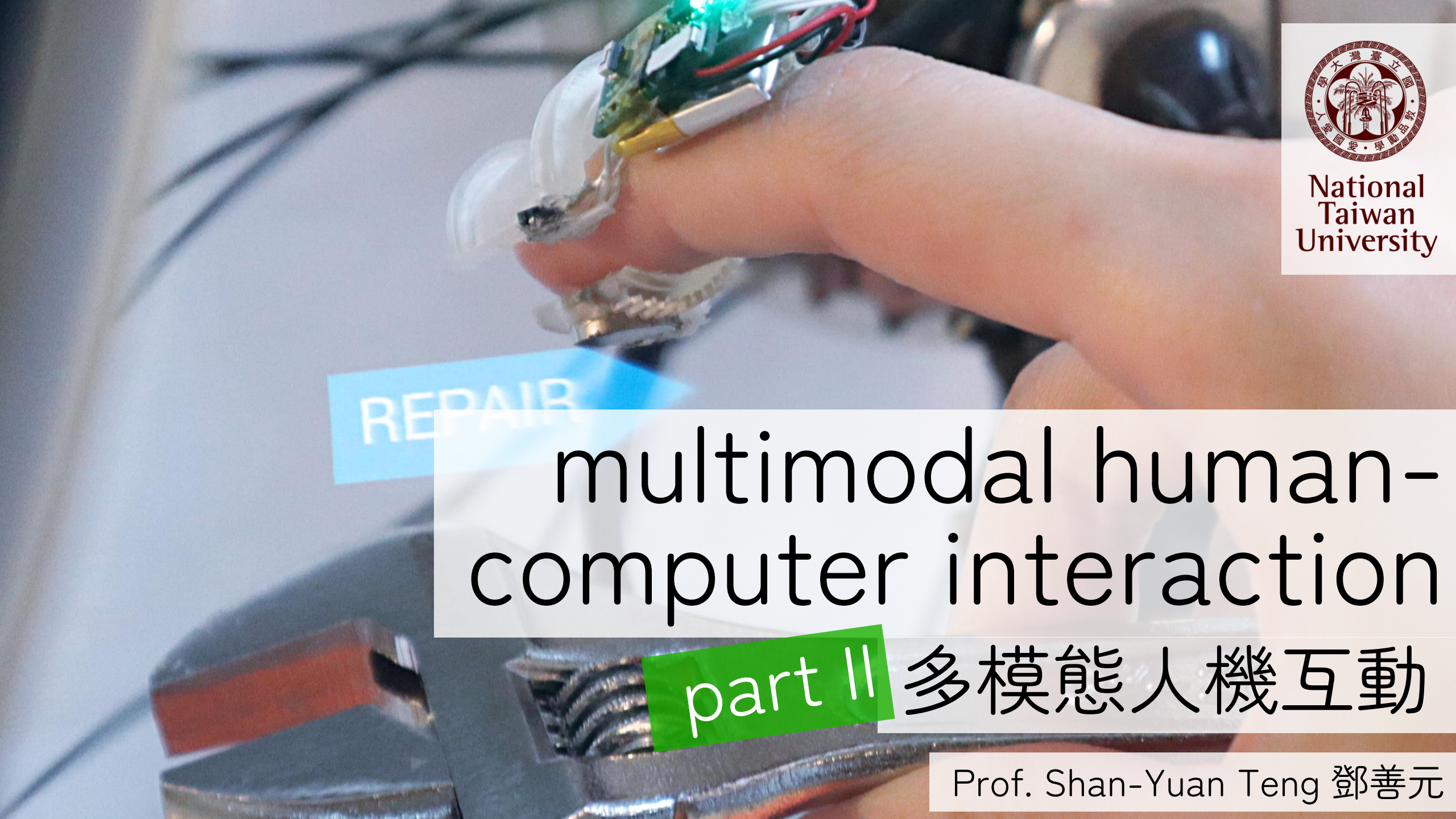


National Taiwan University

multimodal human-computer interaction

part II 多模態人機互動

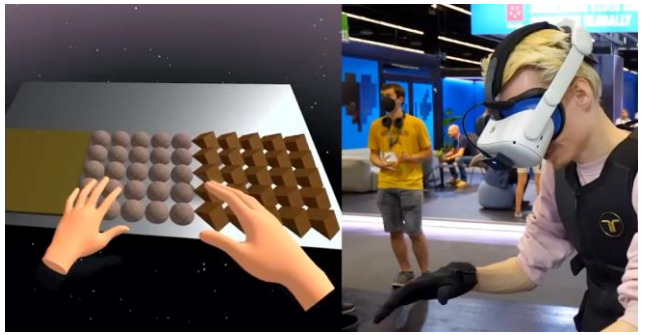
Prof. Shan-Yuan Teng 鄧善元

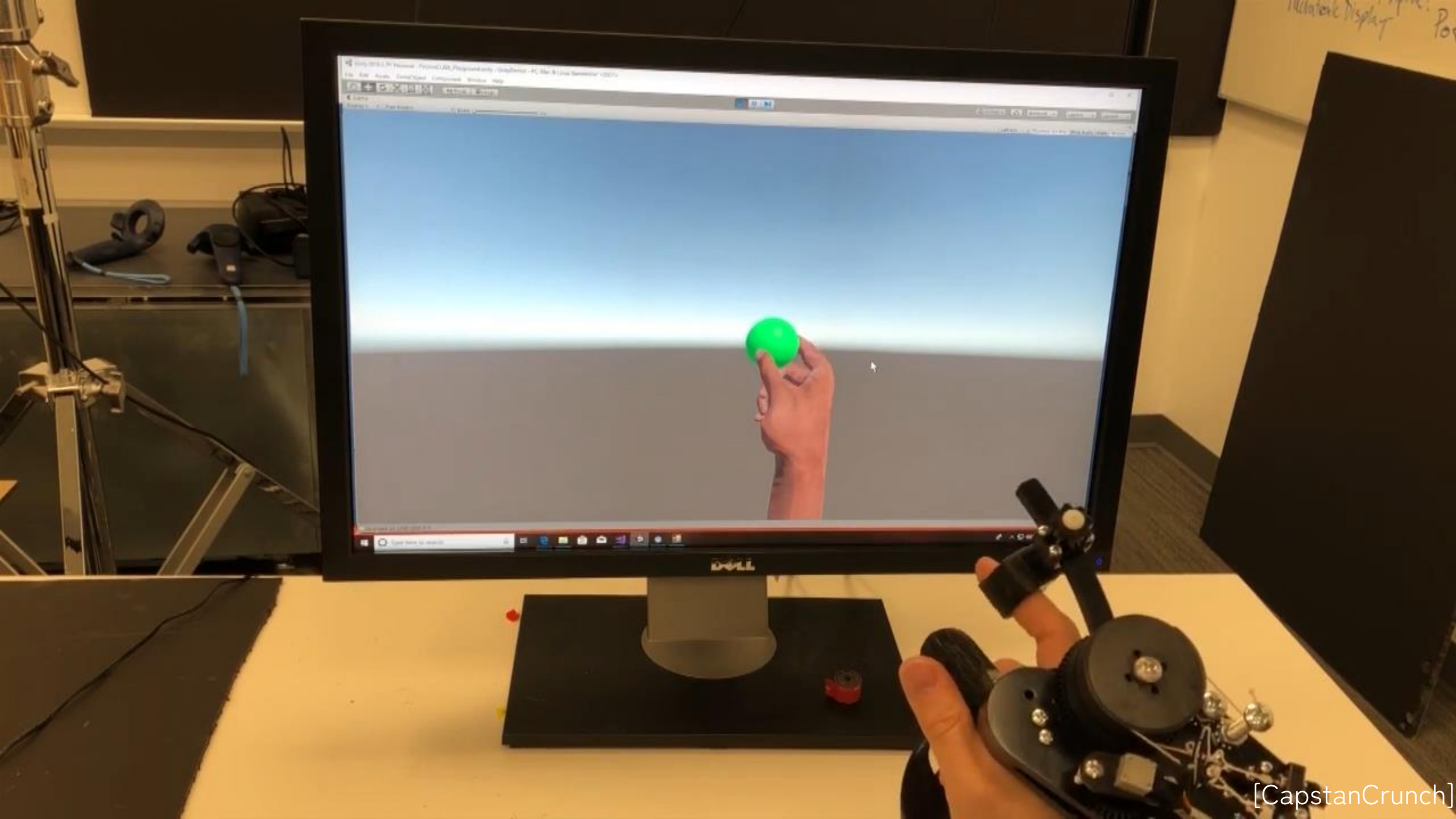
benefits of multimodal HCI:

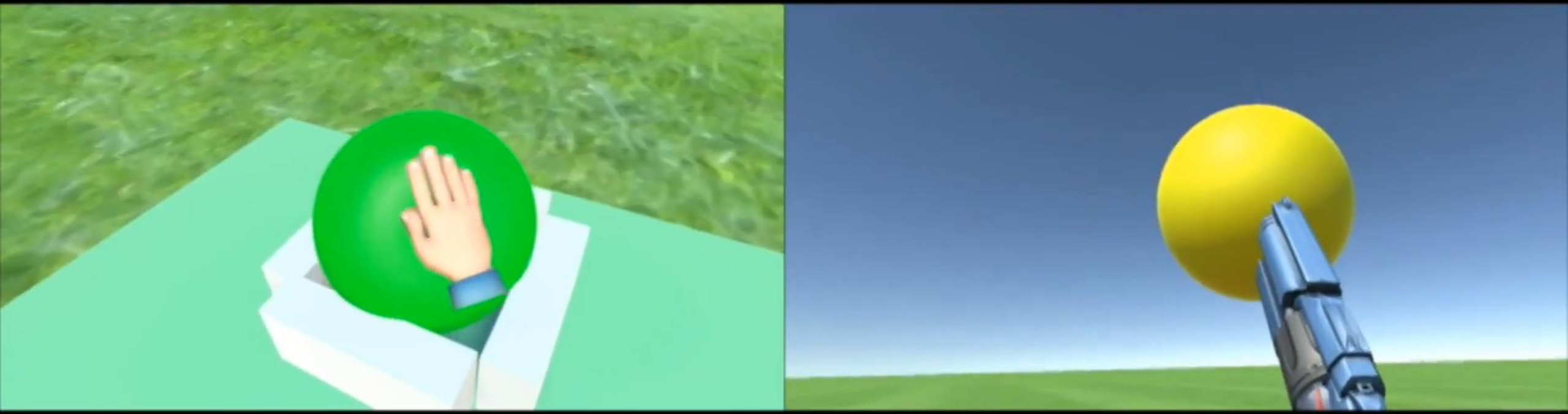
1. efficiency
   make computers more efficient

2. usability
   make computers easier to use

3. fidelity
   make computers feel more realistic

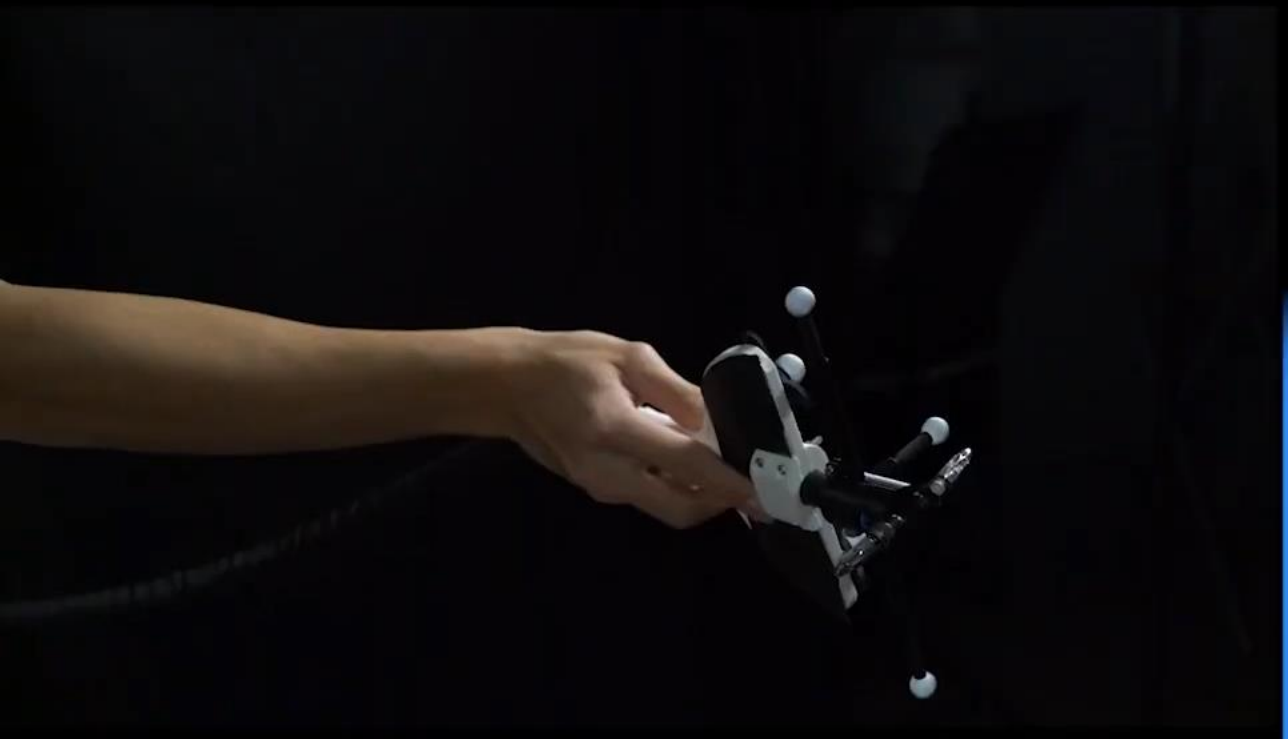


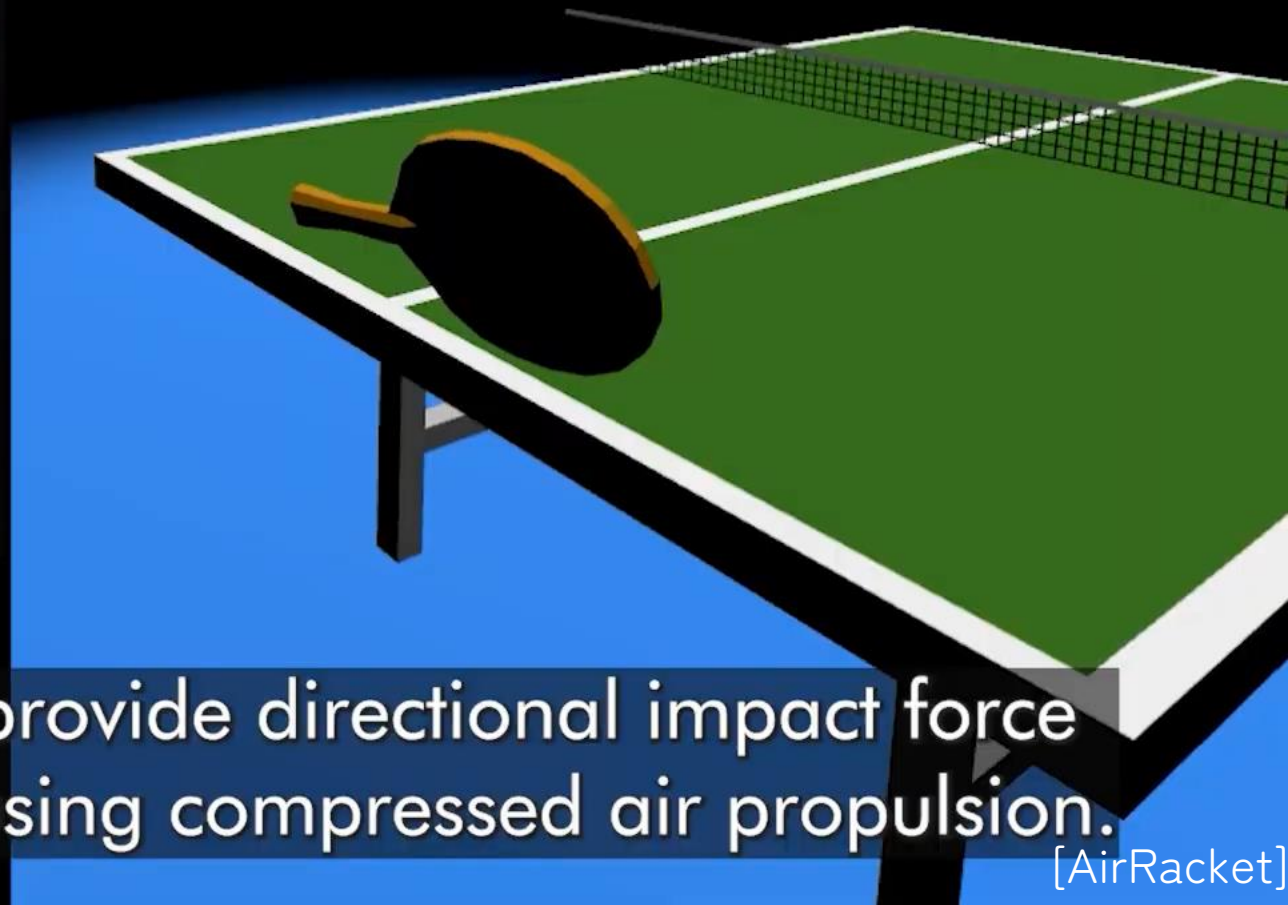

[CapstanCrunch]



To provide multilevel continuously-changed resistive force

[ElasticVR]

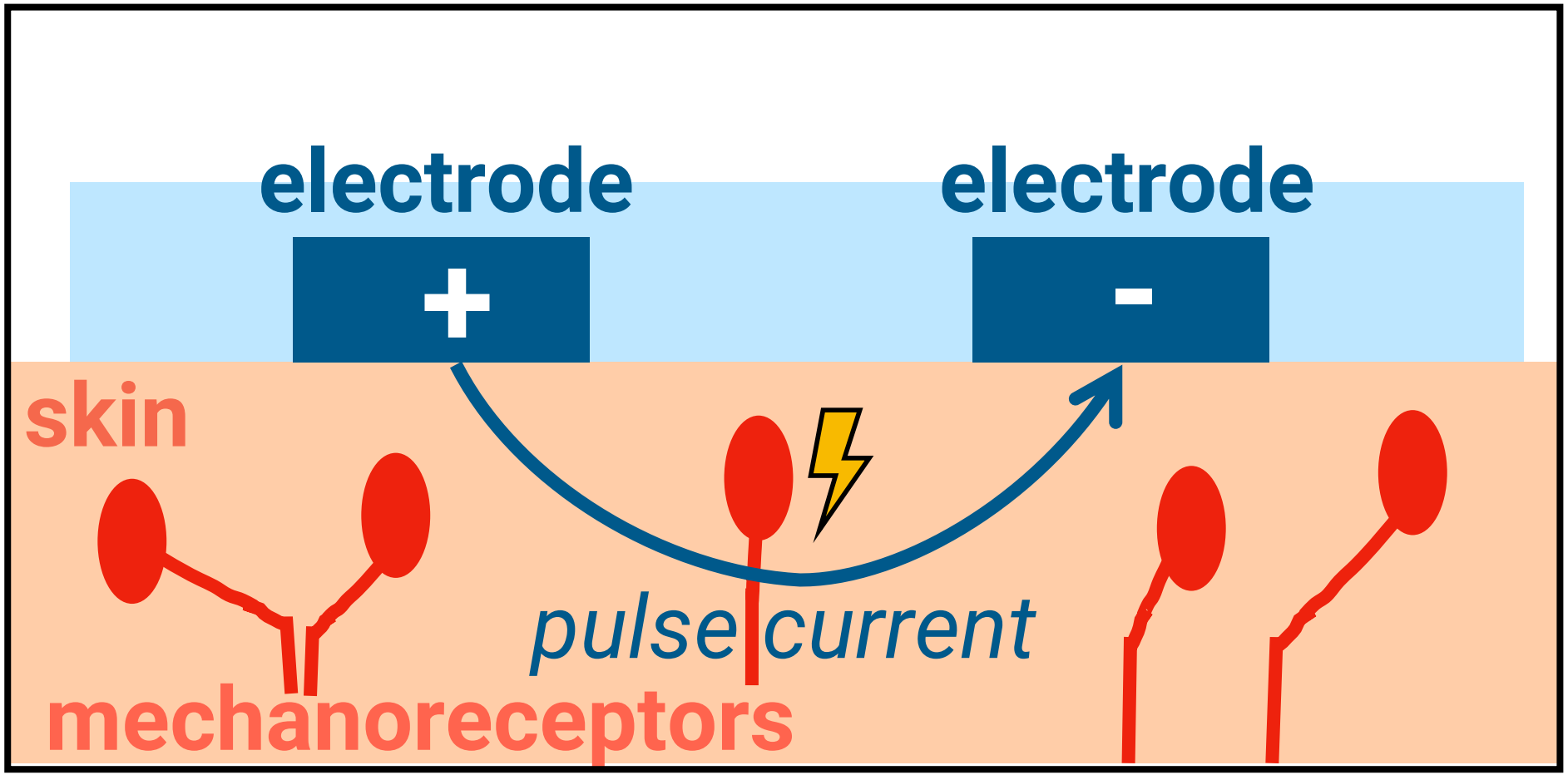




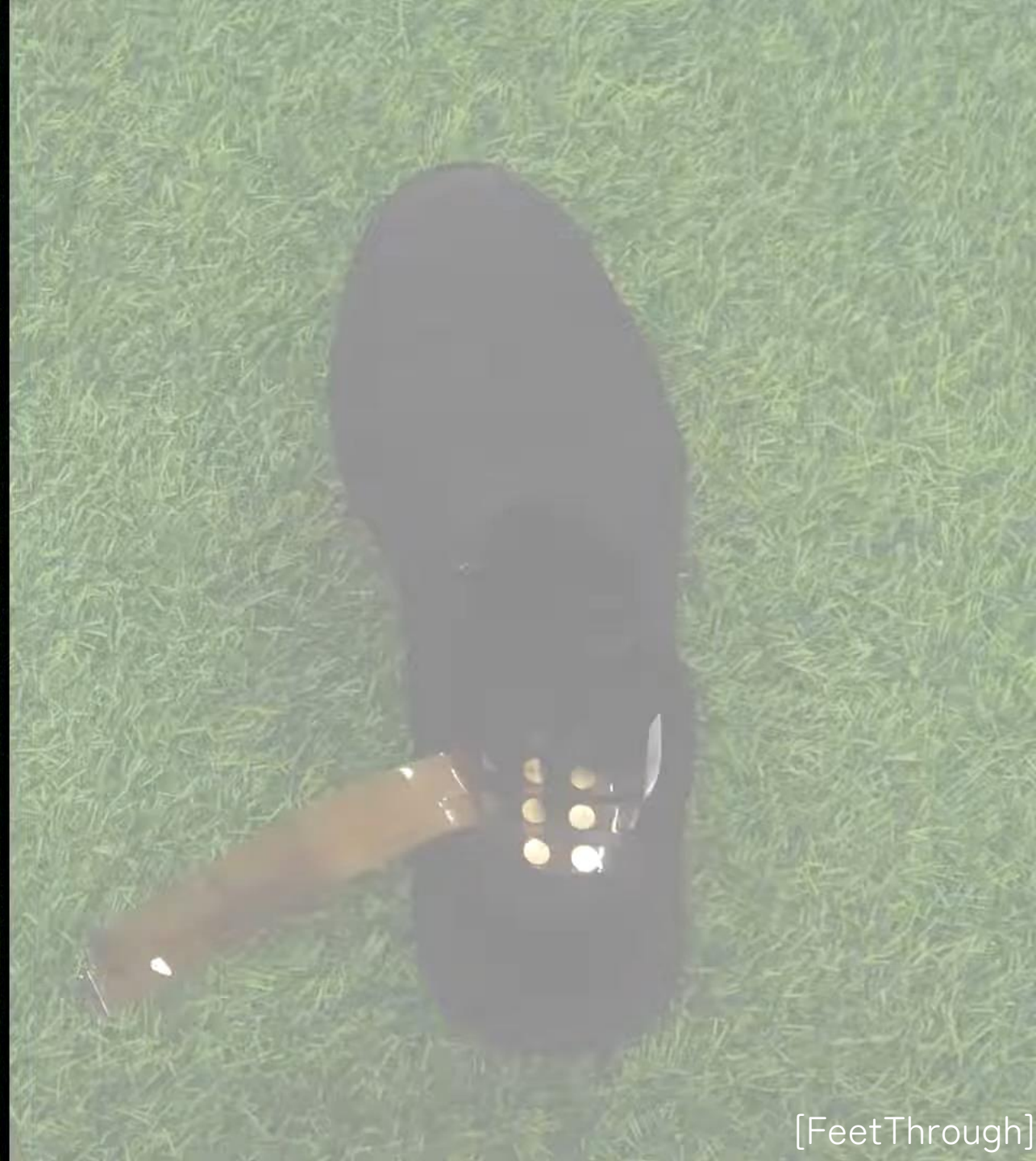
AirRacket is the first device to provide directional impact force feedback for VR racket sports, using compressed air propulsion.

