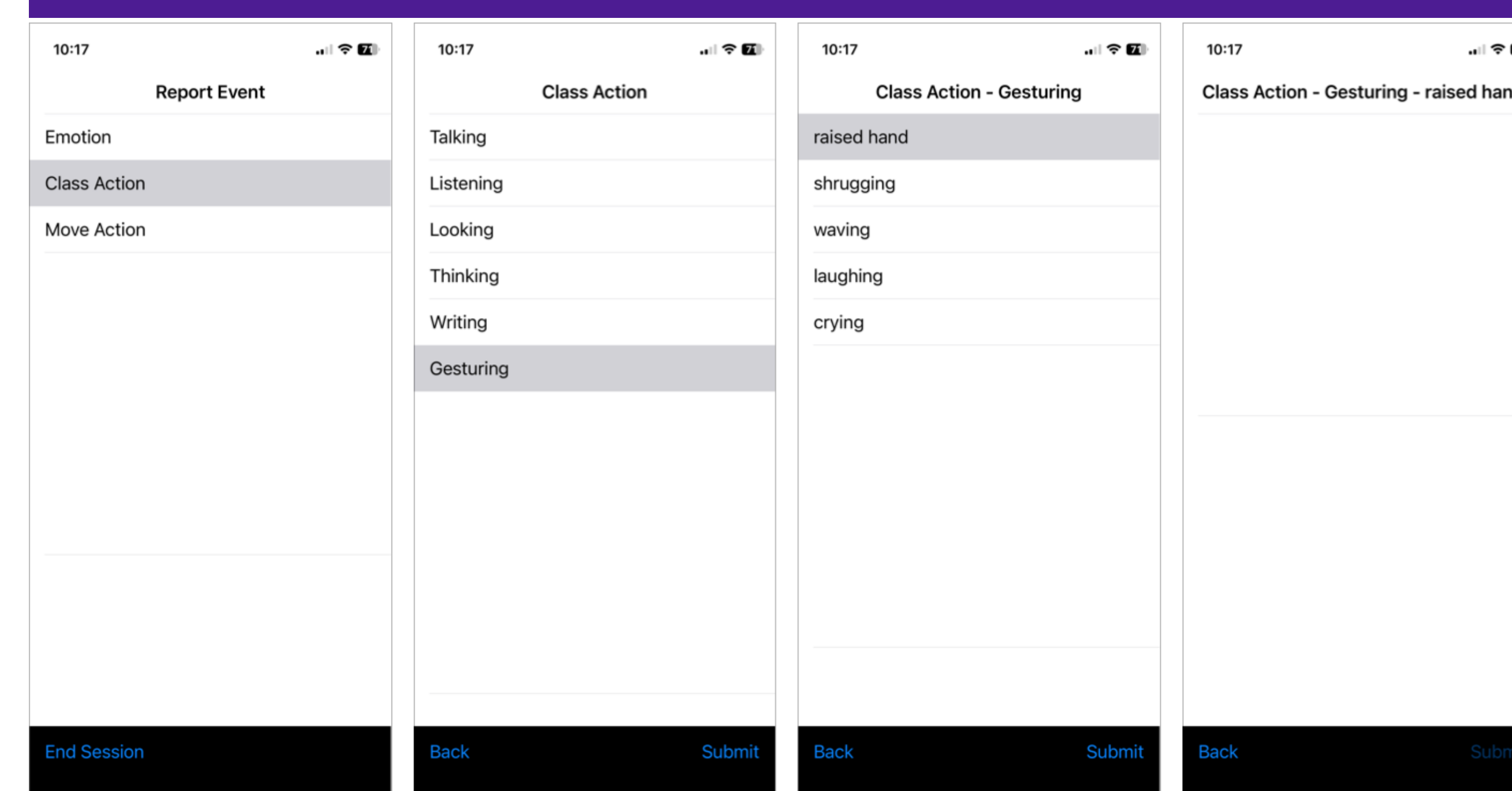


- Potential Extensions:
 - different label UI: pie menu, single click, grid
 - add duration to labels
- Potential Recording Applications
 - classroom – instructor, students in desks
 - group work at tables
 - office – student(s) at office hours
 - faculty meeting together
 - human participants research
 - sporting events
- Additional sensor questions
 - how to overcome limitations of outdoor recording
 - how to efficiently record 3D video
 - can QR code synchronization be added to IOT sensors

Trial (Exp 1)	Drift FCT	Drift CCT	Drift FQR	Drift CQR
1	192.7	fails	0.6	0.1
2	0.9	-0.1	0.9	-0.05
3	1.5	0.5	0.91	0.18
4	-0.2	-1.2	0.7	-0.03
Trial (Exp 2)	Drift FCT	Drift CCT	Drift FQR	Drift CQR
1: iPhone	0.7	-0.25	0.95	-0.04
1: iPad	0.28	-0.59	0.3	0.00
2: iPhone	1.7	0.7	1.6	0.04
2: iPad	1.95	0.93	0.9	-0.00

Average drift numbers in seconds between automatically calibrated times and hand annotated times of the label frames.