[AirRacket]

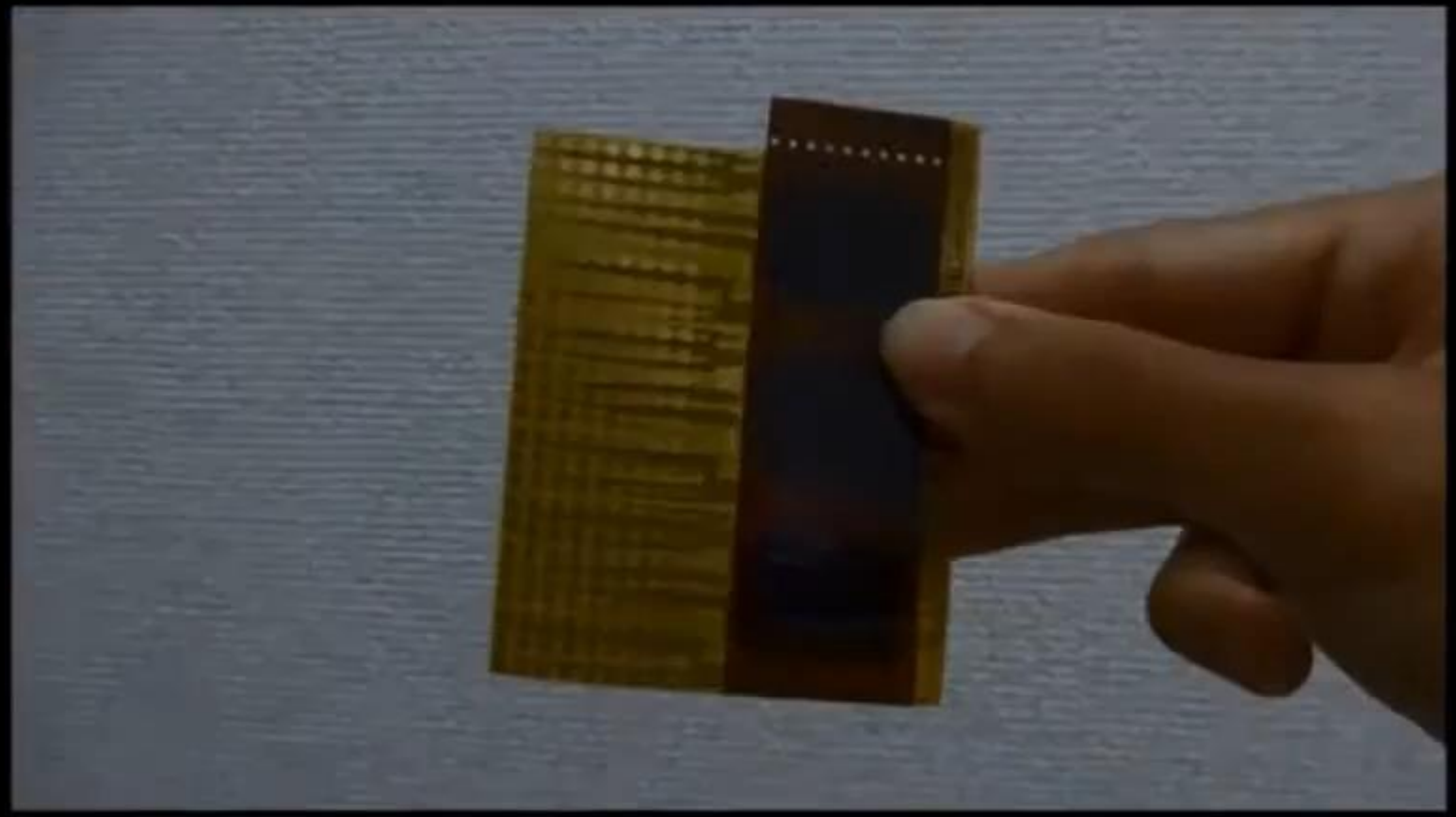
electrotactile feedback




traditional feet haptics = **vibration**


[FeetThrough]

Electro-Tactile Display

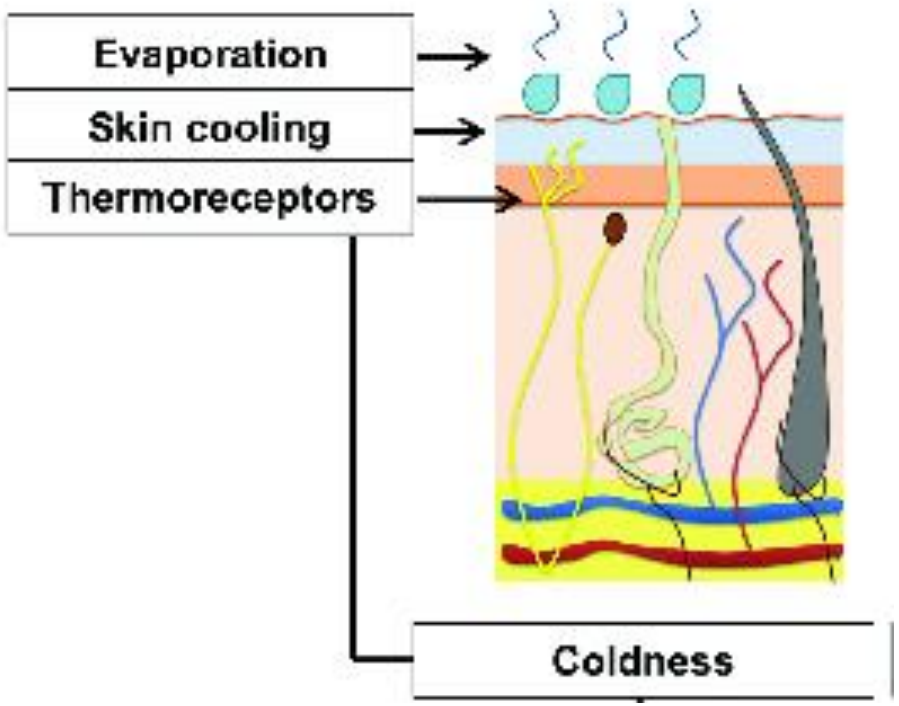


However, electro-tactile display lacks of feedback corresponding to touching motion.

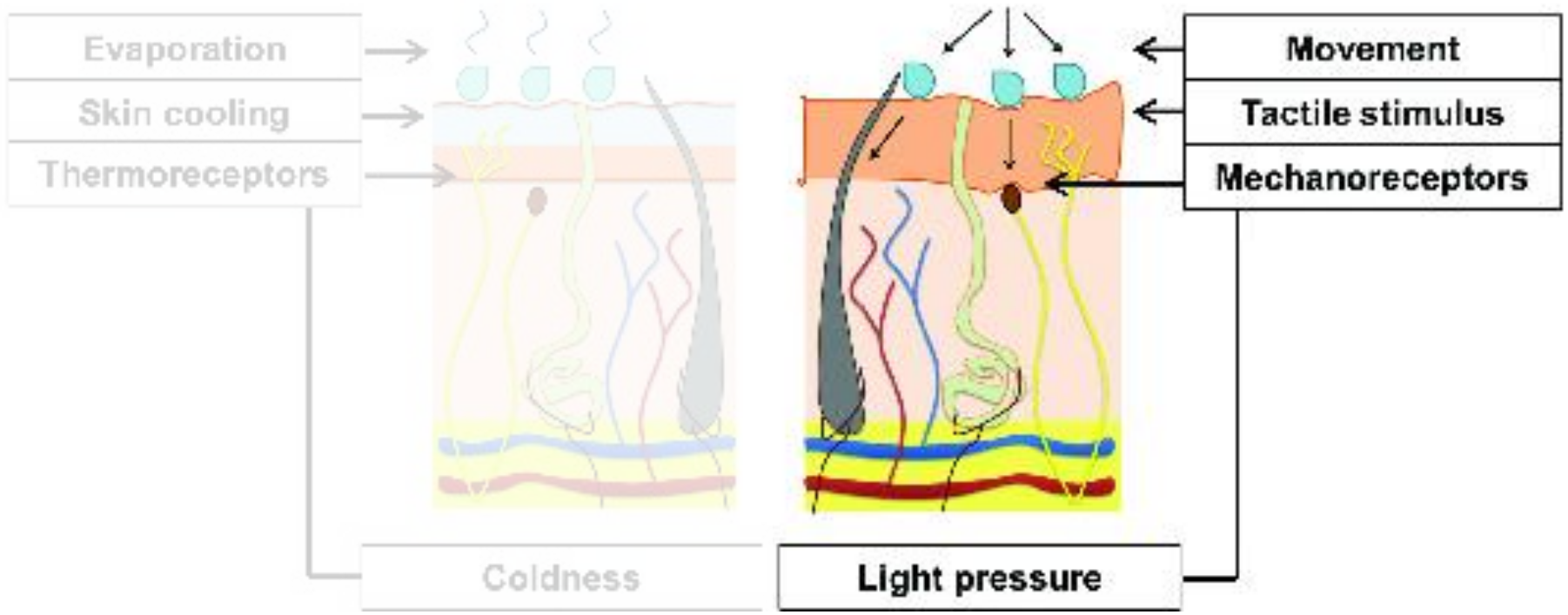
[Takei et al]

how do we sense water (wetness)?

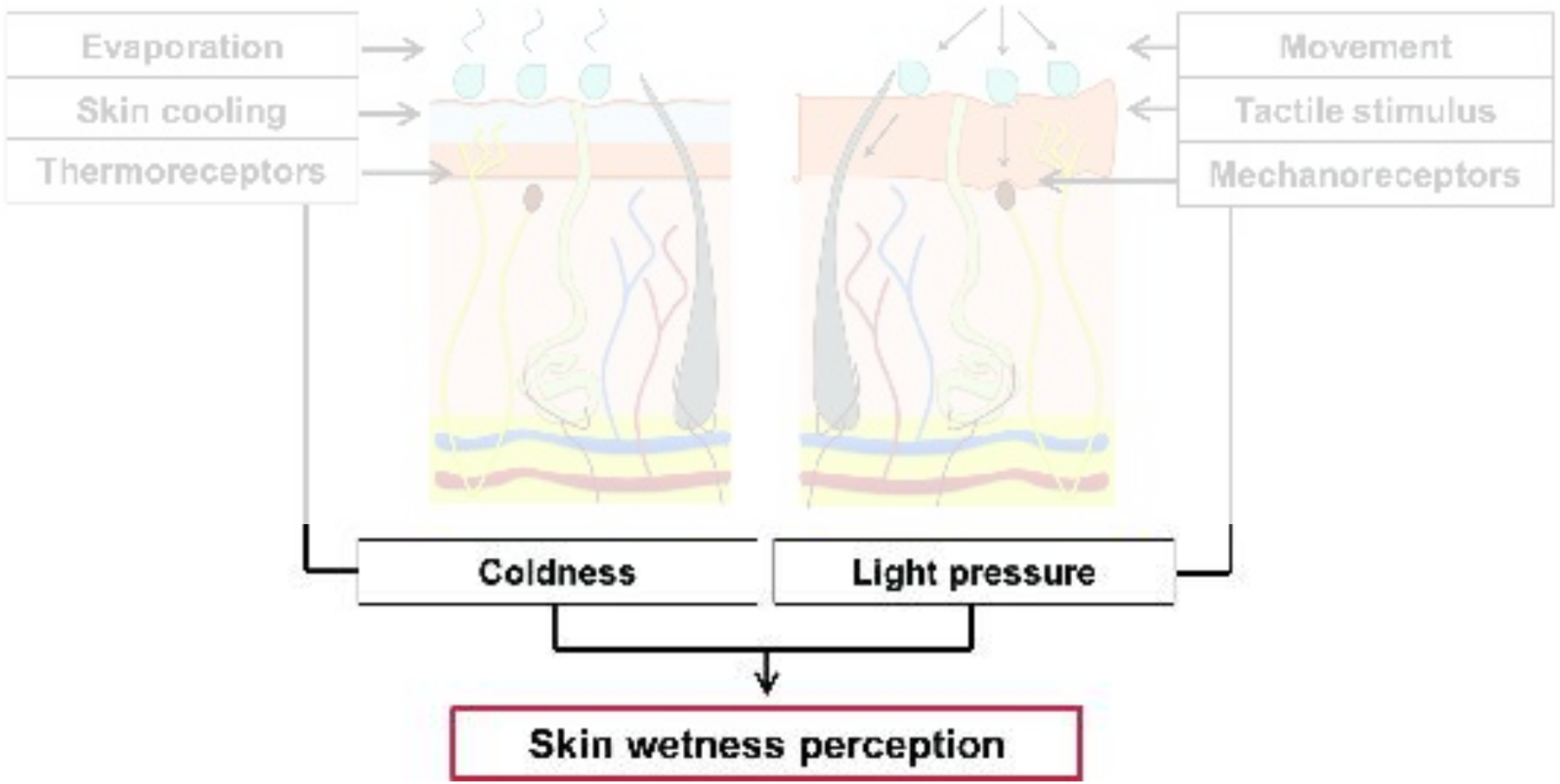
how do we sense water (wetness)?

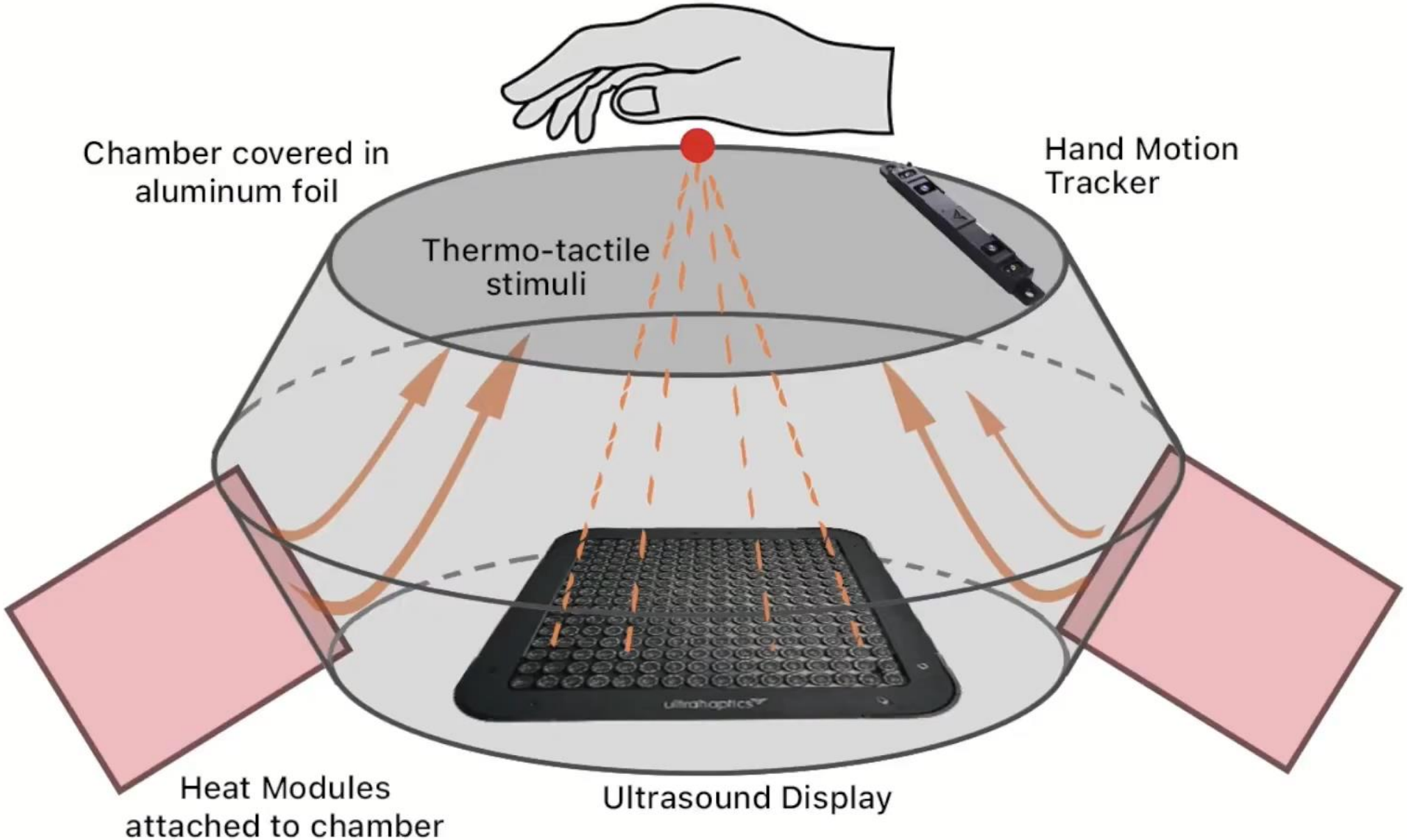


how do we sense water (wetness)?



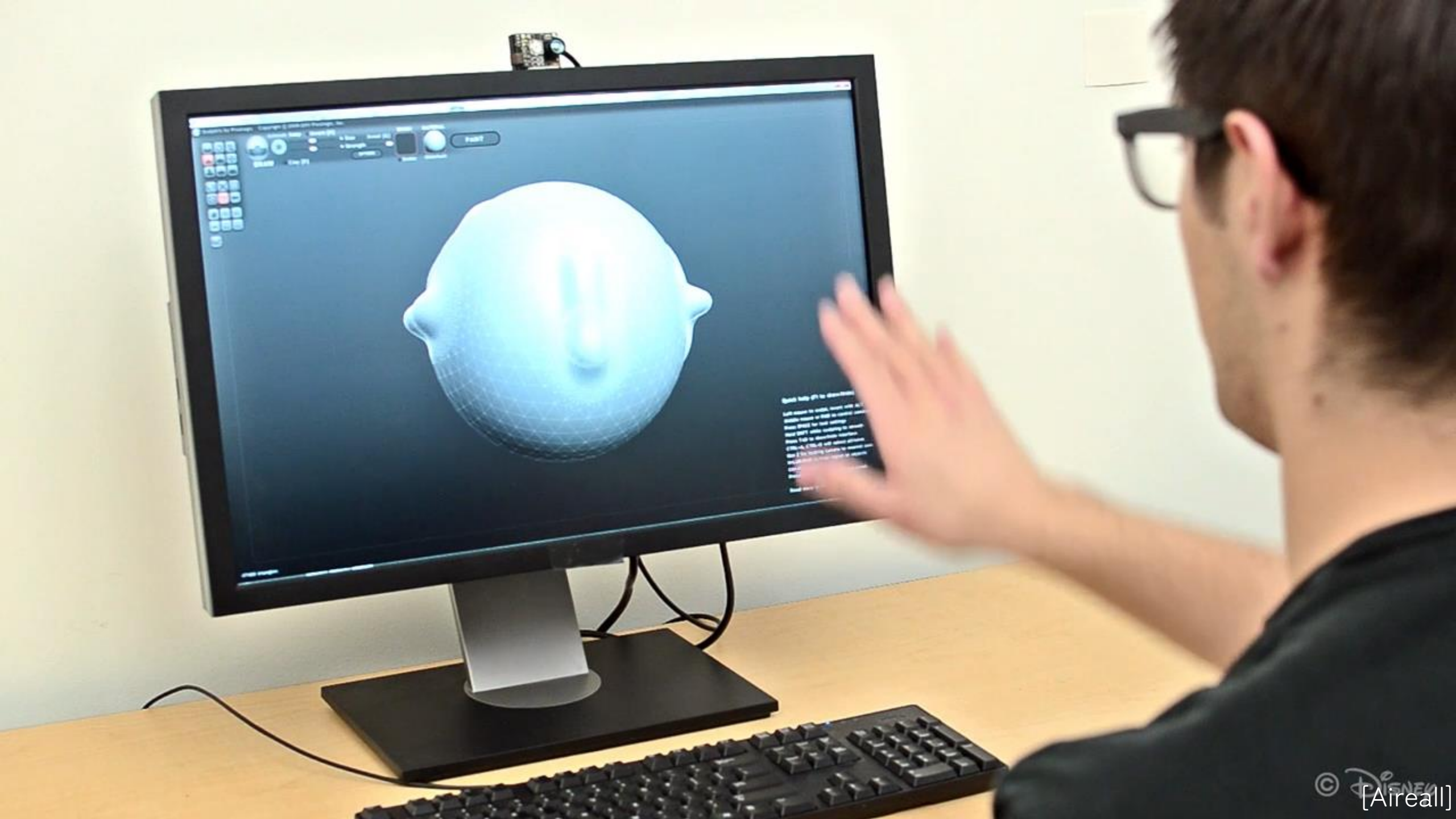
how do we sense water (wetness)?





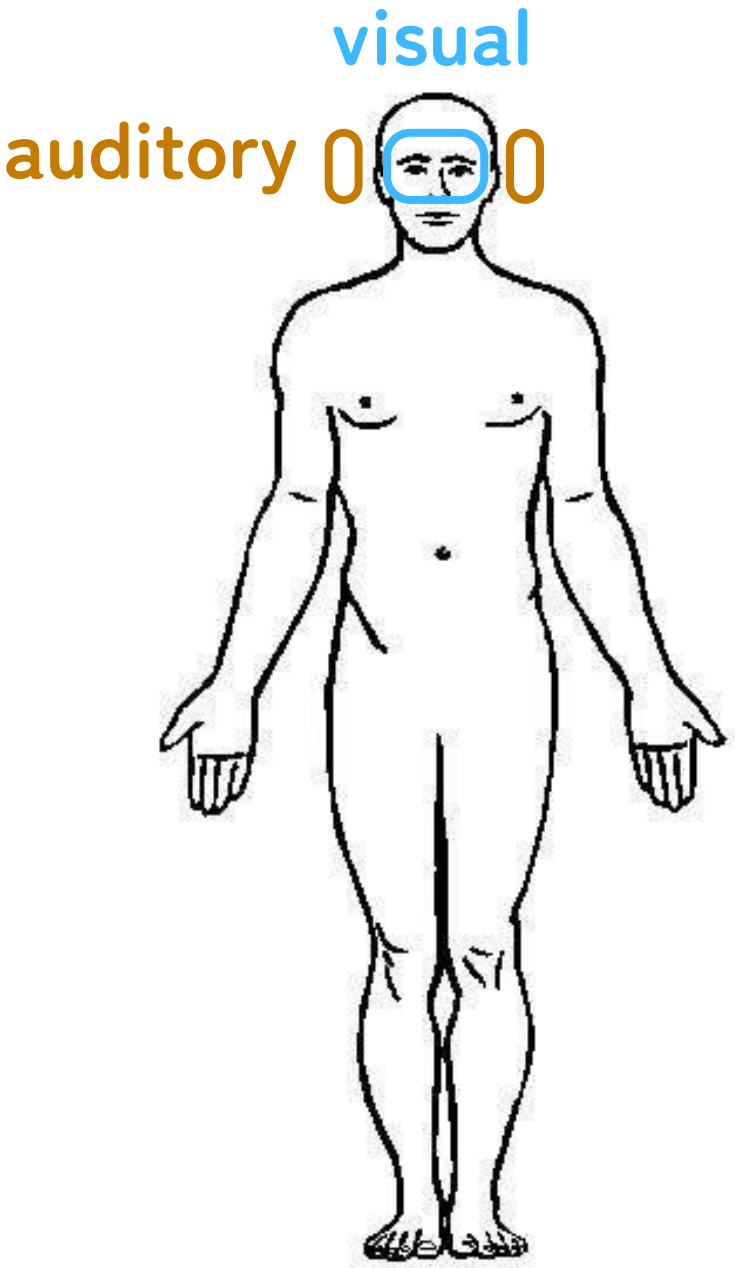
Our proposed system provides thermo-tactile feedback in mid-air.

[Singhal et al]

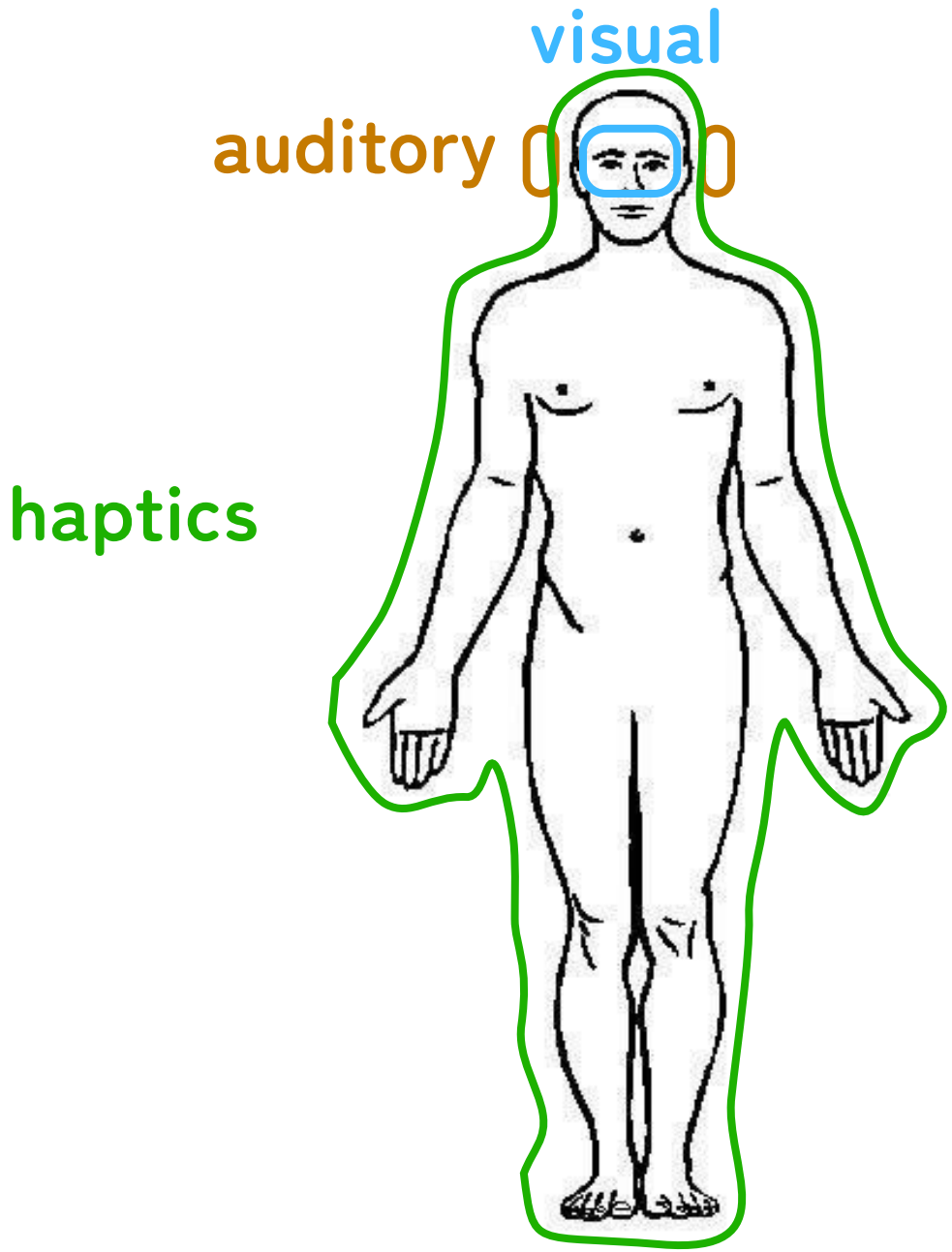

© Disney
[Aireall]

summary:
so much haptics, but very
little we can replicate

summary:
so much haptics, but very little we can replicate

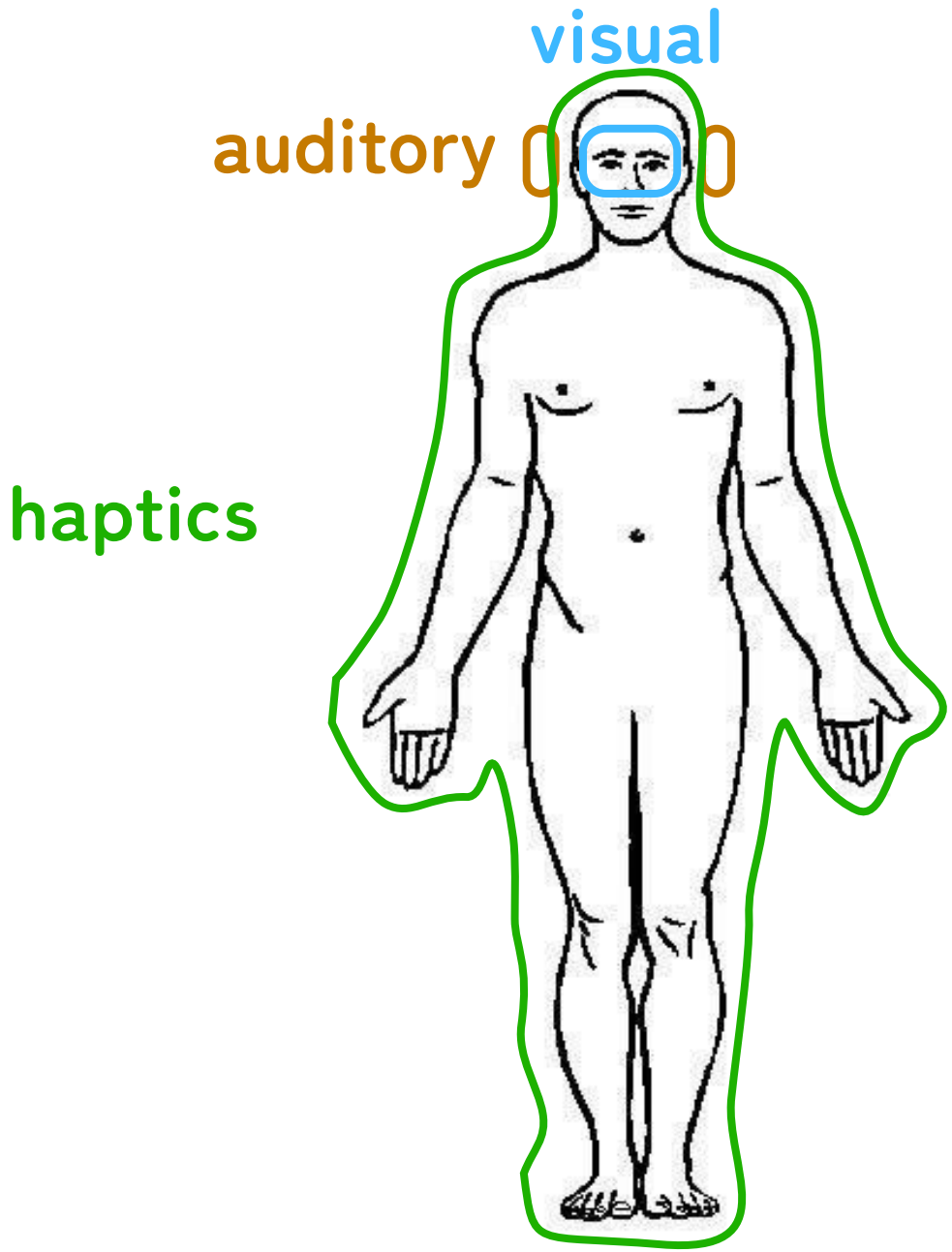


summary:
so much haptics, but very little we can replicate



summary:
so much haptics, but very little we can replicate

open challenge:
big device, small effects



PuPoP:

Pop-up Prop on Palm for Virtual Reality

Shan-Yuan Teng, Tzu-Sheng Kuo, Chi Wang, Chi-huan Chiang, Da-Yuan Huang, Liwei Chan, Bing-Yu Chen

National Taiwan University, National Chiao Tung University, National Taiwan University of Science and Technology

國立臺灣大學
National Taiwan University

國立交通大學
National Chiao Tung University

國立臺灣科技大學

haptic-visual illusion

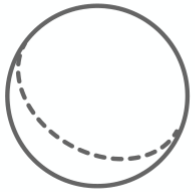
Physical Prop

≈

Acceptance Range

Visual Counterpart

haptic-visual illusion

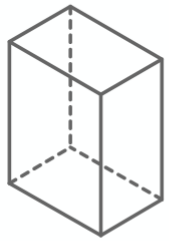


Sphere

Physical Size	***Small***	***Medium***	***Large***
Visual Size Upper Bound	63.2	75.5	95.3
Visual Size Lower Bound	46.0	60.5	73.4

Cylinder

Physical Size	***Small***	***Medium***	***Large***
Visual Size Upper Bound	38.0	54.4	70.8
Visual Size Lower Bound	28.7	42.4	54.1



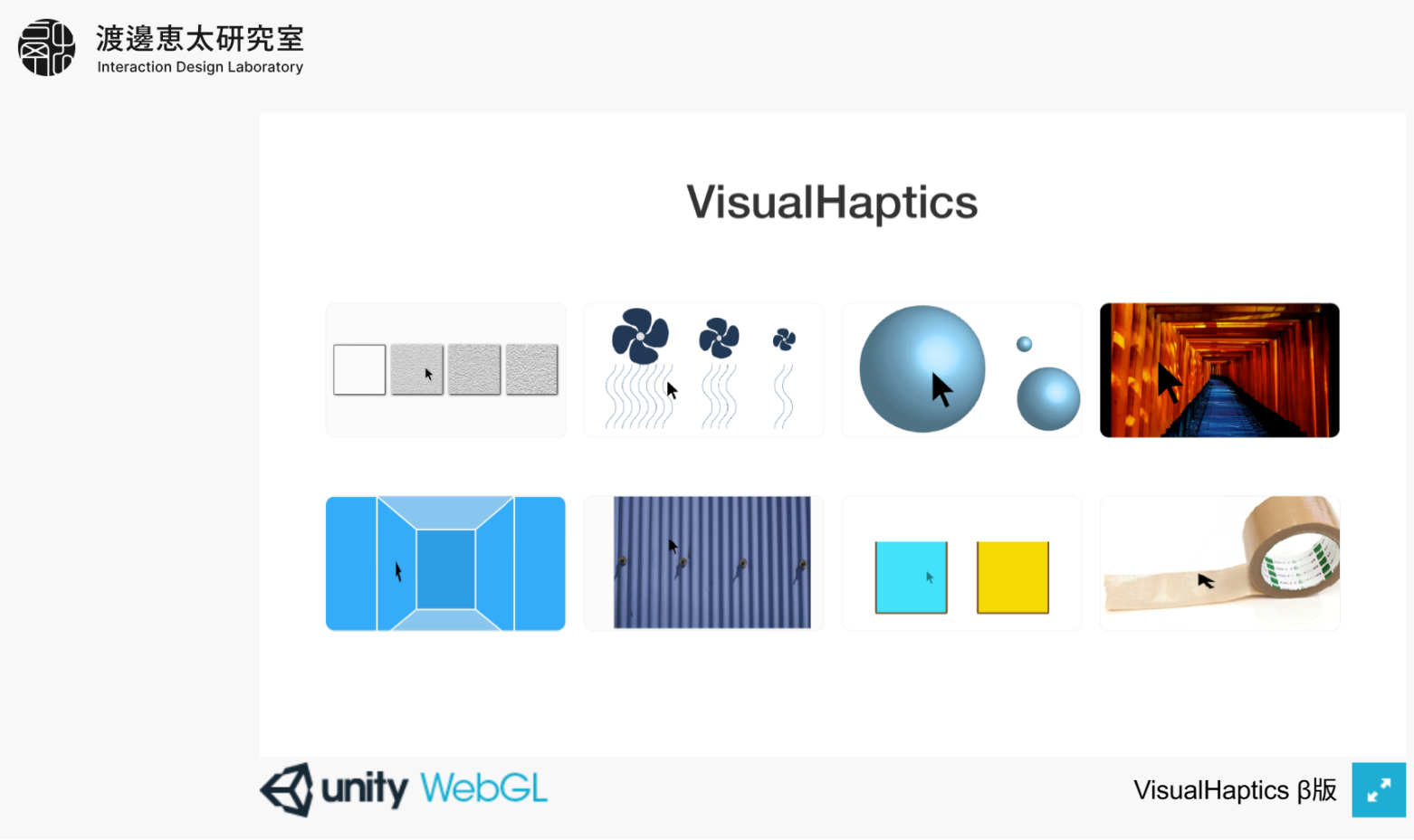
Box

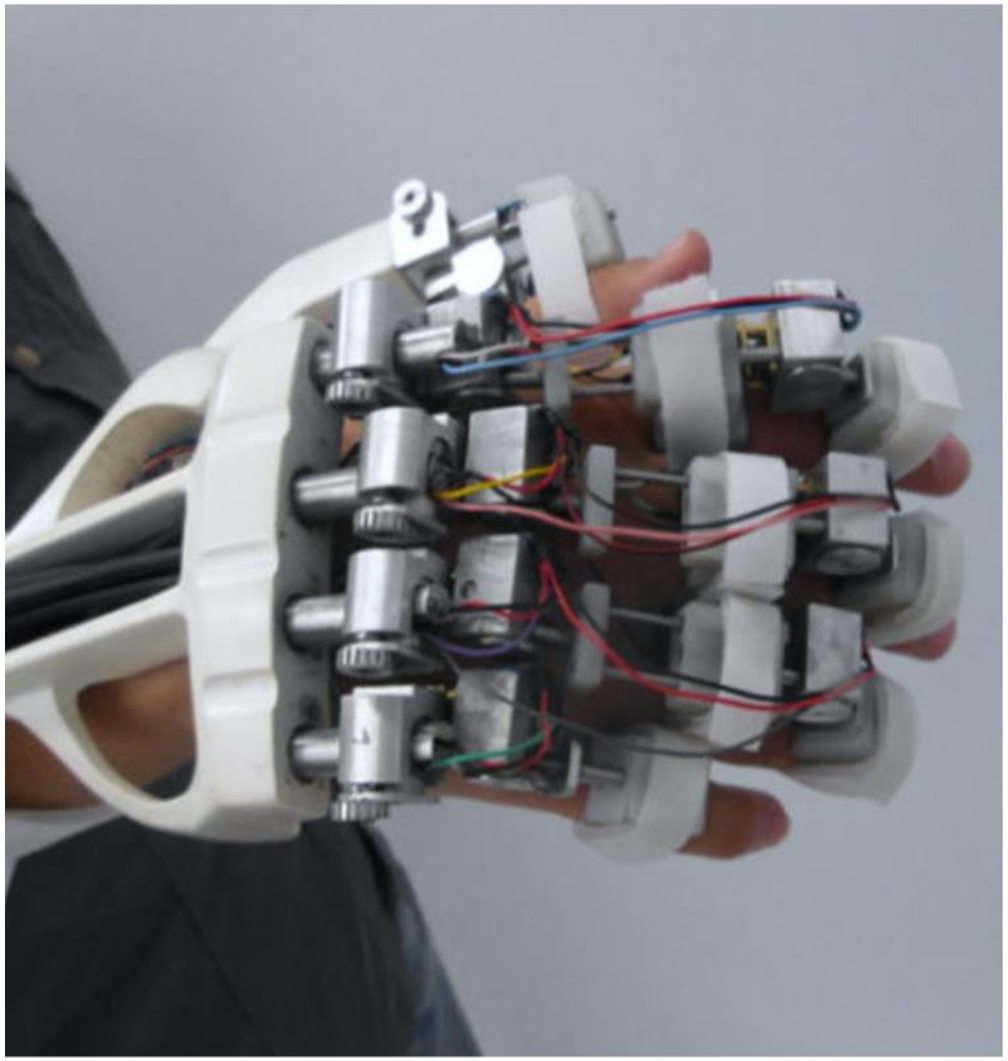
Physical Size	***Small***	***Medium***	***Large***
Visual Size Upper Bound	59.2	69.7	88.7
Visual Size Lower Bound	43.7	53.4	68.2

pseudo-haptics: generating the sense of touch without stimulating the tactile receptors

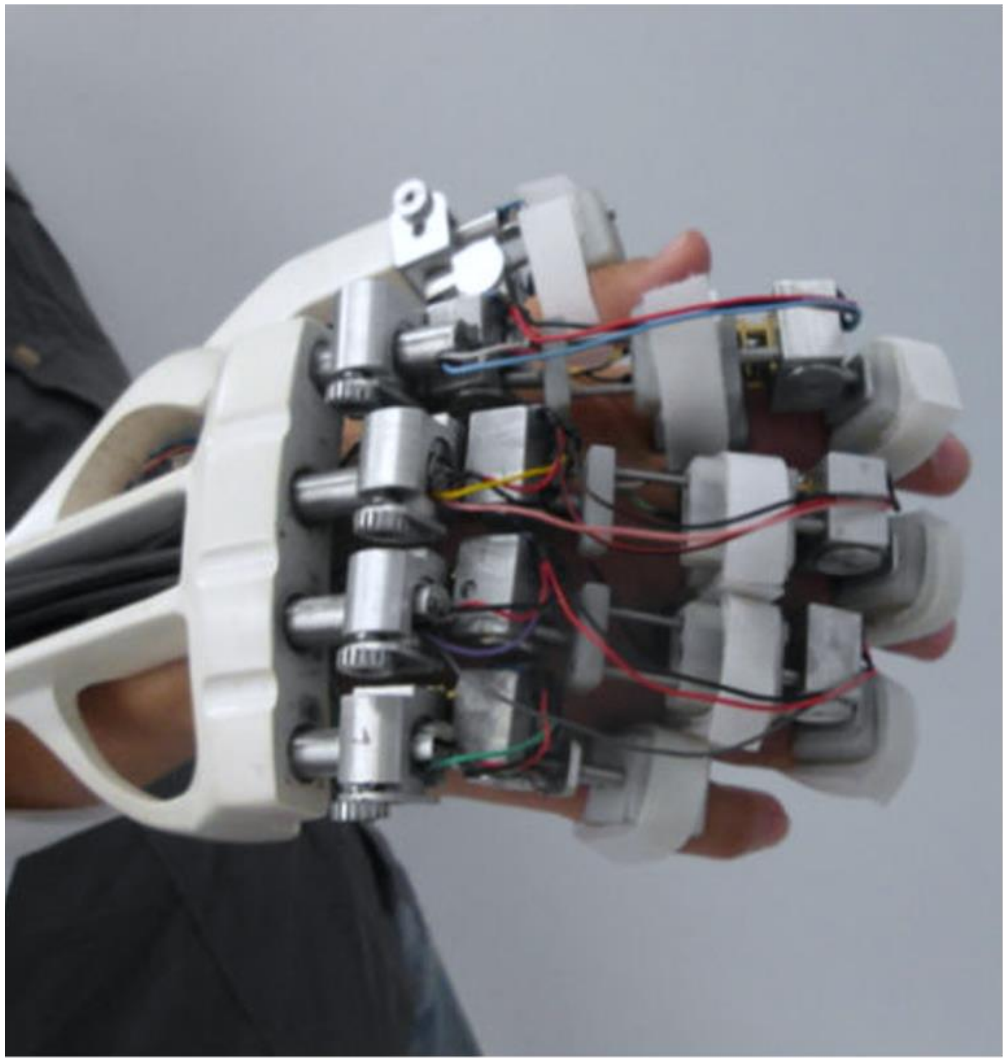
pseudo-haptics: generating the sense of touch without stimulating the tactile receptors

keita-lab.jp/projects/VisualHaptics





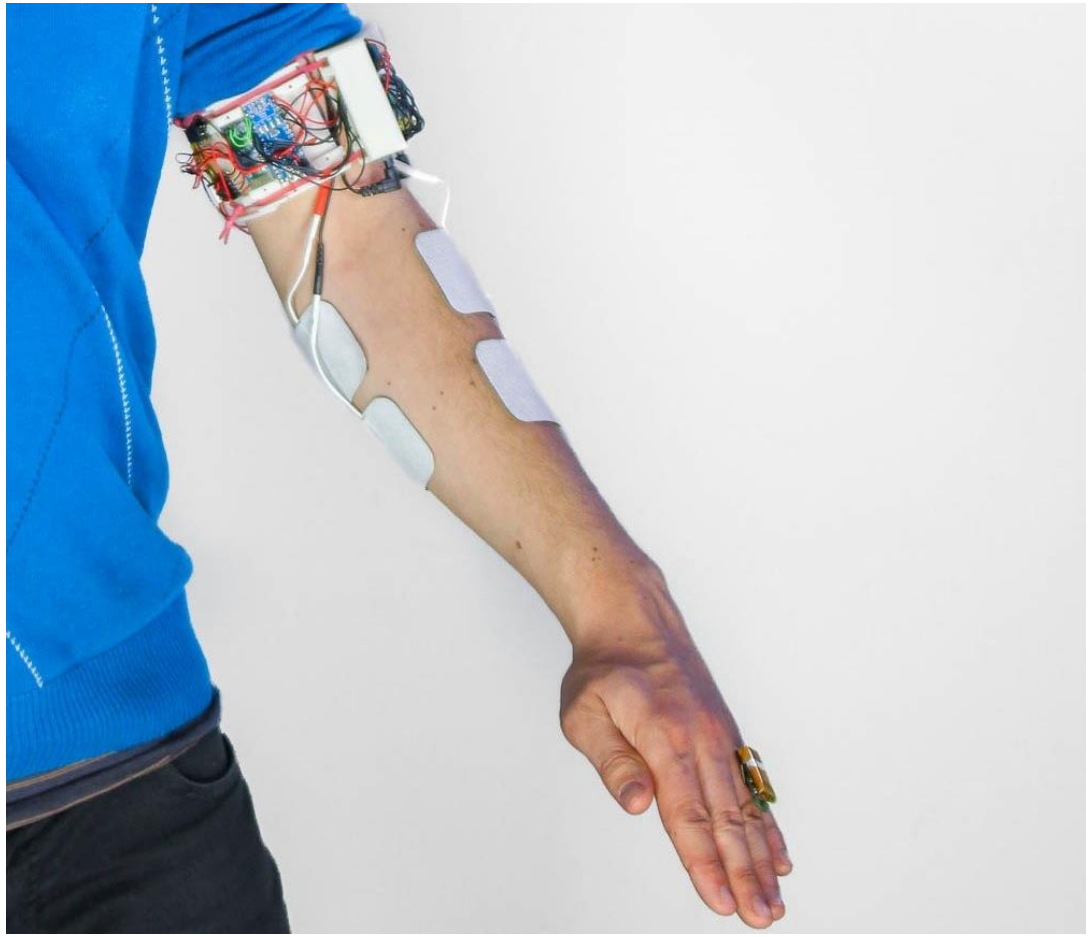
[Sandoval-Gonzalez et al., 2016]



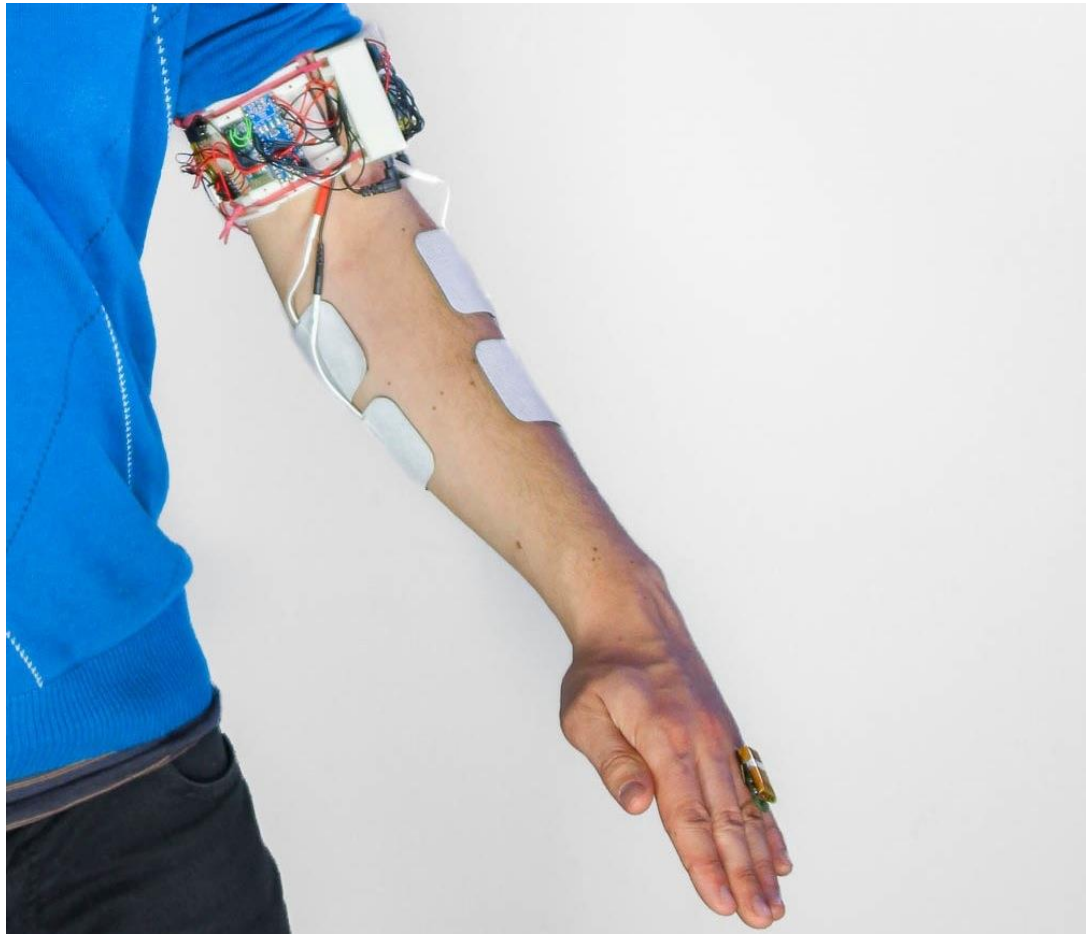
[Sandoval-Gonzalez et al., 2016]

perhaps we can borrow our body?

borrowing body as haptic actuator

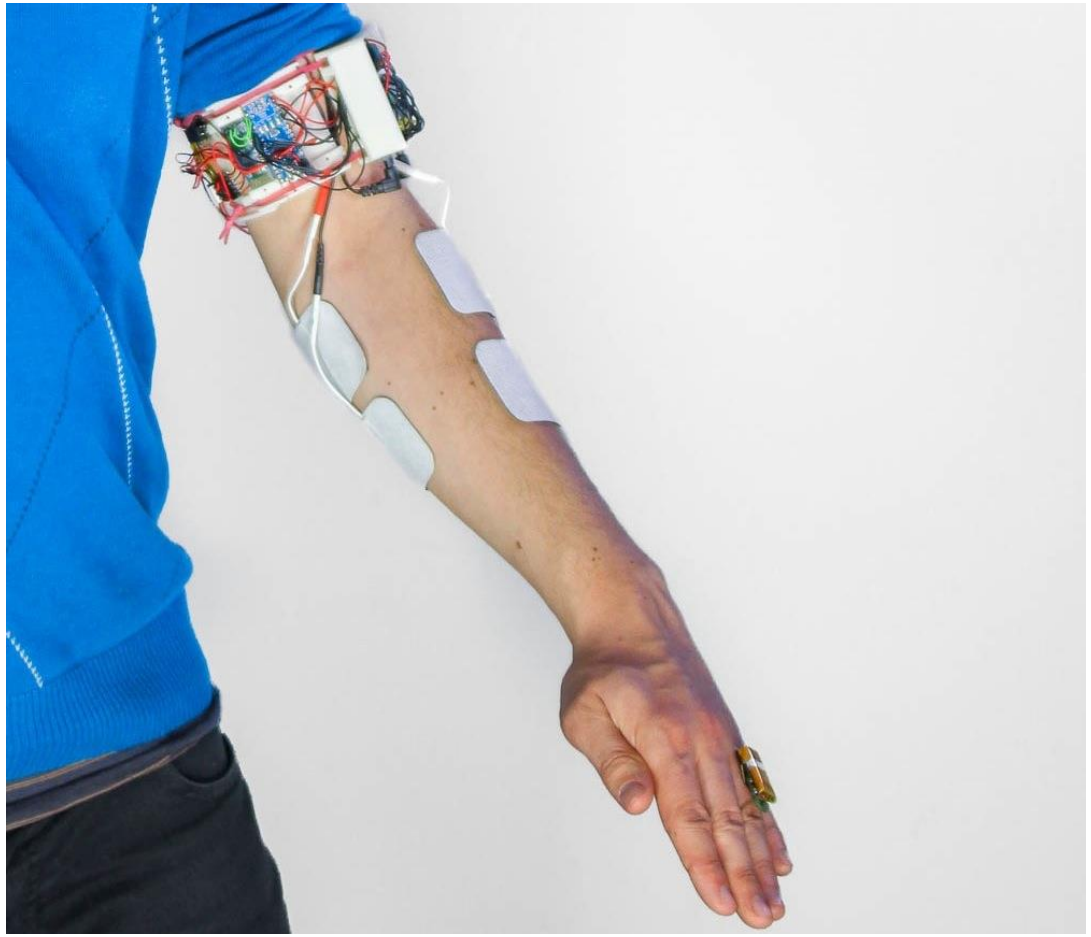


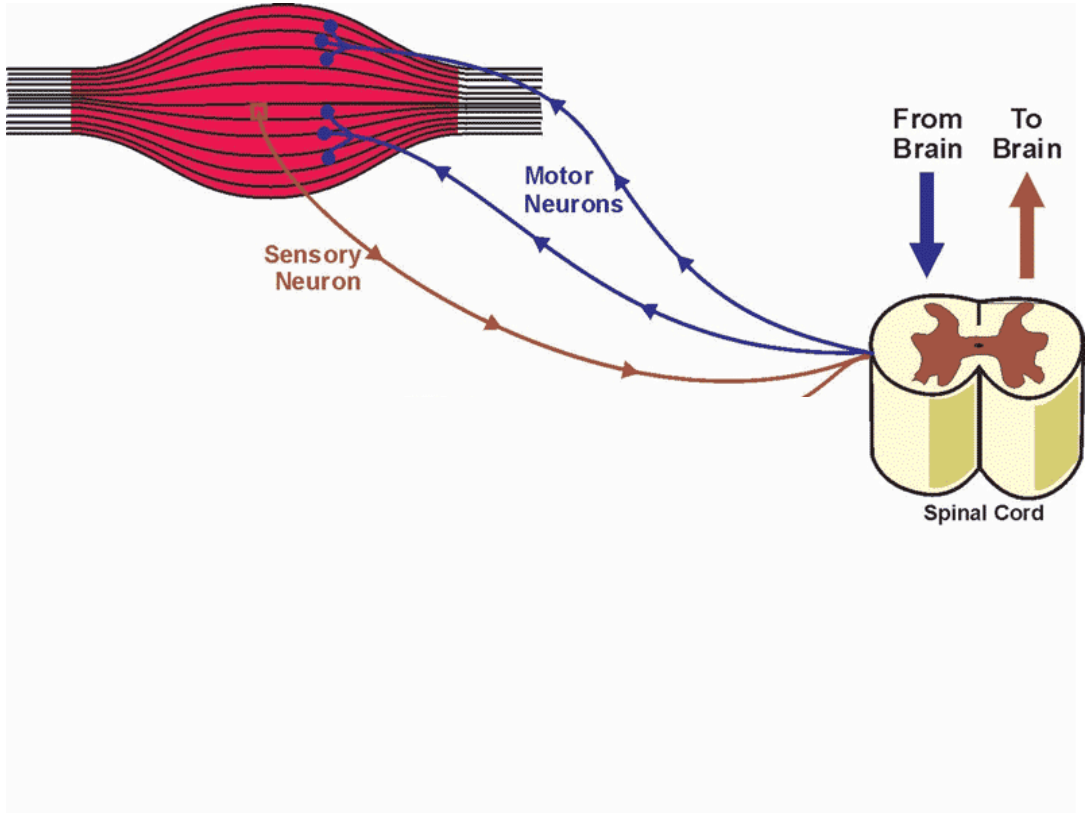
borrowing body as haptic actuator



Electrical Muscle Stimulation (EMS)

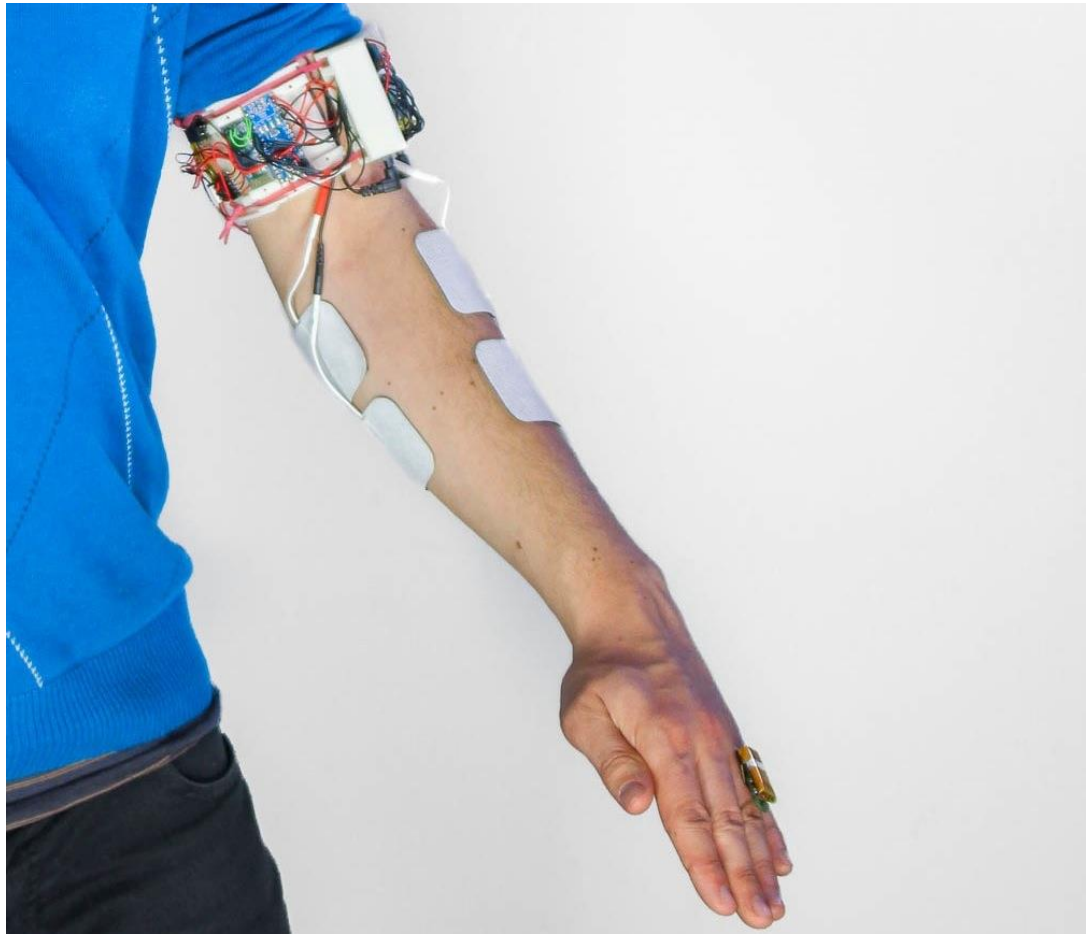
borrowing body as haptic actuator

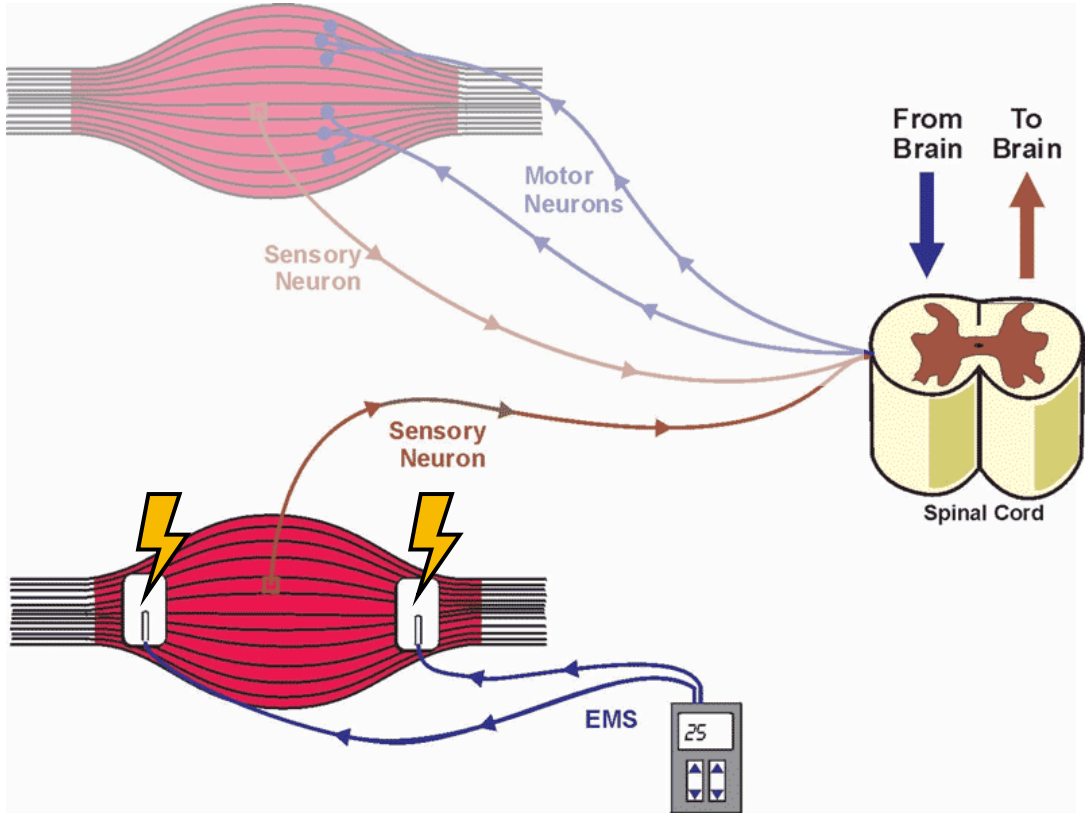




Electrical Muscle Stimulation (EMS)

borrowing body as haptic actuator

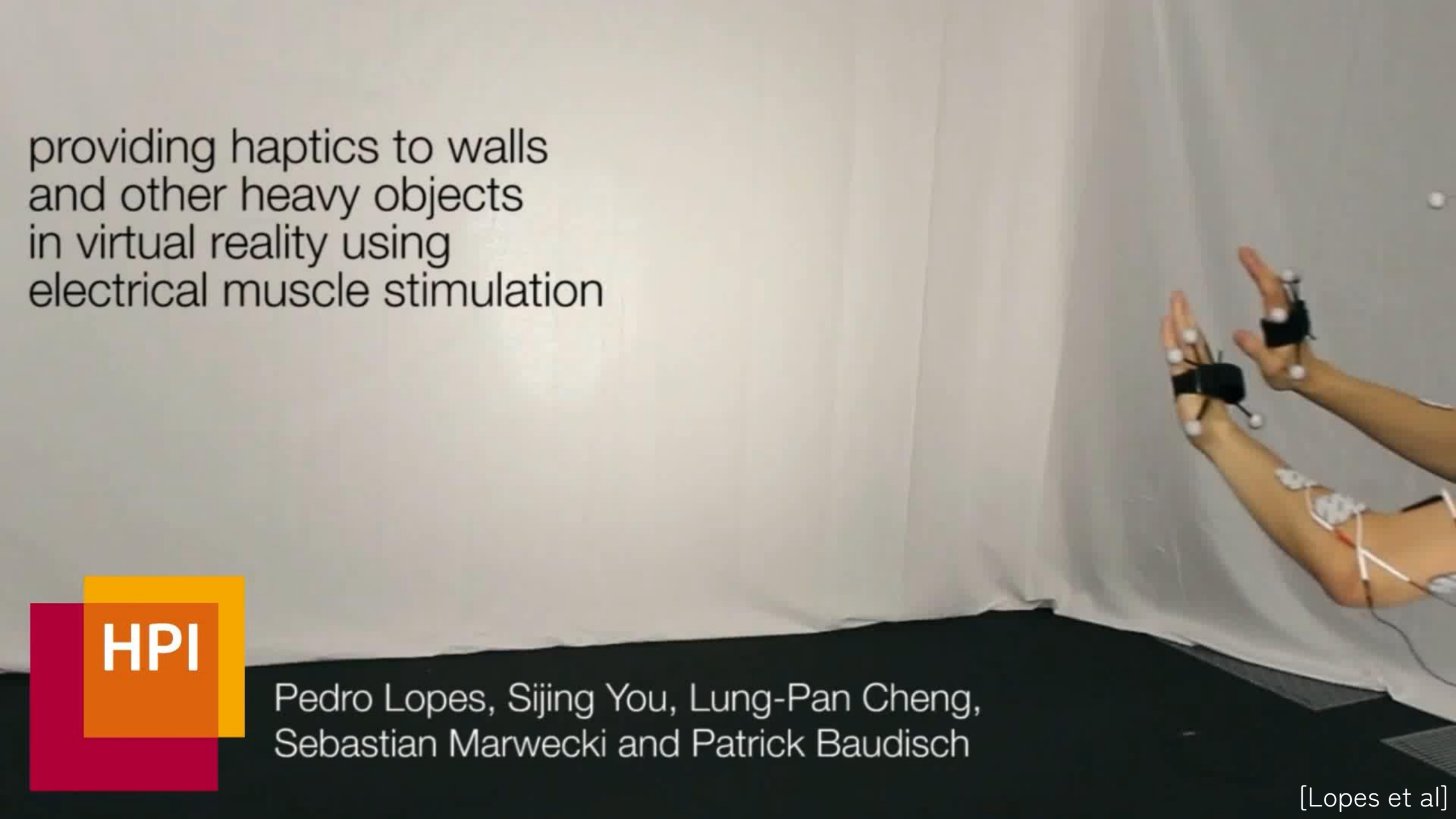




Electrical Muscle Stimulation (EMS)

[Lopes & Baudisch]

providing haptics to walls
and other heavy objects
in virtual reality using
electrical muscle stimulation



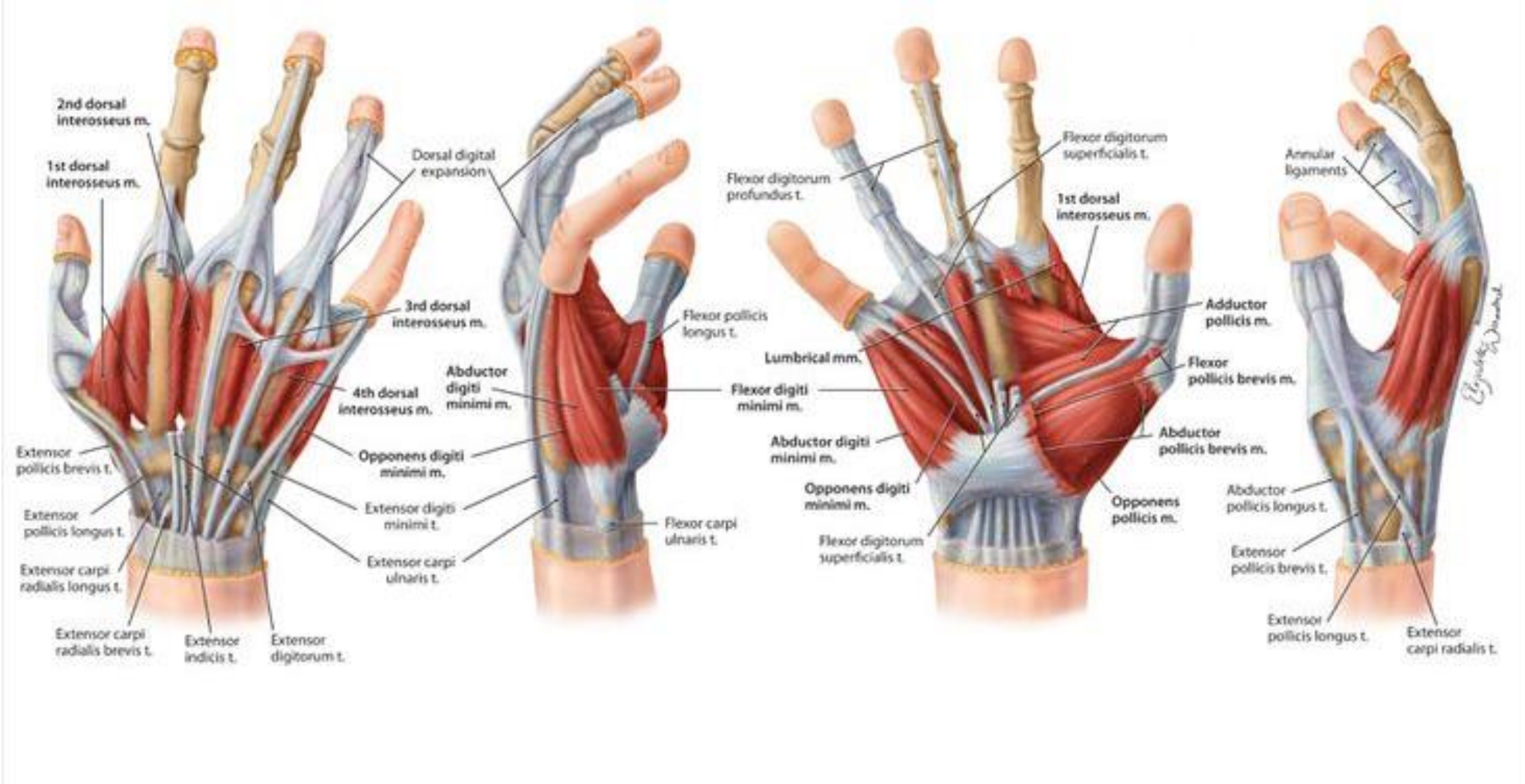
HPI

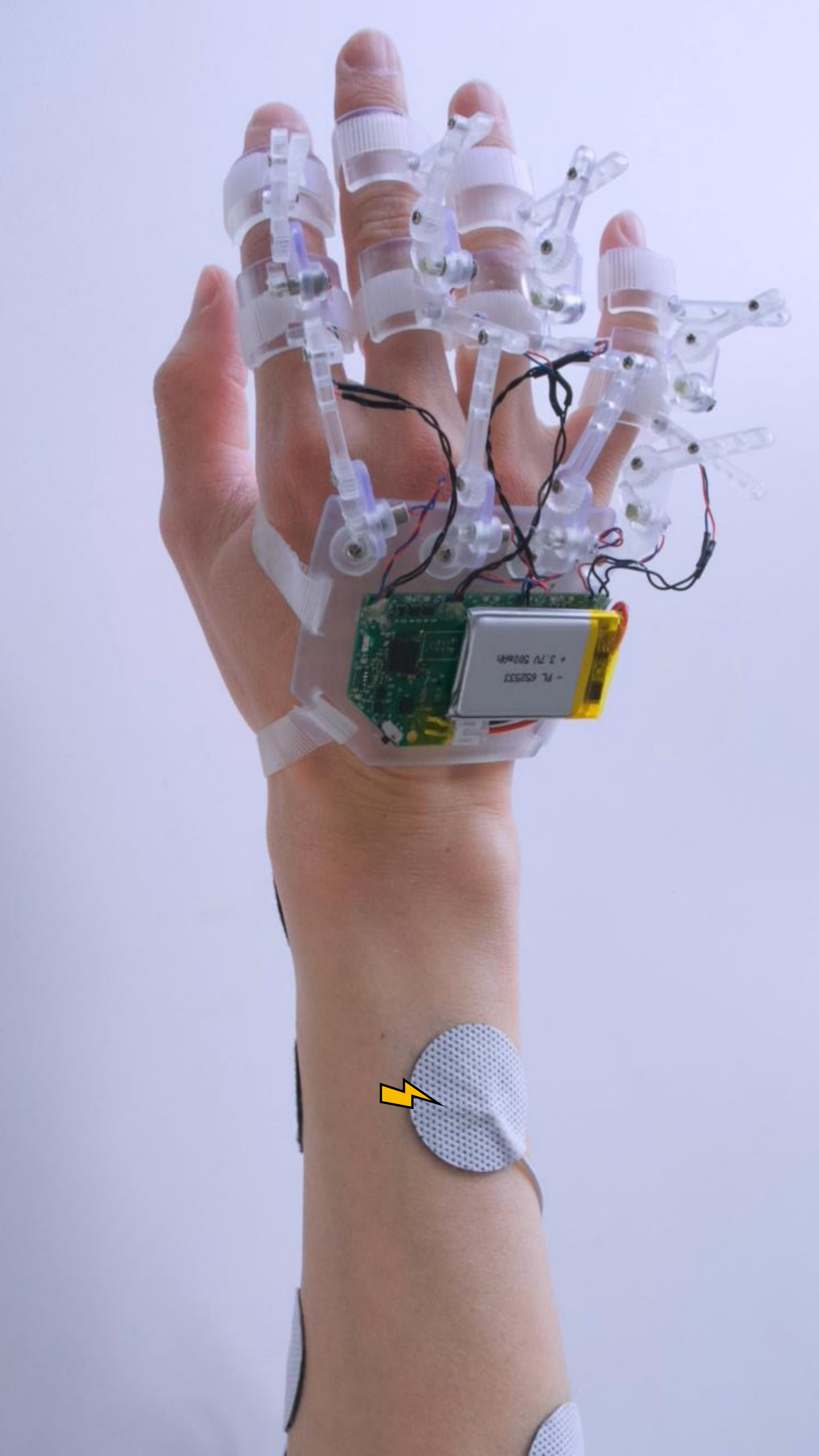
Pedro Lopes, Sijing You, Lung-Pan Cheng,
Sebastian Marwecki and Patrick Baudisch

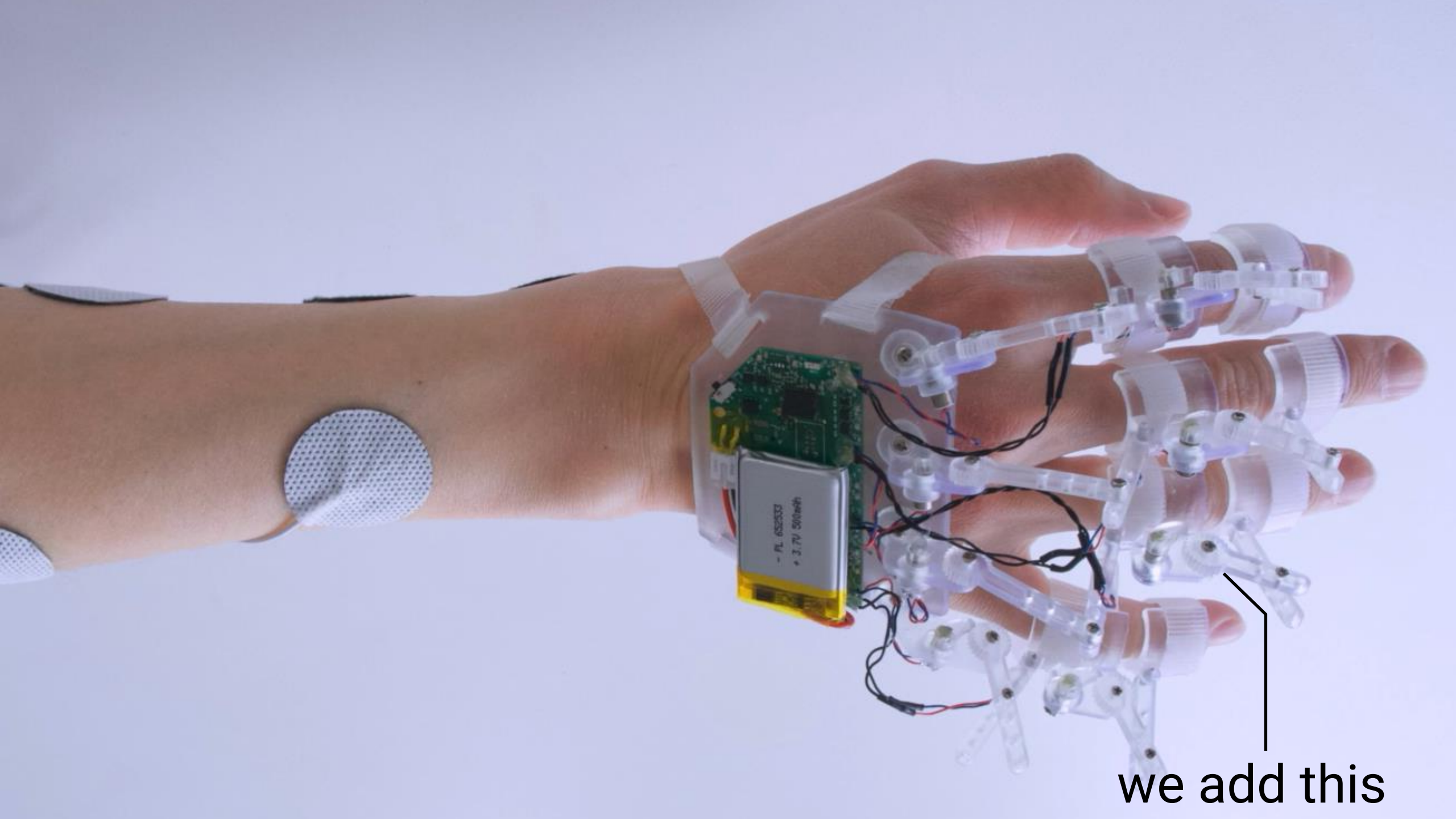
[Lopes et al]

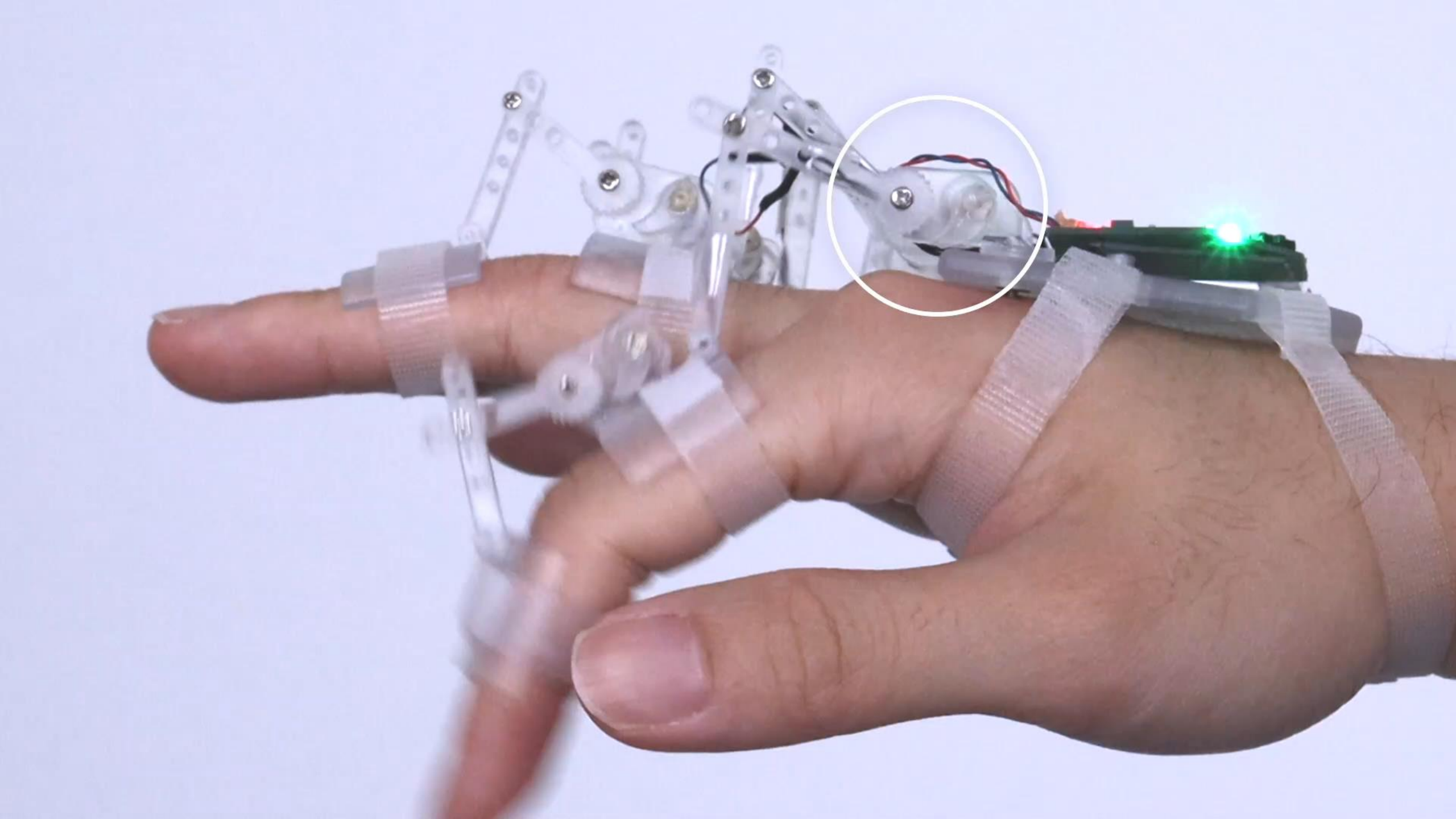
what’s the problem of electrical muscle stimulation?