Example Labeling Sequence: Class Action – raised hand

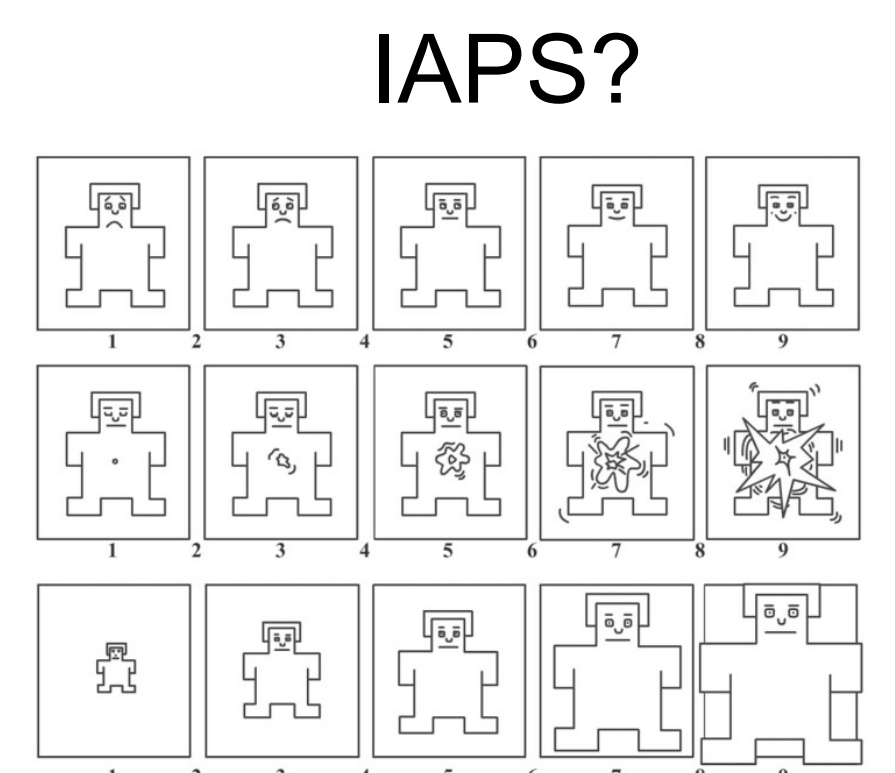


Labeling steps (diagram to the left)

- Precondition: Report Event screen available after calibration
- Notice or experience an action or emotion
- select type of event – Class Action
- select type of Class Action – Gesturing
- select type of Gesturing – raised hand
- Click "Submit" button
- repeat at beginning.
- Postcondition: After ending session show QR code to camera for at least one second.

Use Case: UI Study

- Student develops new UI
- Participants play a board game, some use new and some use original UI.
- Get Participant feedback
- Students rate the labeled video.
- UI comparison examples
 - Pie vs. table view
 - single click vs. table view
 - grid vs. table view
 - pictures vs. text



Use Case: Classroom study

- Determine how many participants to record at a time
- Determine labels
 - Cornell Activity Data Set - 60 (12 activities)
 - Cornell Activity Data Set-120 (10 activities from 10 actions)
 - International Affective Picture System (IAPS)
 - a custom label set
- Students self report as class is recorded.

Other Sensors



- Respeaker – 3D microphone
- Localize sound in space



- OAK-D – depth camera
- 2D video, plus depth

- Research challenges
 - differentiate people locations with other sound locations
- Research challenges:
 - efficient depth video
 - cross sensor calibration

Related Work

- S. Sridhar, P. Misra, G. S. Gill, and J. Warrior, "Cheepsync: a time synchronization service for resource constrained bluetooth le advertisers," IEEE Communications Magazine, vol. 54, no. 1, pp. 136–143, January 2016.
- J. Silva, D. Cabral, C. Fernandes, and N. Correia, "Real-time annotation of video objects on tablet computers," in Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, 2012, pp. 1–9.