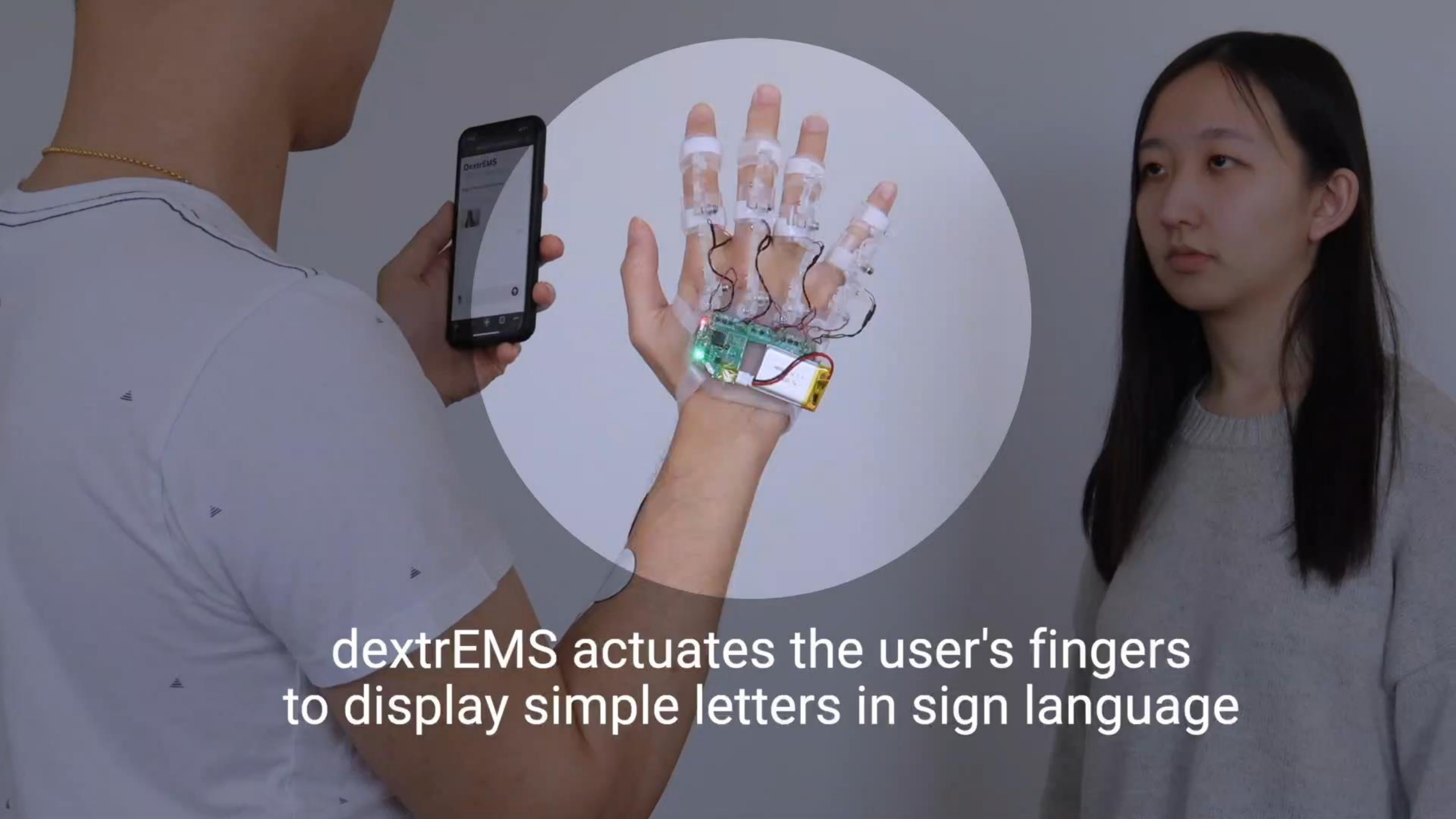
what's the problem of electrical muscle stimulation?





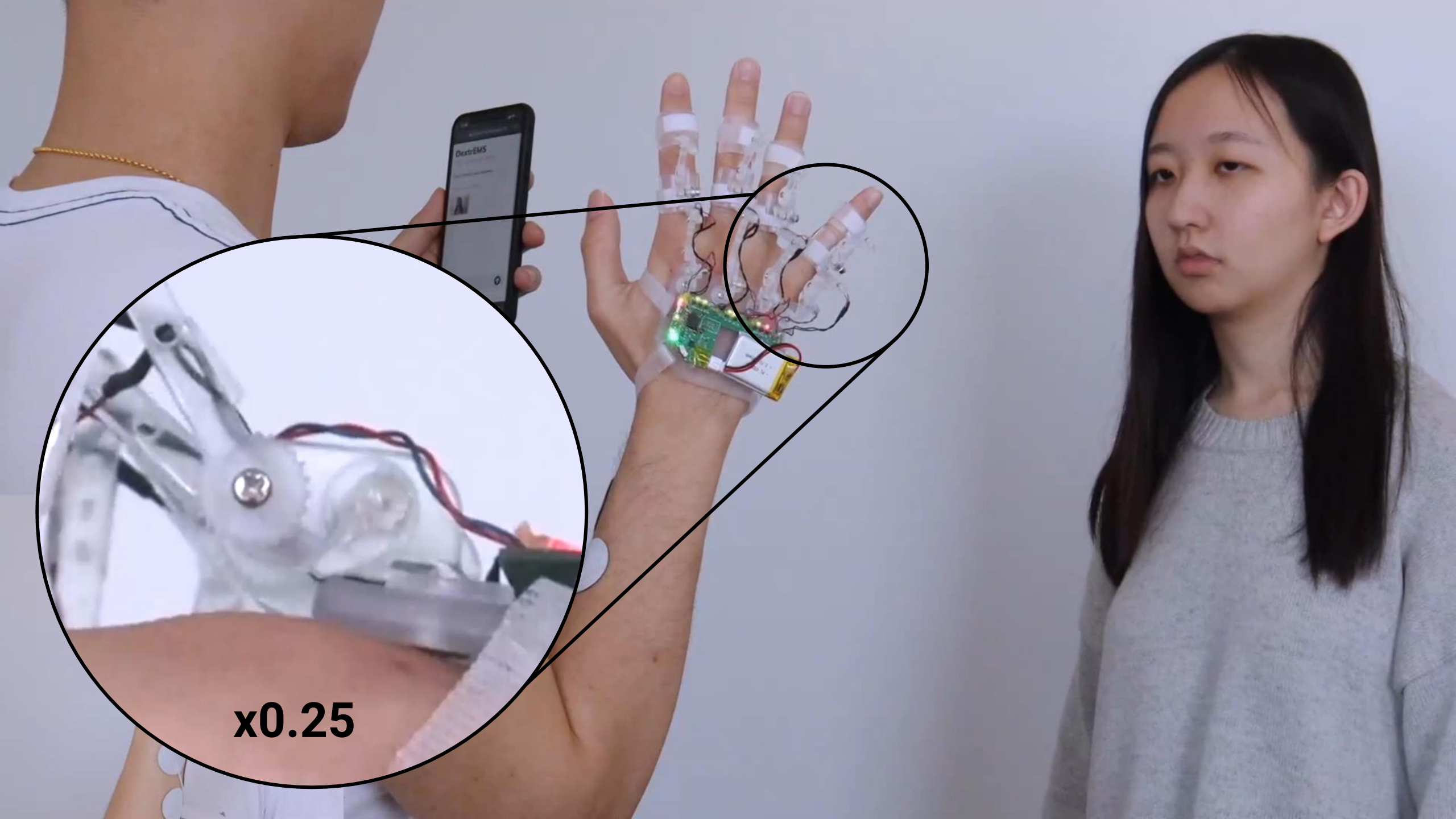

we add this

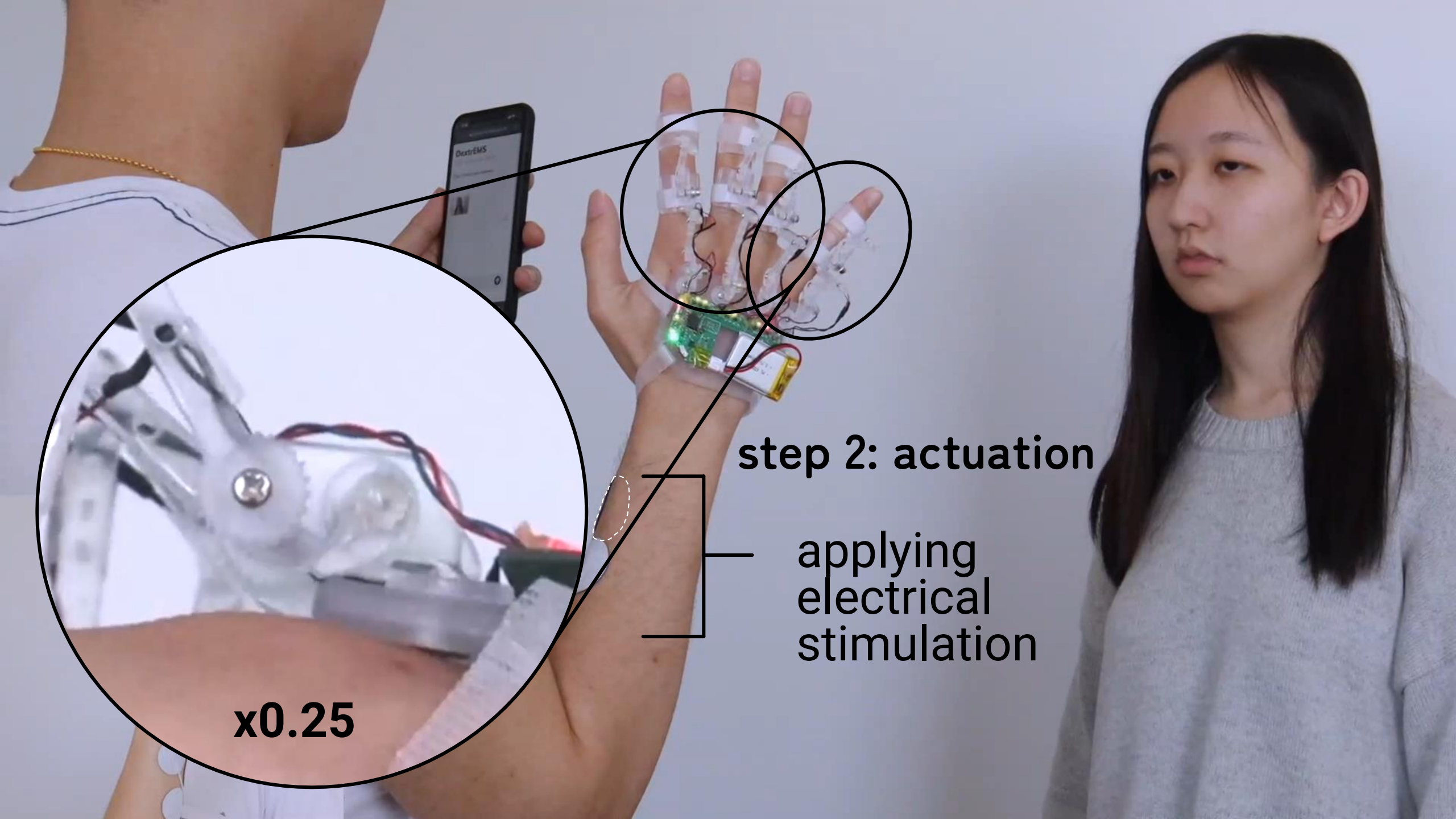



dextrEMS actuates the user's fingers
to display simple letters in sign language


DeanEMS

step 1: locking


DextrEMS
x0.25


step 2: actuation
applying electrical stimulation
x0.25

[UIST '21]

(more) benefits of multimodal HCI:

(more) benefits of multimodal HCI:

4. learnability
   make skills easier to learn

5. accessibility
   make computers work for more people

multimodal cues helps learning

multimodal cues helps learning

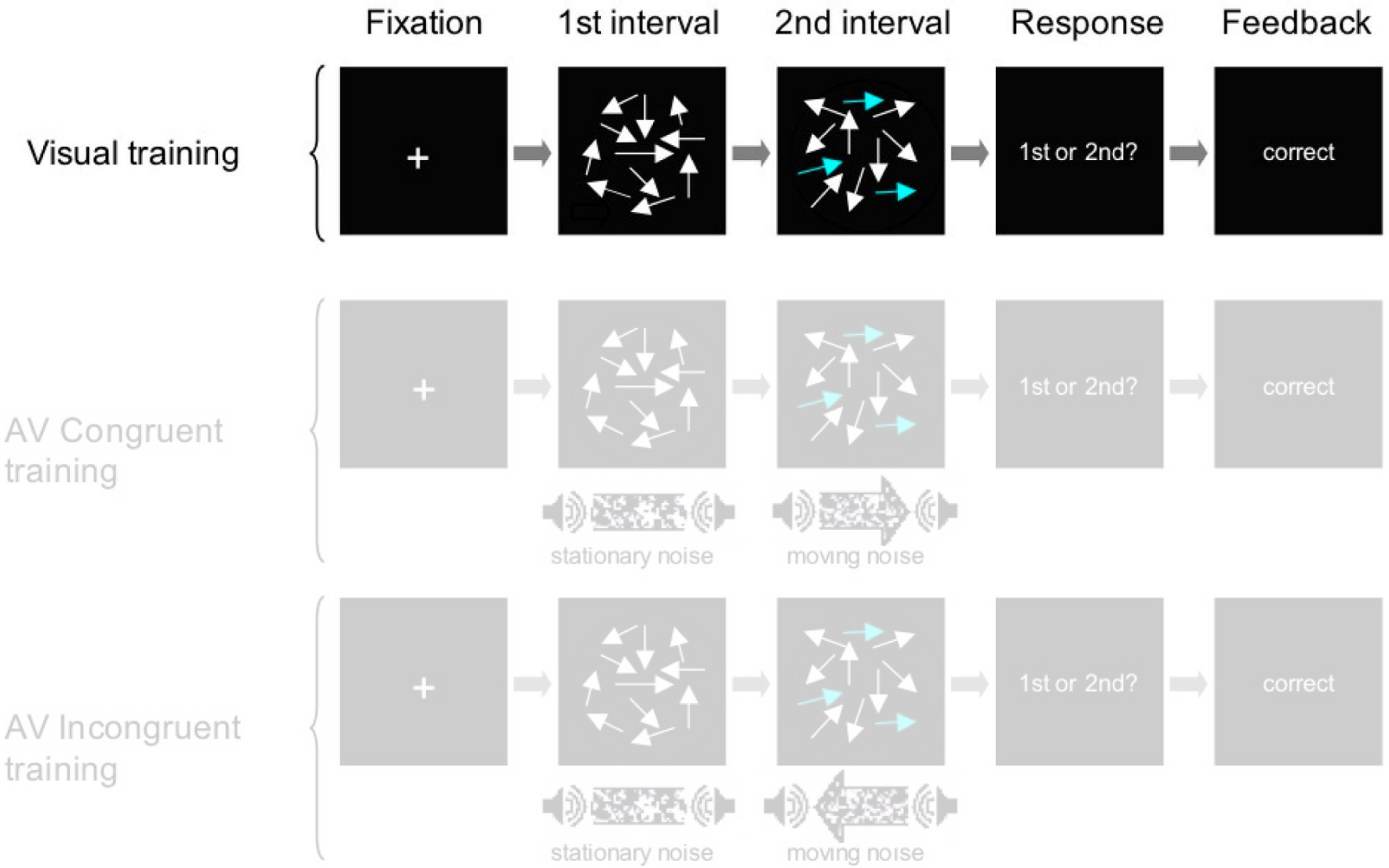
周大慶

multimodal cues helps learning

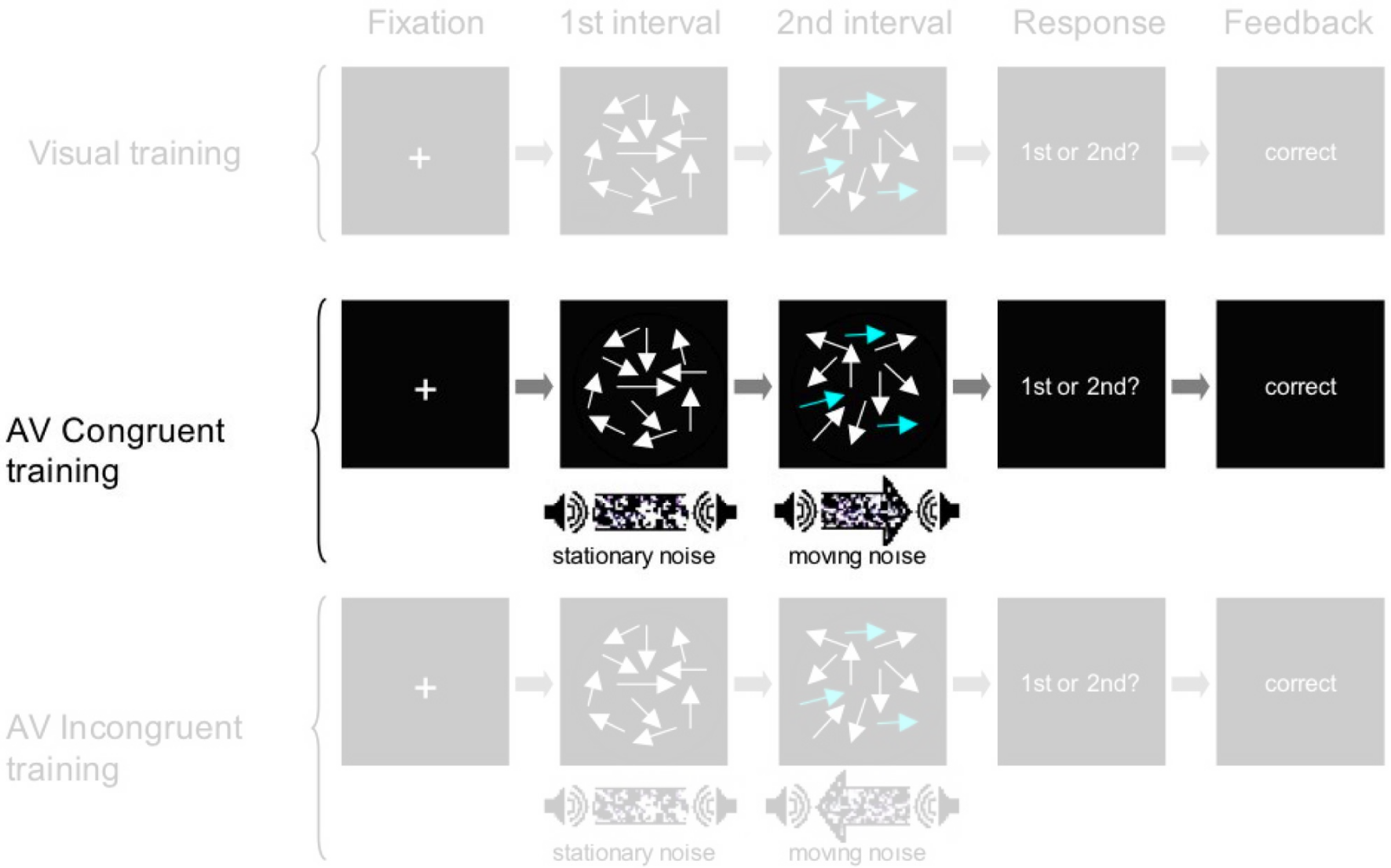
*even when multimodal cues are absent!

周大慶

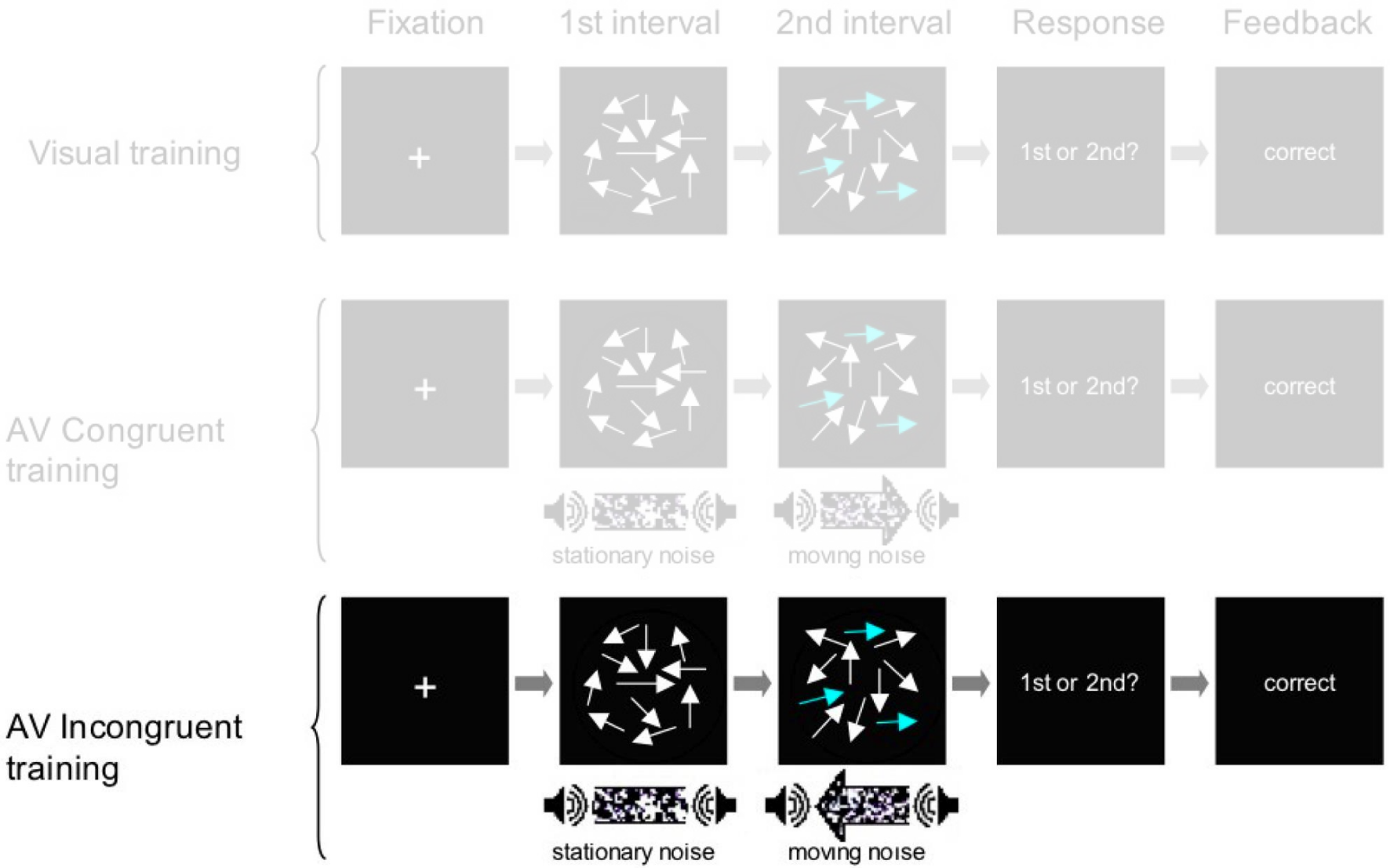
multimodal cues helps learning



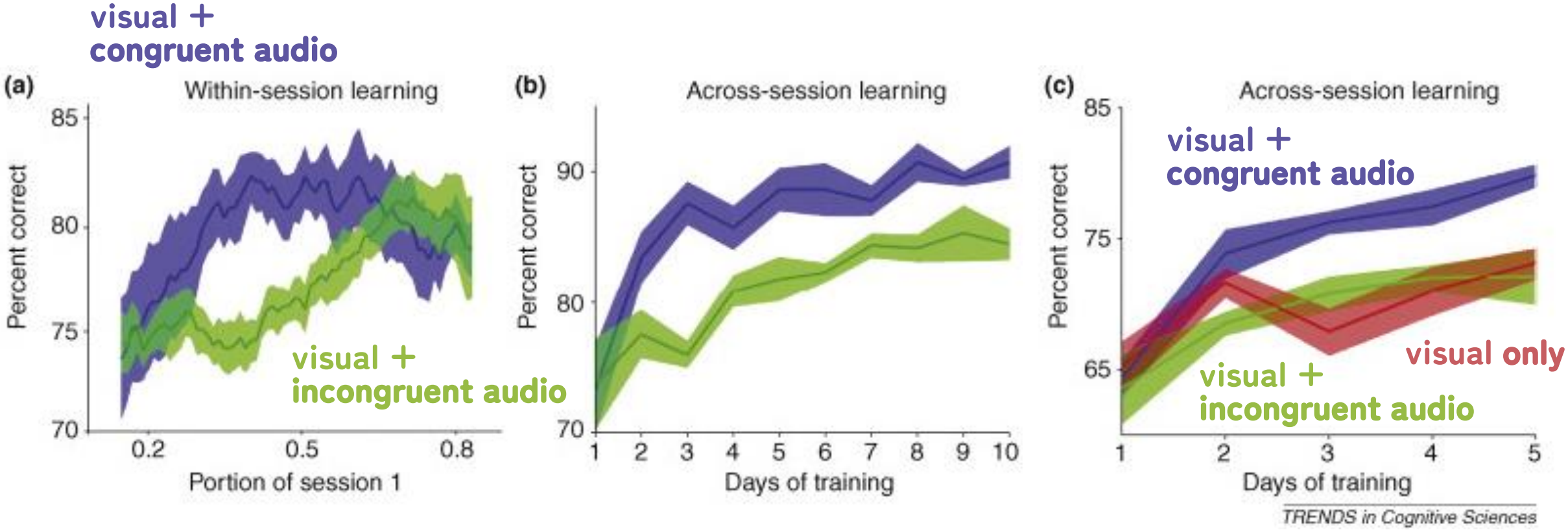
multimodal cues helps learning



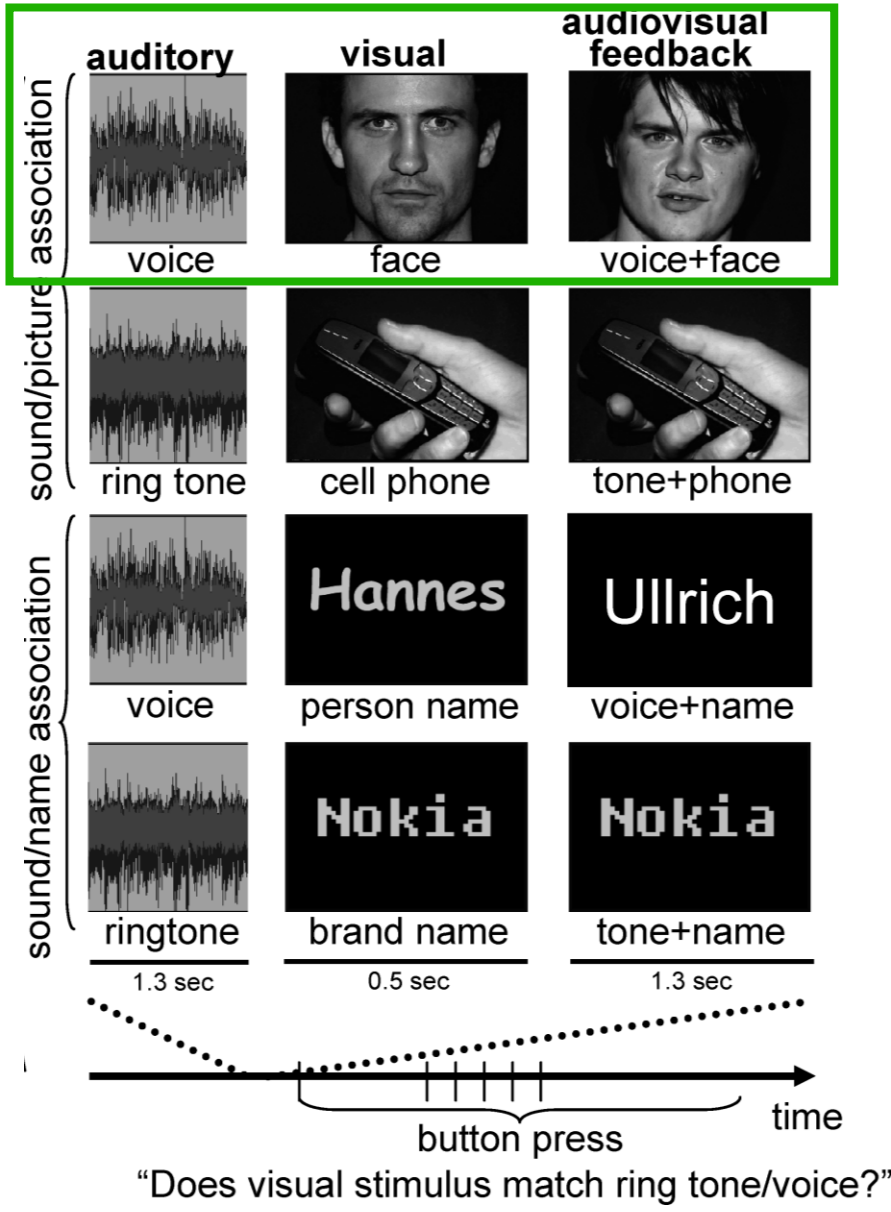
multimodal cues helps learning



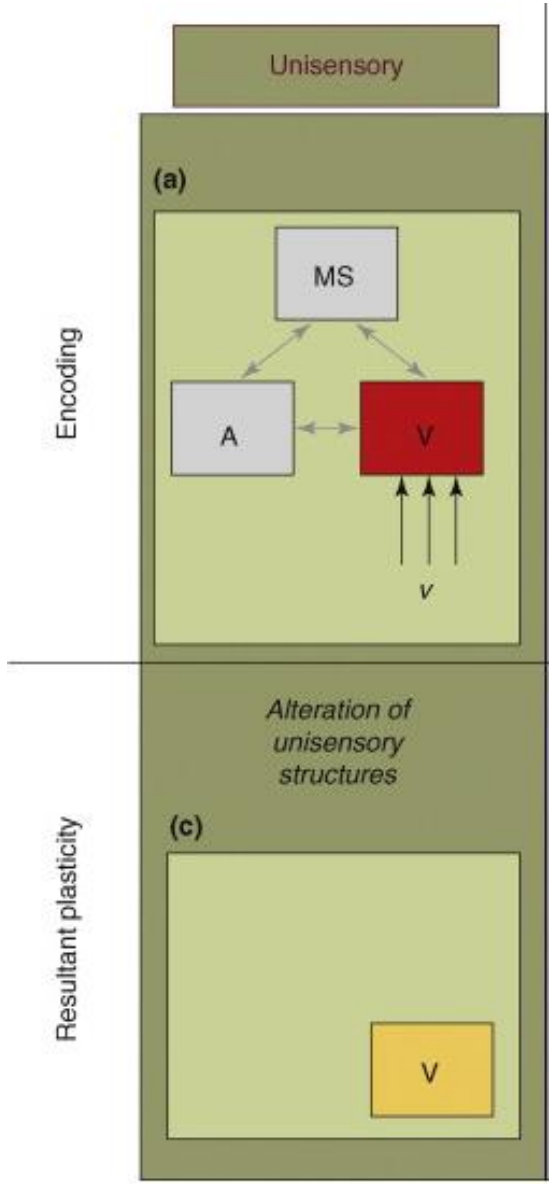
visual + **congruent audio** helps visual tasks



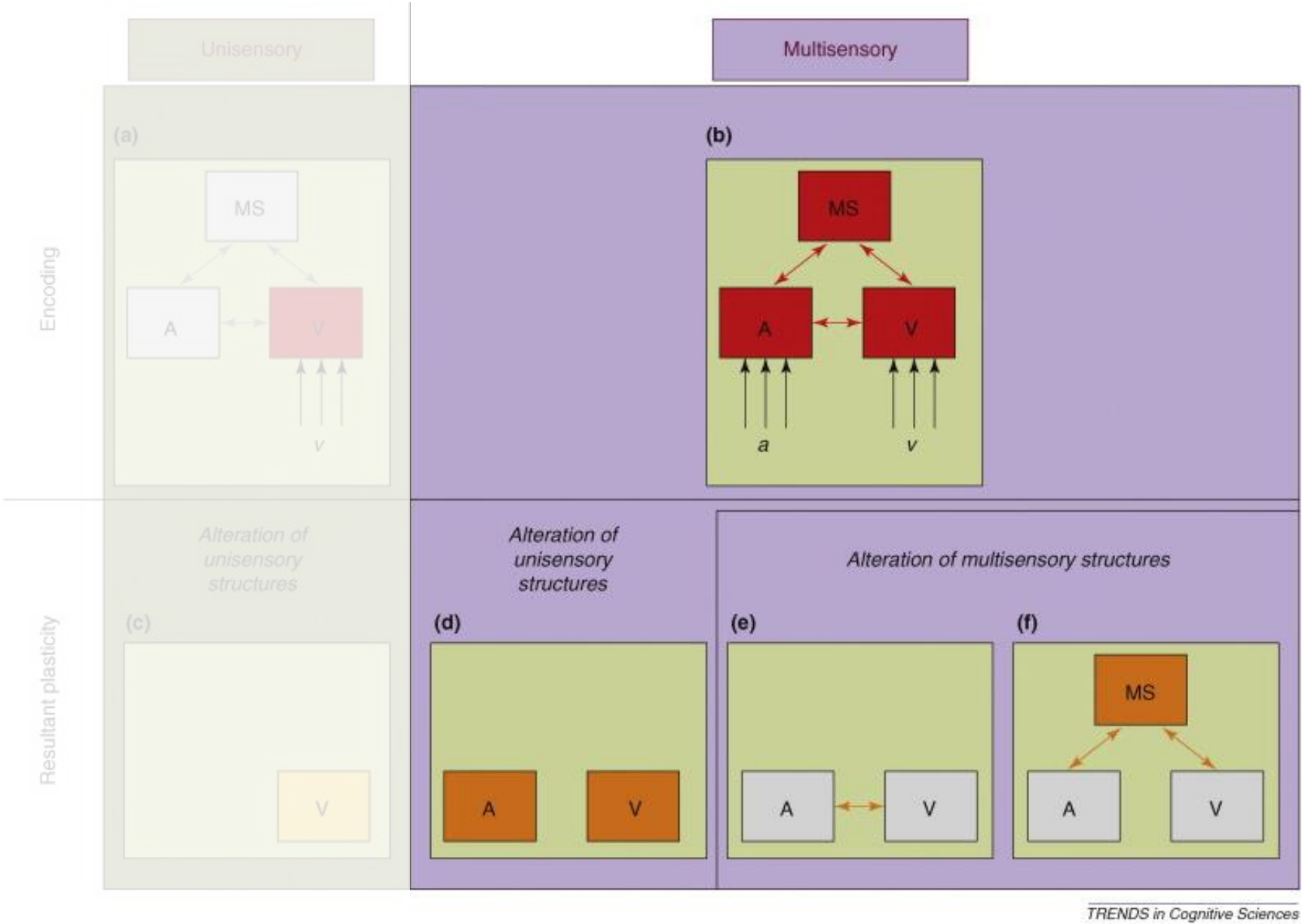
audio + **congruent visual** helps audio tasks



multimodal encoding model



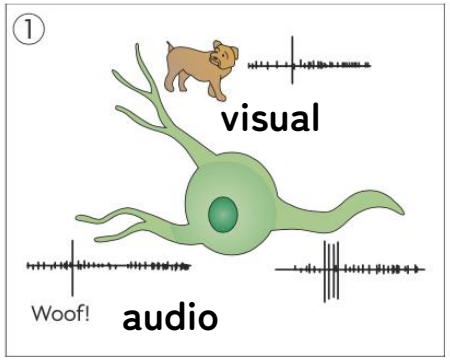
multimodal encoding model



multisensory integration (at neuron cell level)

multisensory integration (at neuron cell level)

multisensory integration (at neuron cell level)



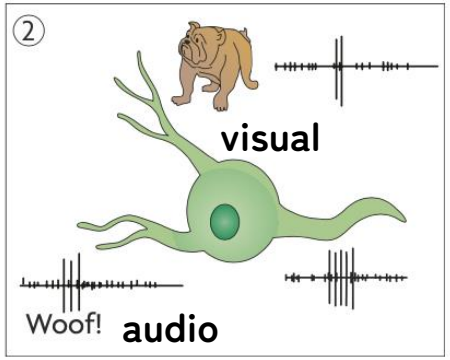
superadditive

> visual + audio

multisensory integration (at neuron cell level)

superadditive

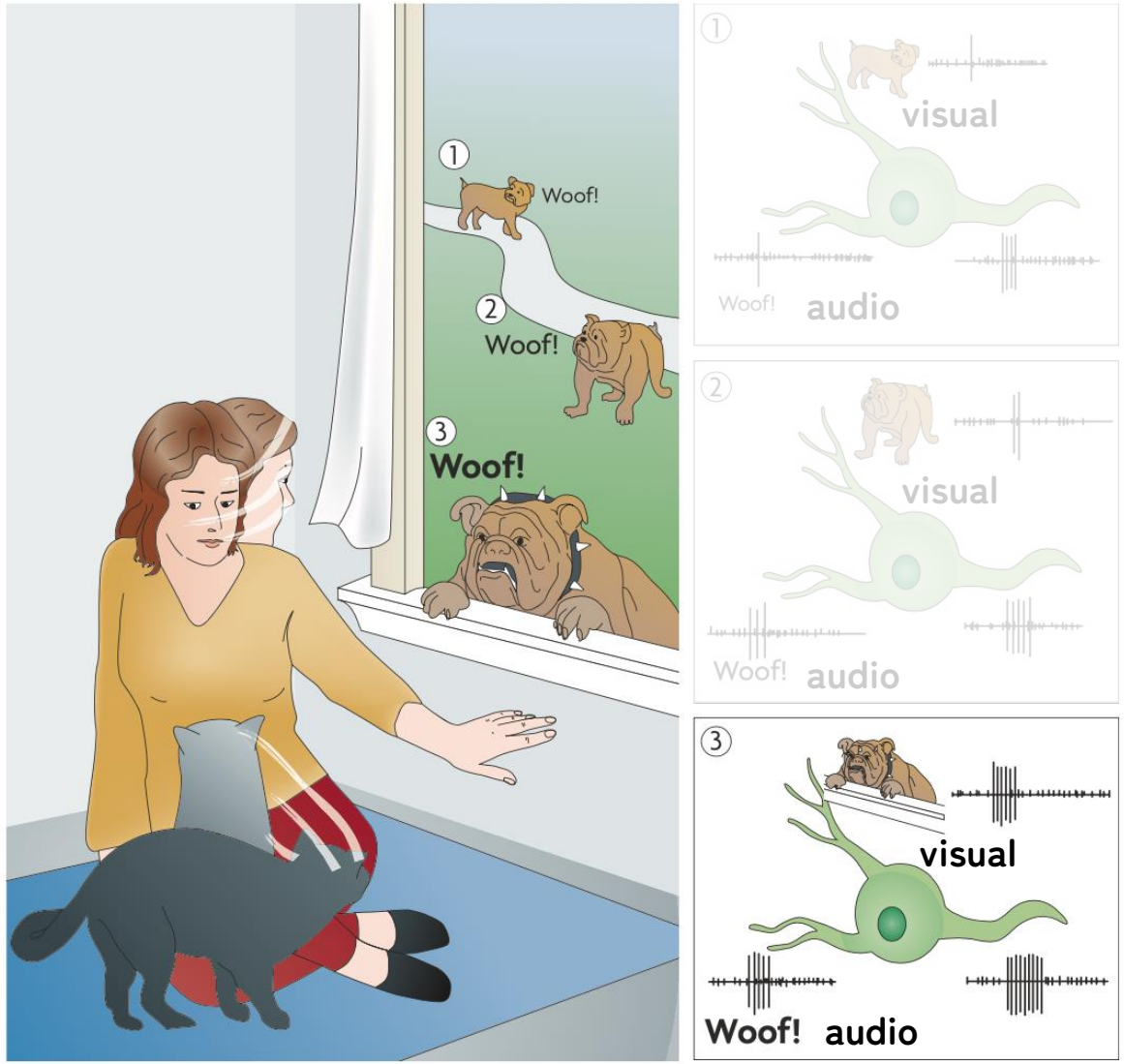
> visual + audio



additive

= visual + audio

multisensory integration (at neuron cell level)



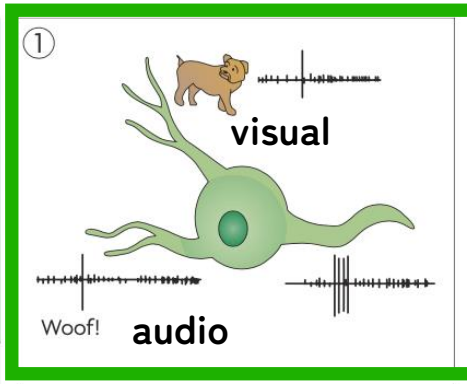
superadditive
> visual + audio

additive
= visual + audio

subadditive
< visual + audio

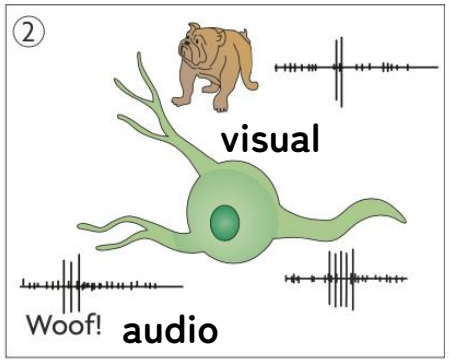
multisensory integration (at neuronal level)

when signals are weak (uncertain)



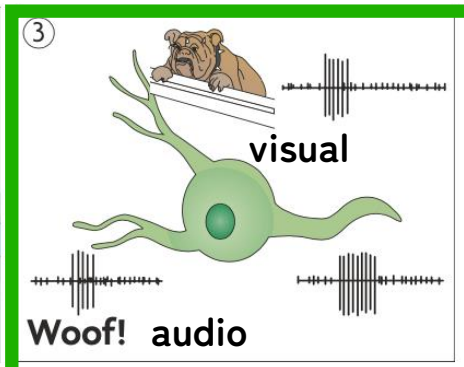
superadditive

> visual + audio



additive

= visual + audio



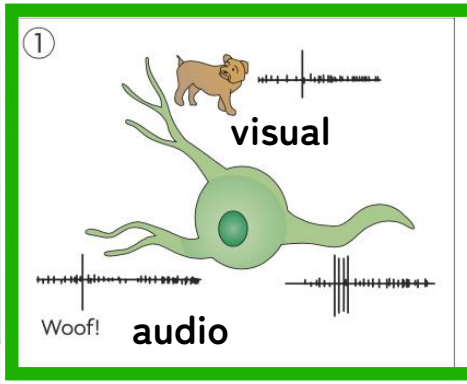
subadditive

< visual + audio

multisensory integration (at neuronal level)

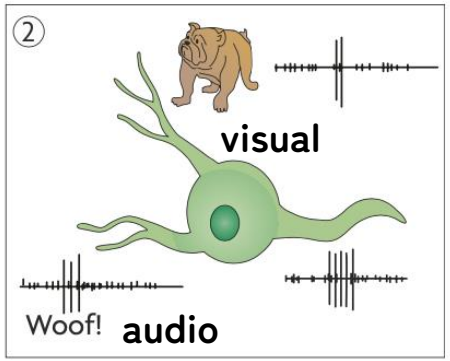
when signals are weak (uncertain)

why?



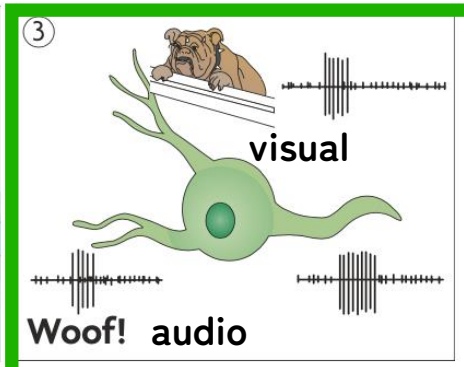
superadditive

> visual + audio



additive

= visual + audio



subadditive

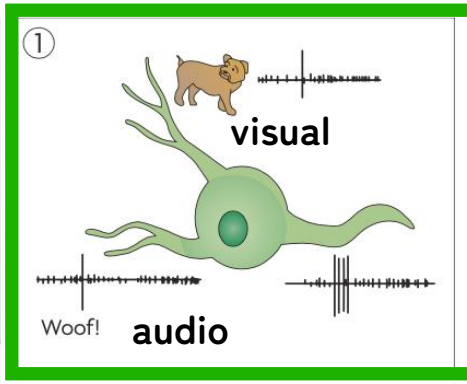
< visual + audio

multisensory integration (at neuronal level)

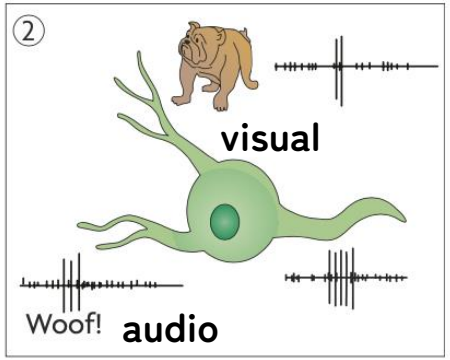
when signals are weak (uncertain)

why?

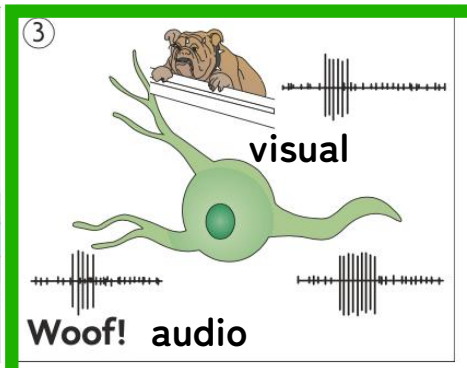
can **shorten the interval** between sensory encoding and motor-command formation & create a **unitary** perceptual experience



superadditive
> visual + audio



additive
= visual + audio

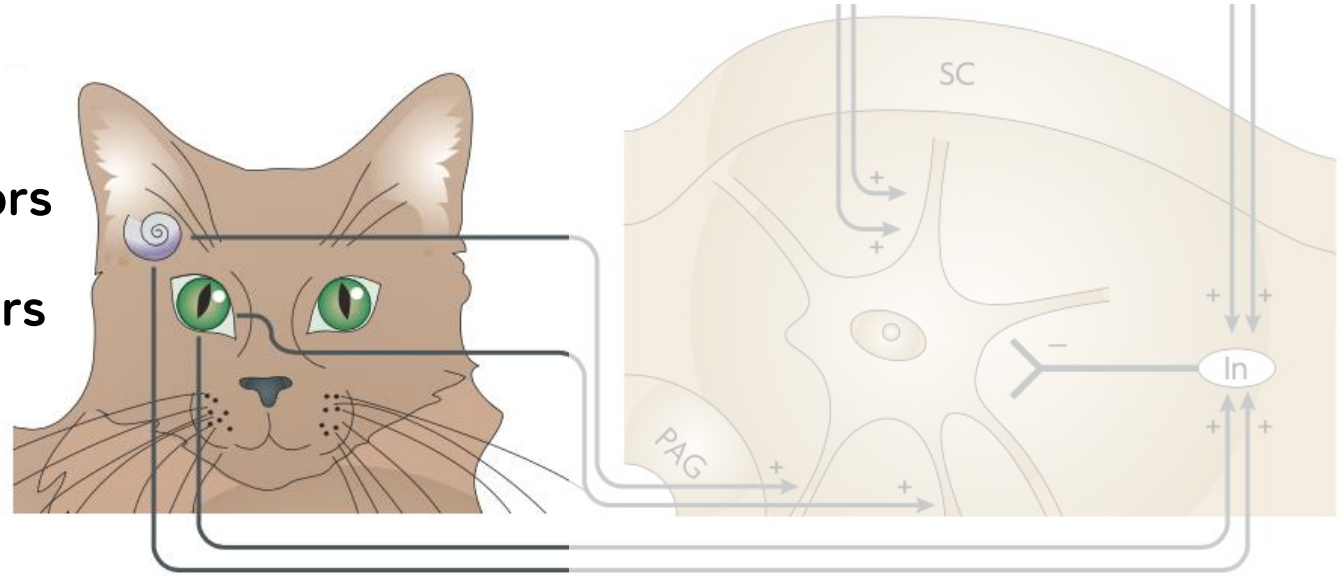


subadditive
< visual + audio

where integration happens?
(in cats)

auditory receptors

visual receptors



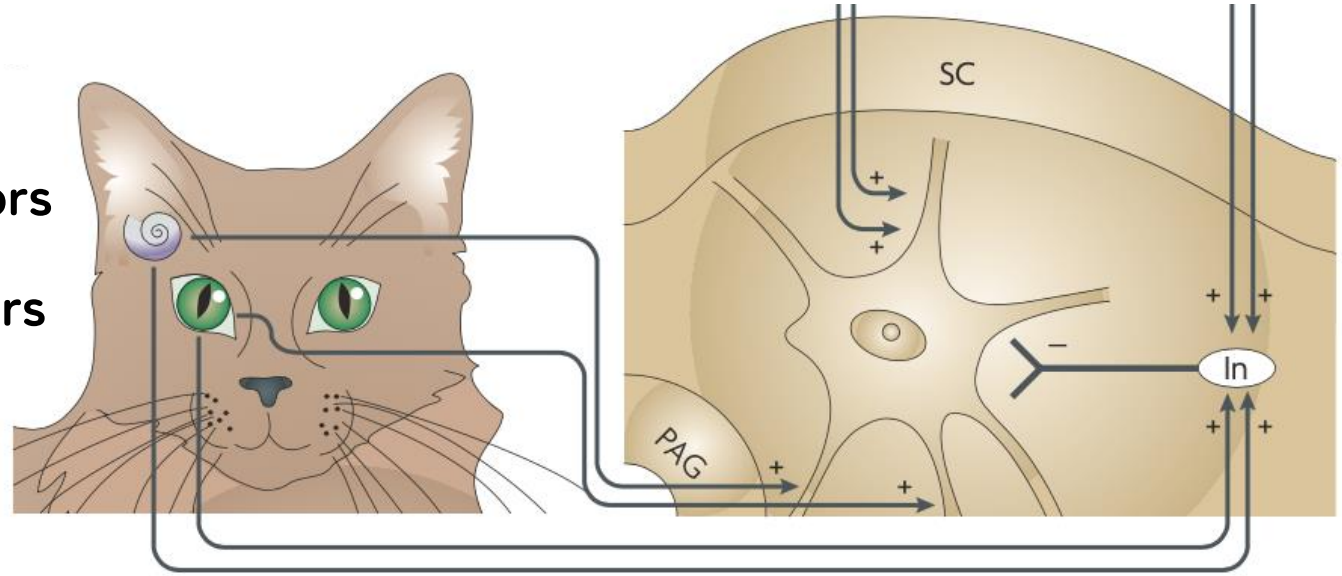
superior colliculus

gaze shift,
integration, etc

where integration happens? (in cats)

auditory receptors

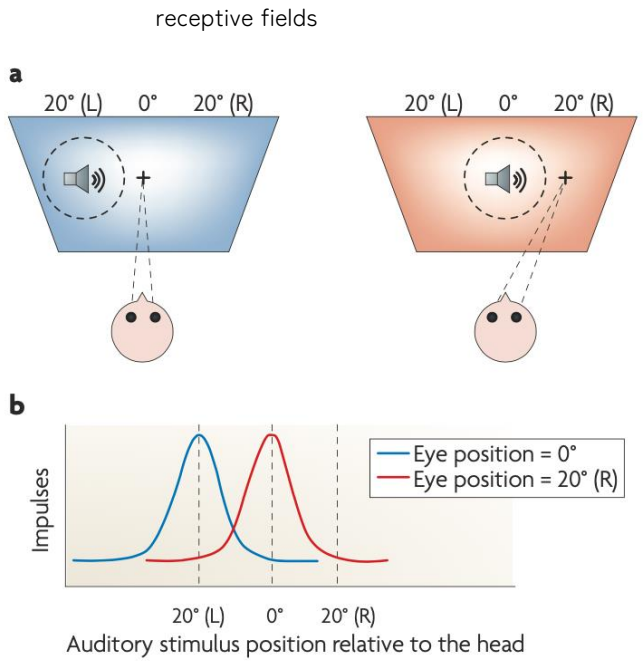
visual receptors



superior colliculus

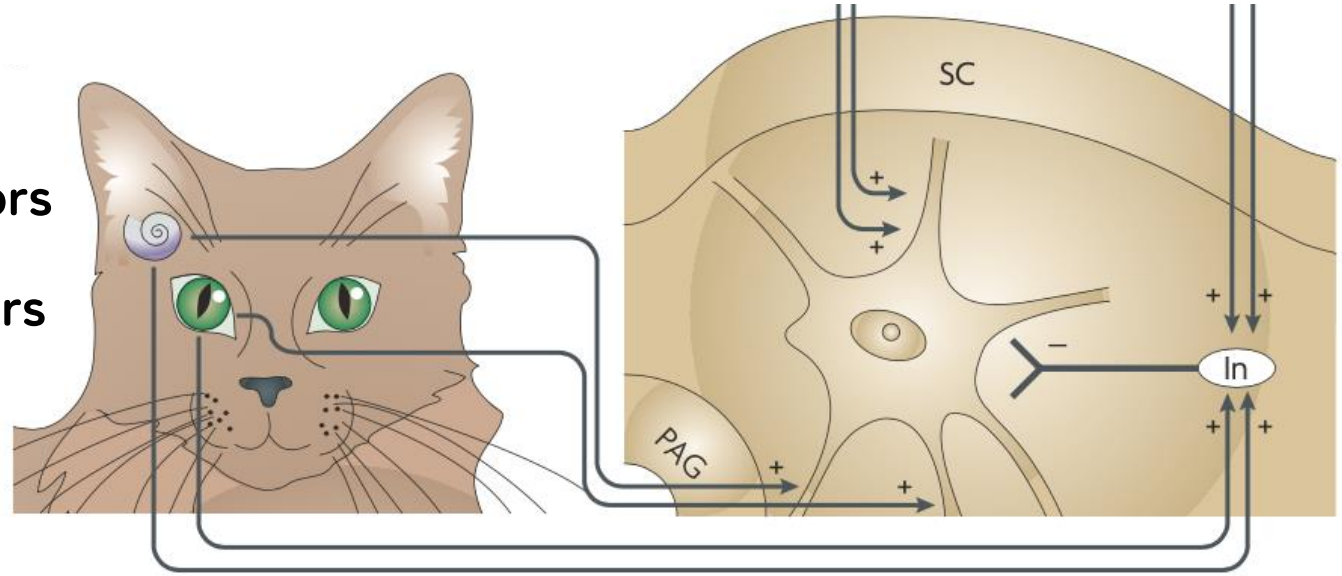
gaze shift, integration, etc

where integration happens? (in cats)



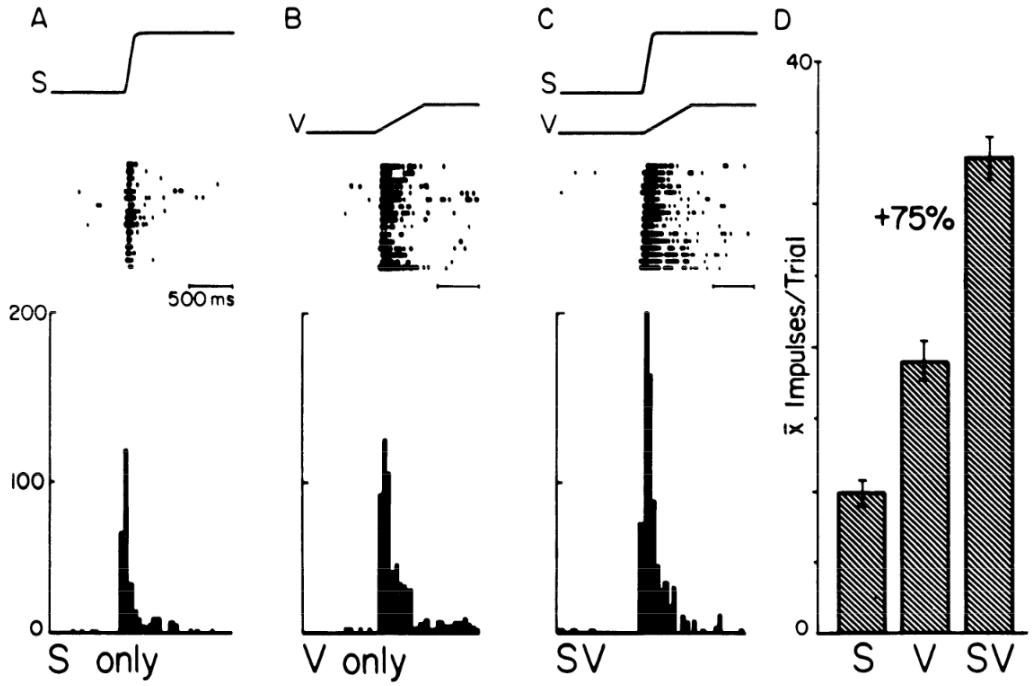
auditory receptors

visual receptors



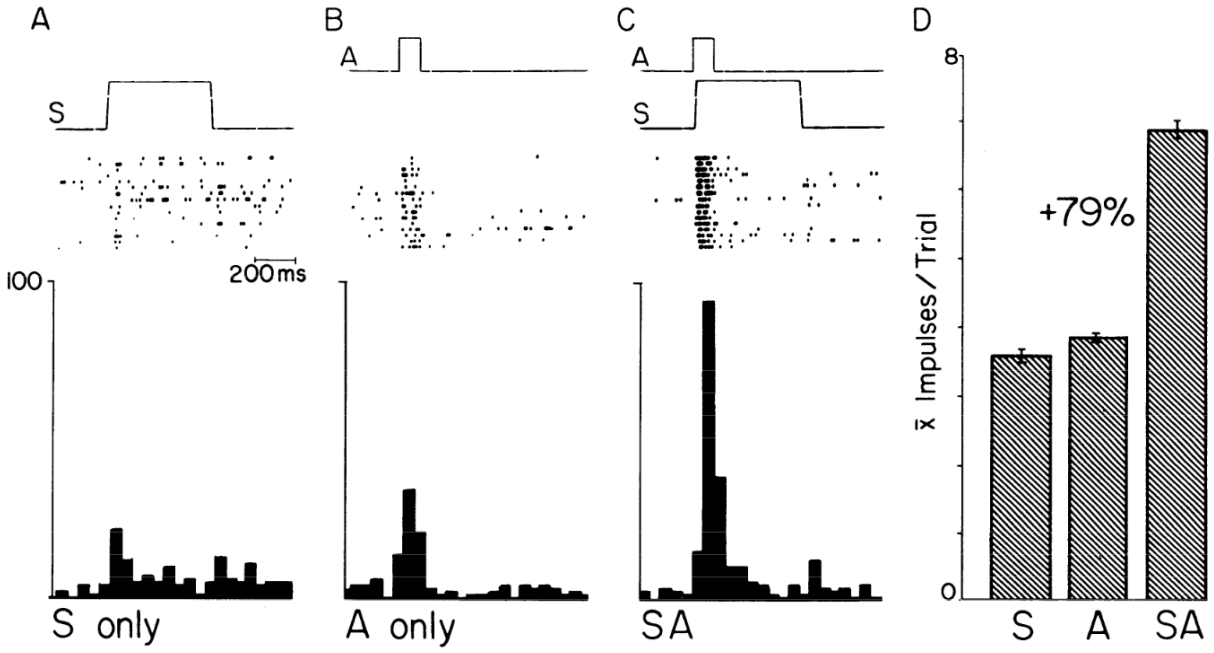
superior colliculus

gaze shift, integration, etc



somato-sensory only (vibration)

visual only



somato-sensory only

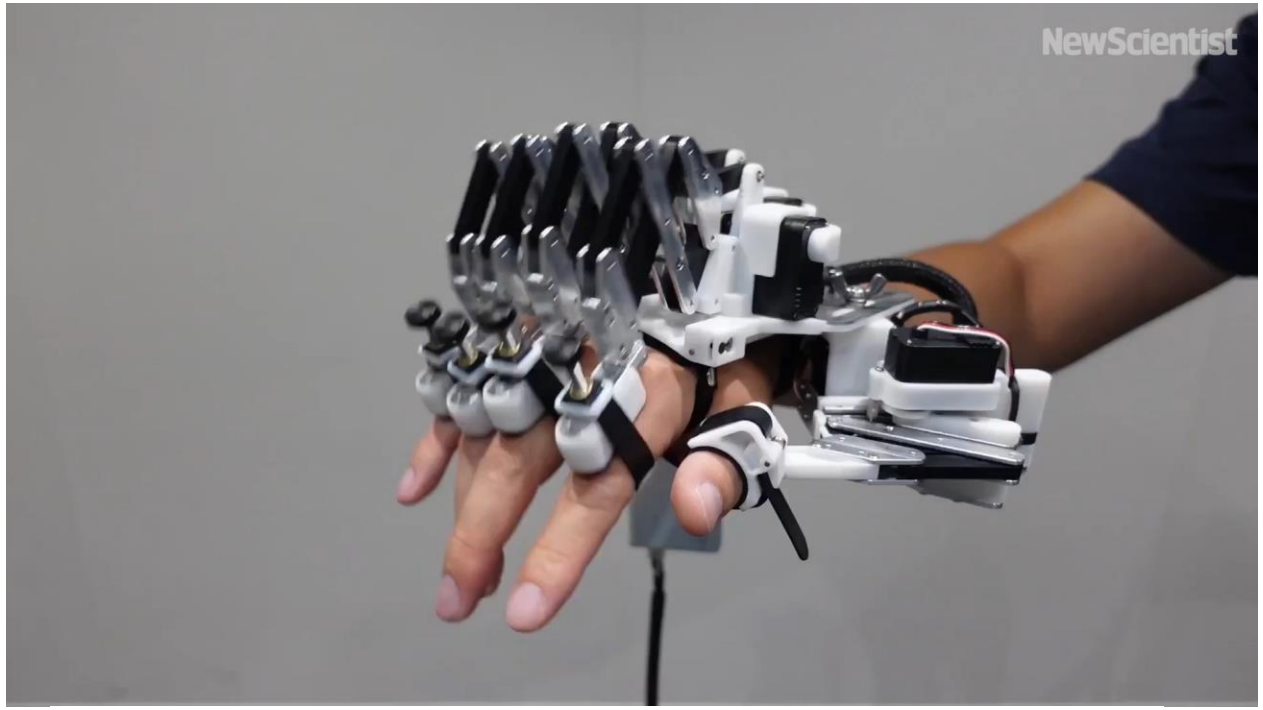
audio only

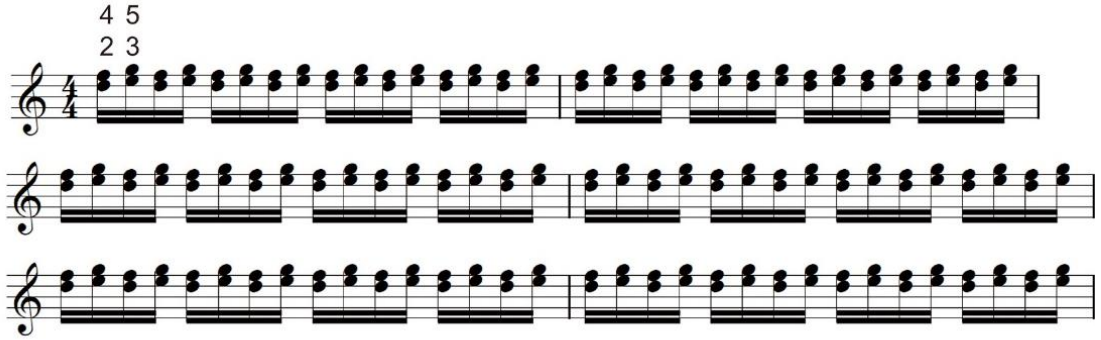
[Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration, 1986]

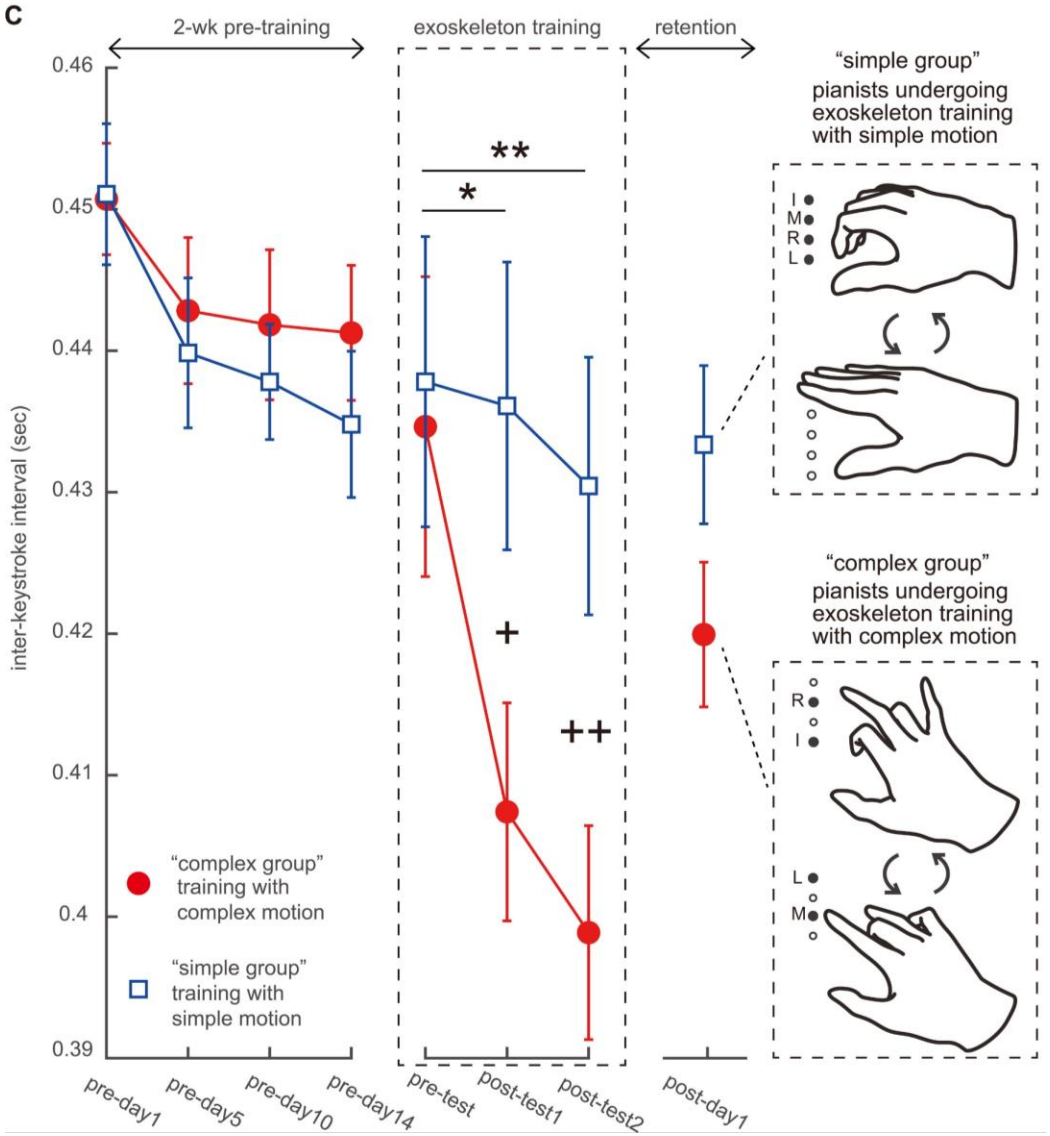
haptics helps **motor learning**


this user **learns** a new melody
with electrical muscle stimulation (EMS)
AKAI
PROFESSIONAL
LPK25
LAPTOP PERFORMANCE KEYBOARD

go beyond traditional training







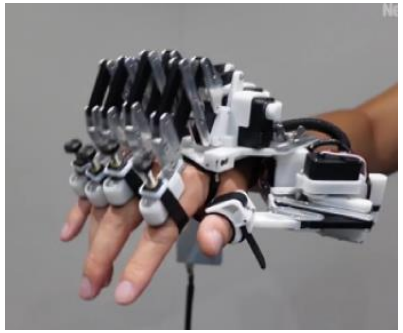
multimodal learning

multimodal learning

... merging other modalities?

benefits of multimodal HCI:

4. learnability
   make skills easier to learn

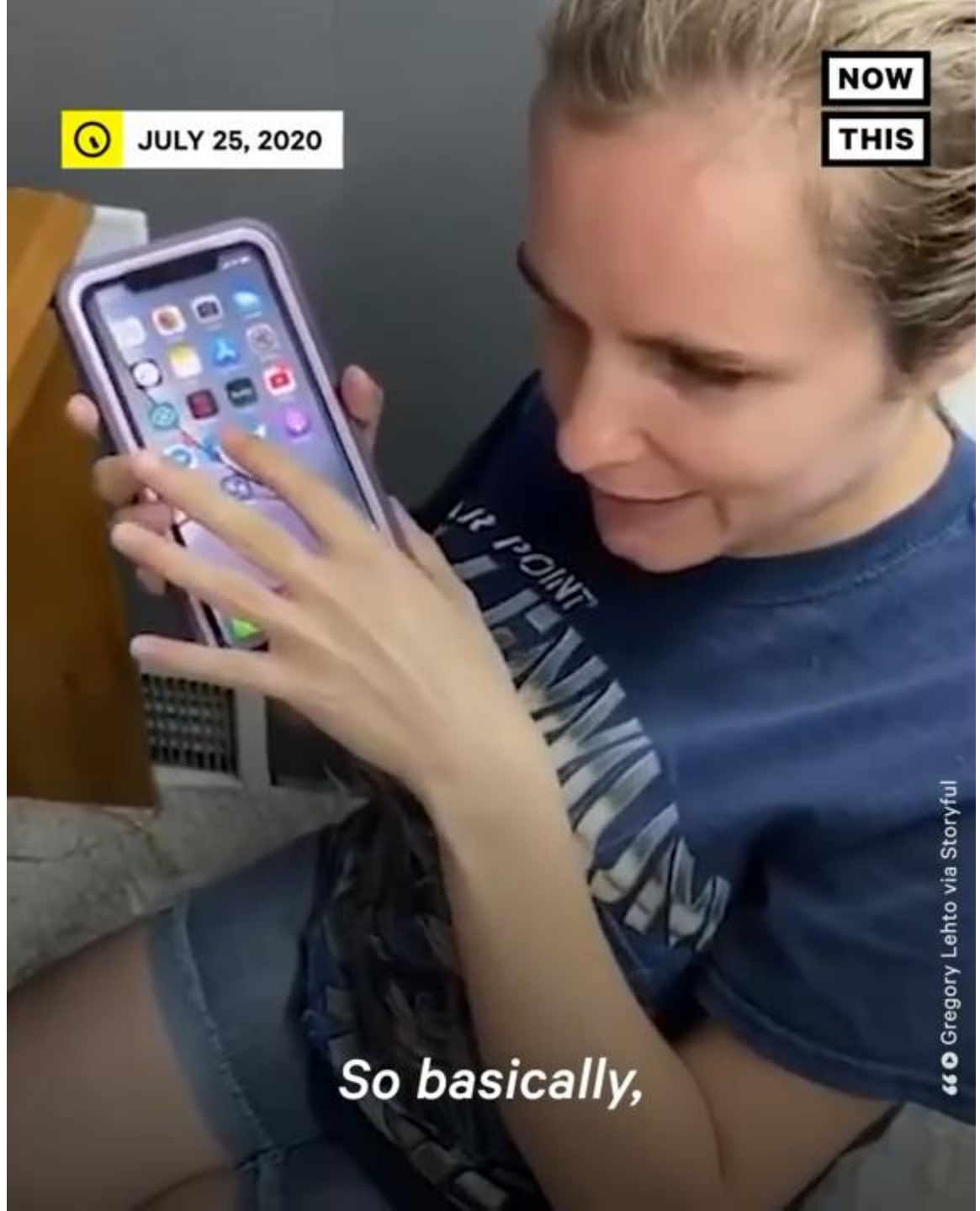


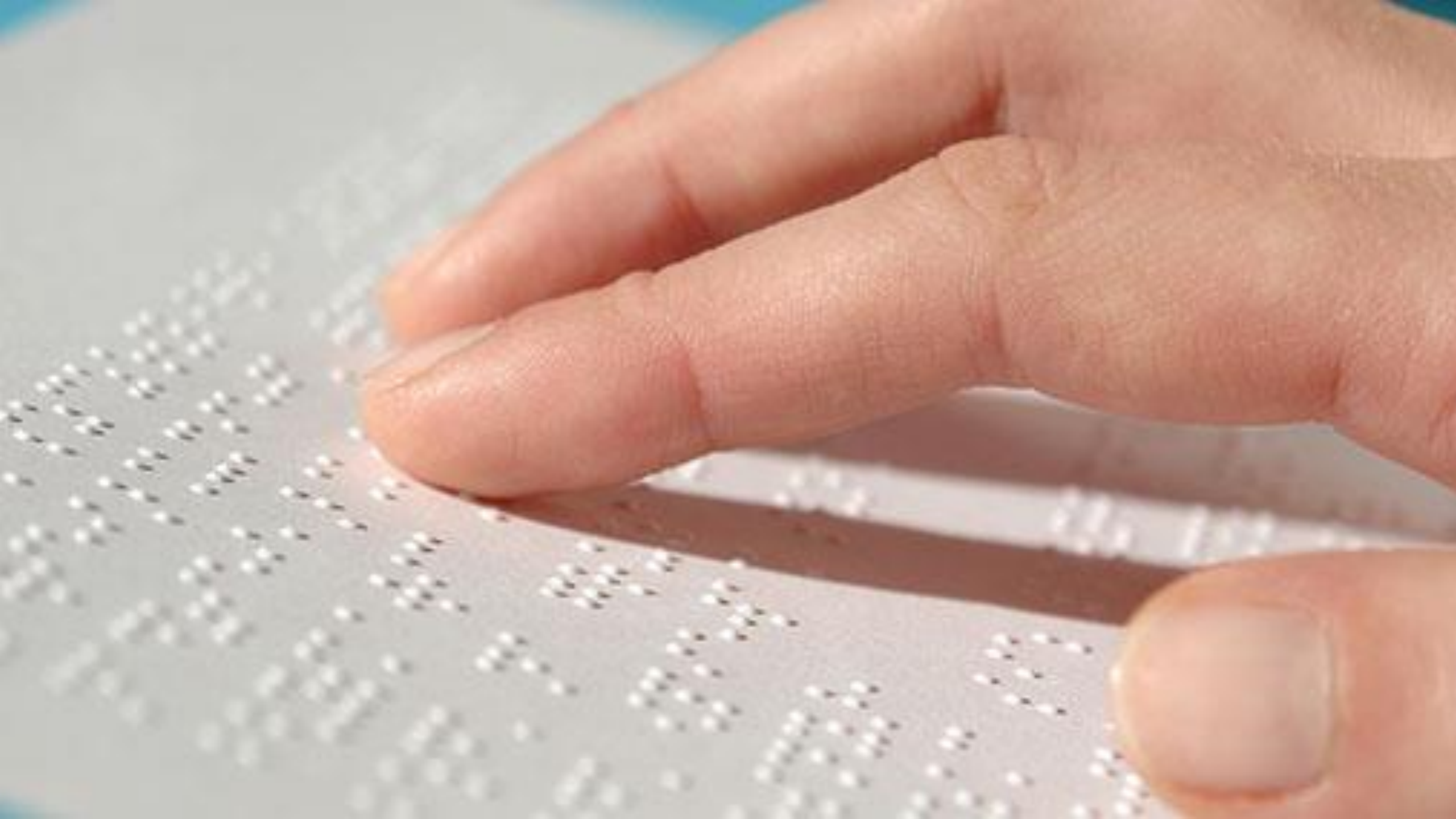
5. accessibility
   make computers work for more people

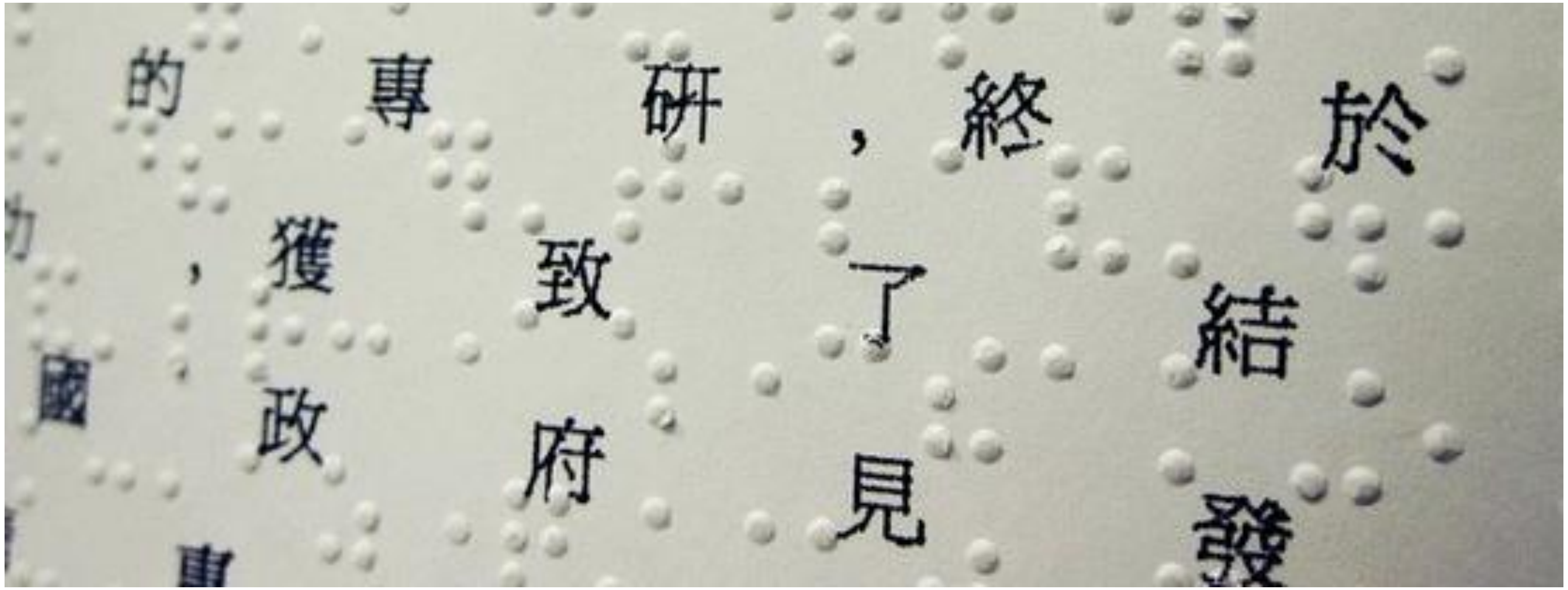
benefits of multimodal HCI:

4. learnability
   make skills easier to learn

5. accessibility
   make computers work for more people


JULY 25, 2020
NOW
THIS
So basically,
Gregory Lehto via Storyful





點			字			記			號		
⠝	⠞	⠈	⠓	⠱	⠐	⠅	⠡	⠐	⠗	⠩	⠐
ㄉ	ㄧㄢ	ˇ	ㄗ	ㄭ	ˋ	ㄐ	ㄧ	ˋ	ㄏ	ㄠ	ˋ

Forehead Electro-tactile Display for Vision Substitution

Hiroyuki Kajimoto[1]
The University of Tokyo

Yonezo Kanno[2]
EyePlusPlus, Inc.

Susumu Tachi[3]
The University of Tokyo

ABSTRACT

The Forehead Retina System-composed of a small camera and 512 electrodes on the forehead-captures the view in front, extracts outlines from the view, and converts the outlines to tactile sensation by electrical stimulation. Forehead skin was chosen for stimulation in consideration of usability. An electrotactile display was used so that the system becomes small and durable. A gel layer was designed to prevent an unpleasant sensation during stimulation. The system primarily aims to enable the visually impaired to "see" the surrounding environment..

Keywords: tactile display, electrocutaneous display, electrical stimulation, TVSS, visual prosthesis, virtual reality

1 INTRODUCTION

According to a WHO report in 2003, up to 45 million people are completely blind and 135 million people have low vision [11]. However, there exists no standard vision substitution system for their use in their daily lives. The goal of our project is to provide a cheap, lightweight, yet fully functional system that provides rich and dynamic 2D information to the blind.

The Forehead Retina System (FRS)—composed of a small camera and 512 electrodes on the forehead—captures the view in front, extracts outlines from the view, and converts the outlines to tactile sensation by electrical stimulation (Fig. 1). Using this device, the users can “see” the surrounding environment with their forehead skin, without using their eyes.

Although the system is a combination technology of computer vision, tactile display, and wearable interface, this paper mainly discusses the display aspect of the system because forehead electrical stimulation seems quite unusual.

1.1 Related works

There have been numerous proposals to restore the sense of sight by surgical means [1][2]. Direct stimulation of the visual

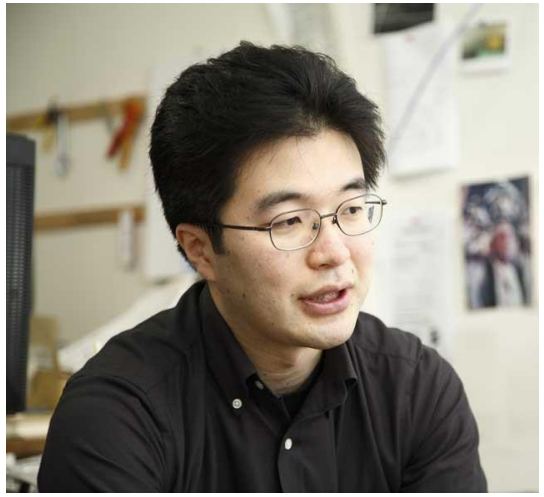
Fig. 1 Forehead Retina System (FRS).

On the other hand, the vision to tactile conversion system was first developed by Collins [3], who used 400 vibration motors on the skin of the back and a CCD (Charge Coupled Device) camera. He called the system a tactile vision substitution system (TVSS).

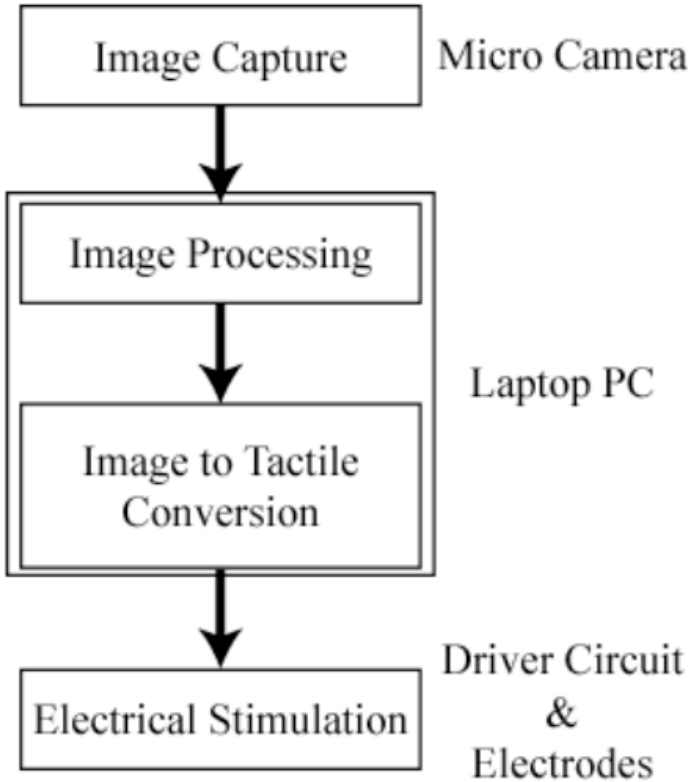
However, although many similar systems have been proposed since then, they have not become widely used for the following two reasons:

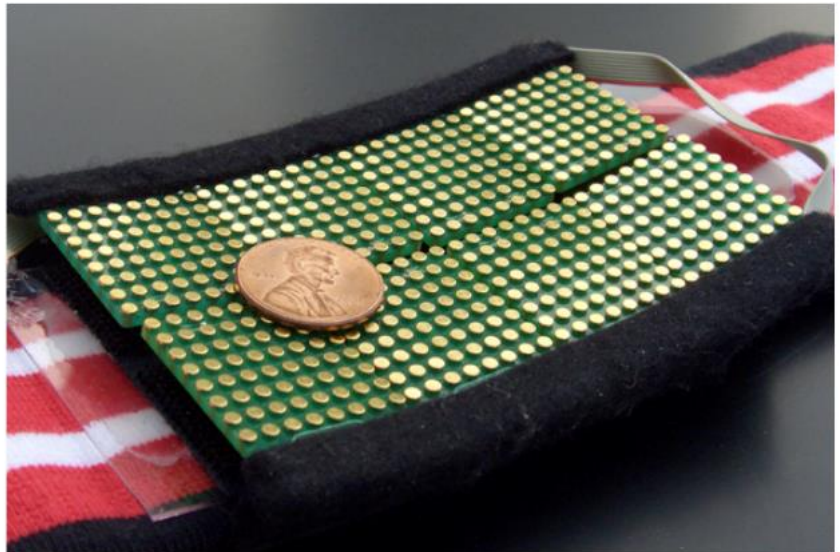
One is the technical aspect of the stimulation. For tactile devices using a large number of mechanical actuators, the system becomes heavy, expensive, and requires a considerable amount of power.

To make the device small and energy efficient, we must



Hiroyuki Kajimoto





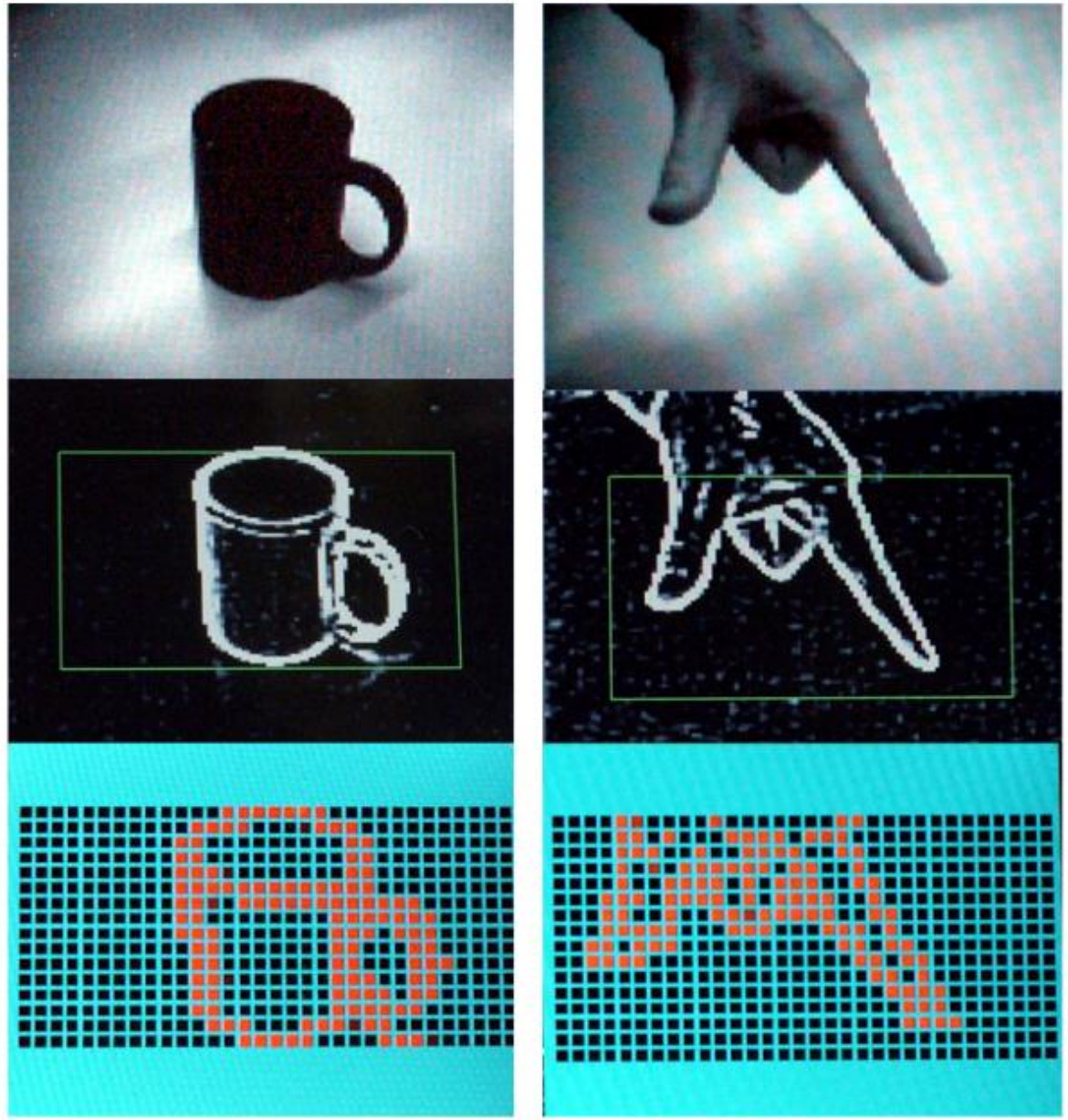


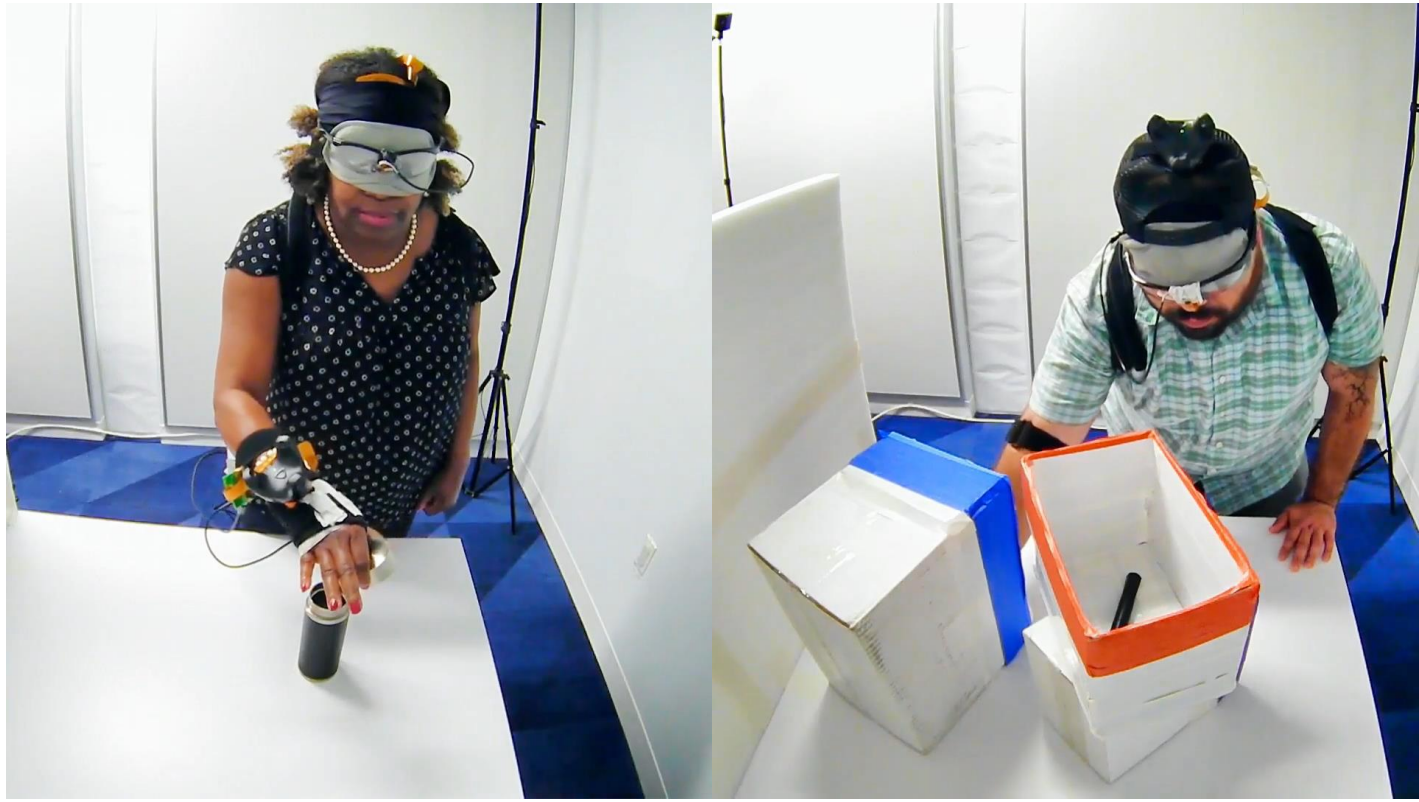
Fig. 4 The captured image is converted to a tactile pattern. (Top) Raw image. (Middle) Edge extraction. (Bottom) Tactile pattern.

BrainPort

traditional sensory substitution
sees from **eyes' perspective**

obstacles recognition

traditional sensory substitution
sees from **eyes' perspective**

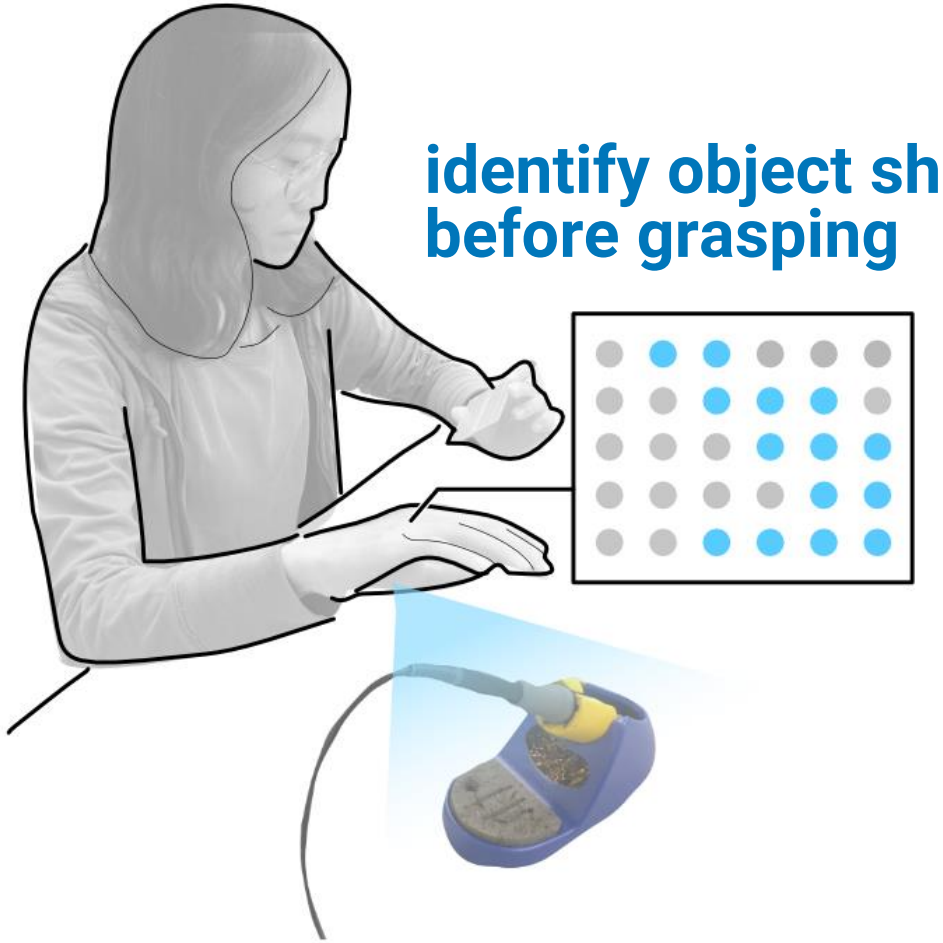


use **hands** for grasping

traditional sensory substitution
sees from **eyes' perspective**

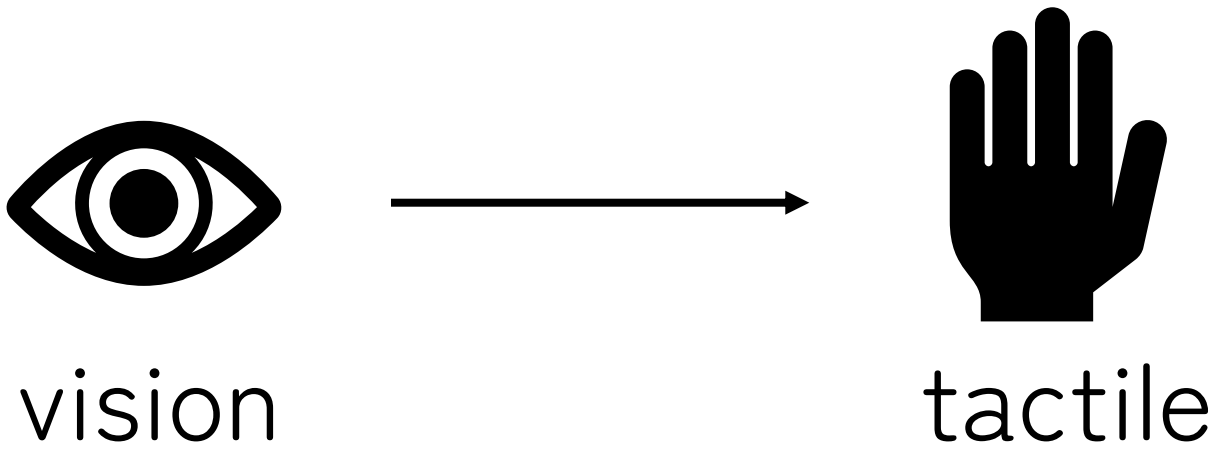
we propose seeing from
hands' perspective

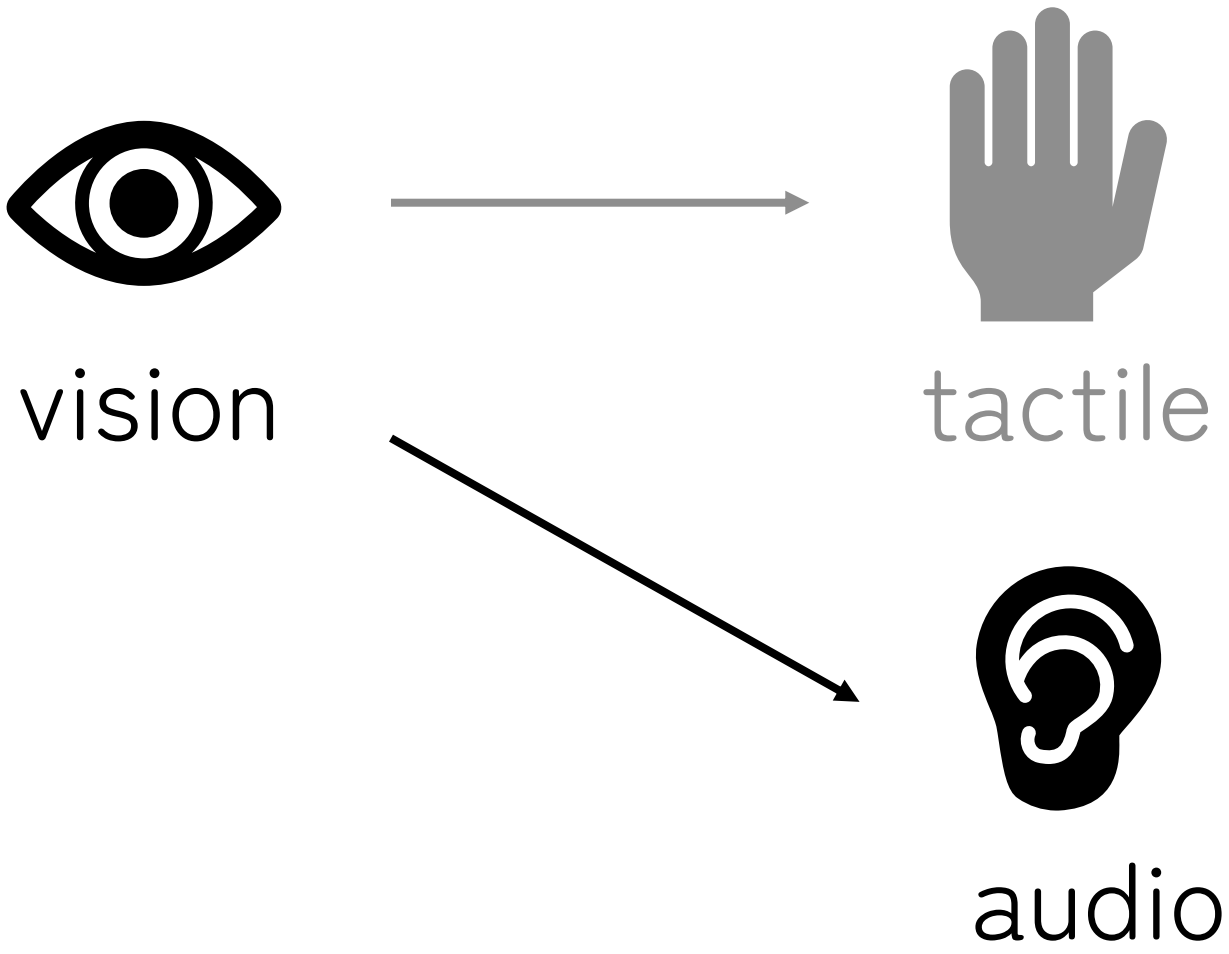
identify object shapes before grasping



this **blind user** wants to use a **hot kettle**


our approach can be **used together** with
traditional tactile perspective (e.g., eyes)


vision
tactile

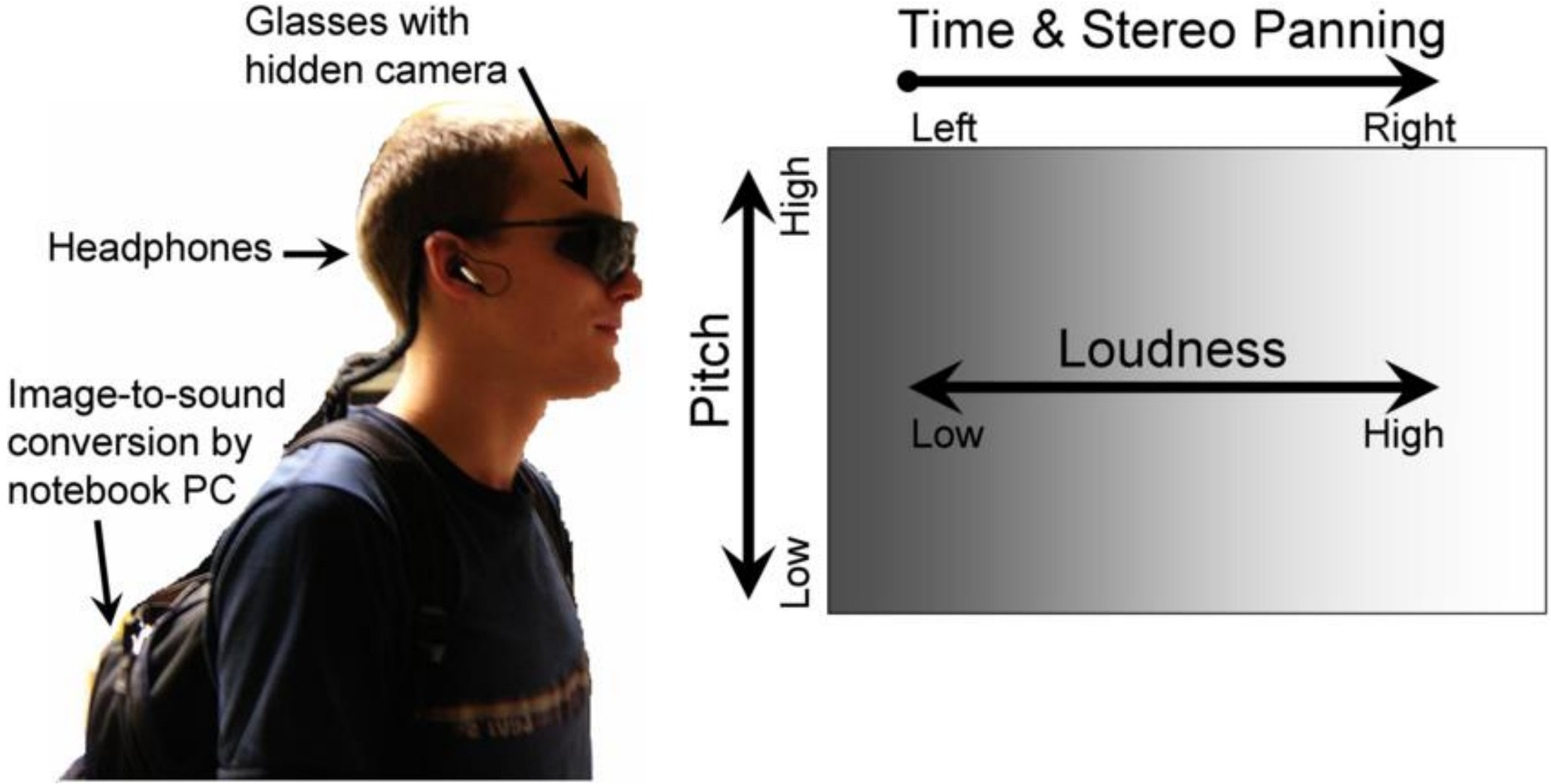

vision
tactile
audio

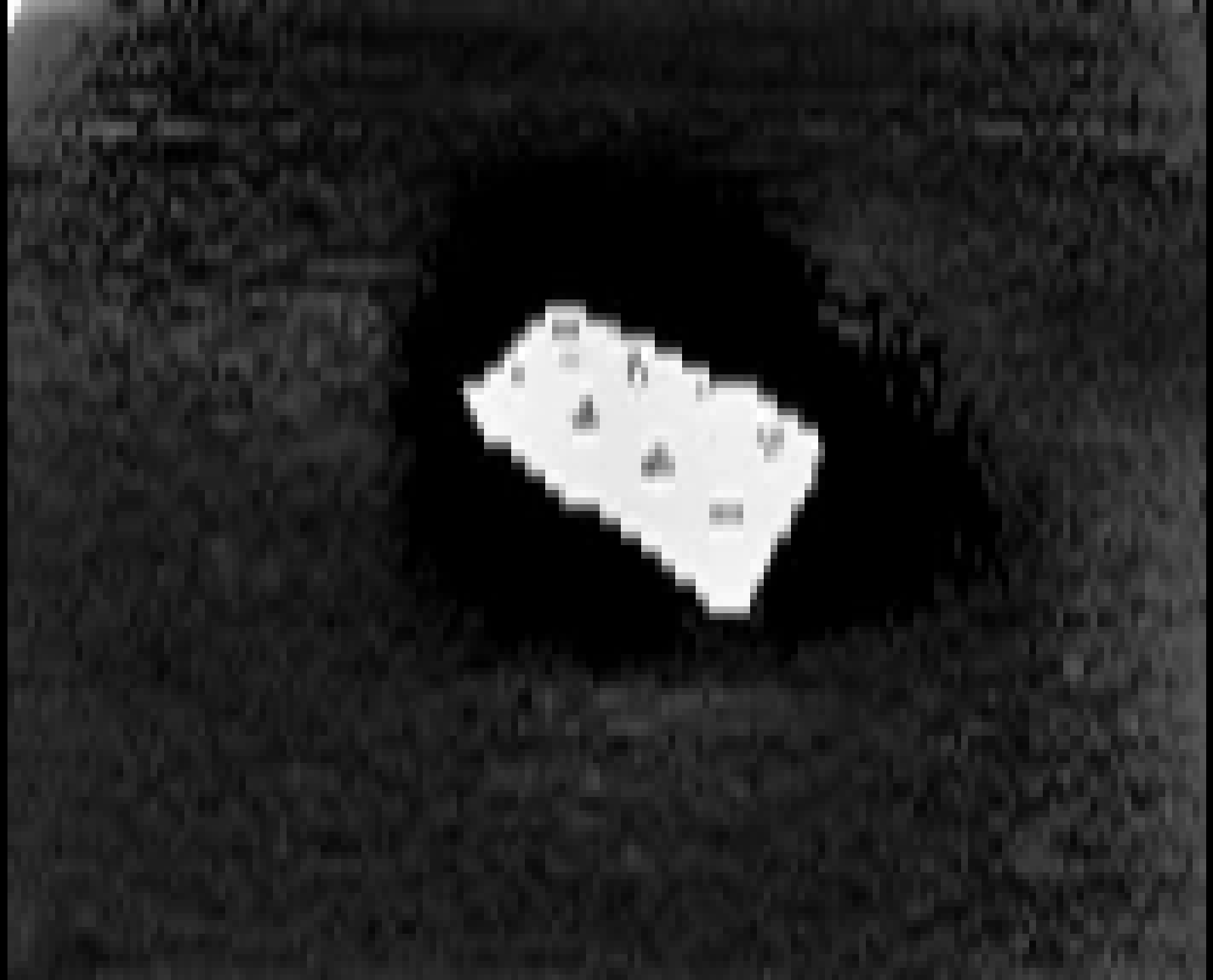
WorldScribe

Towards Context-Aware Live Visual Descriptions

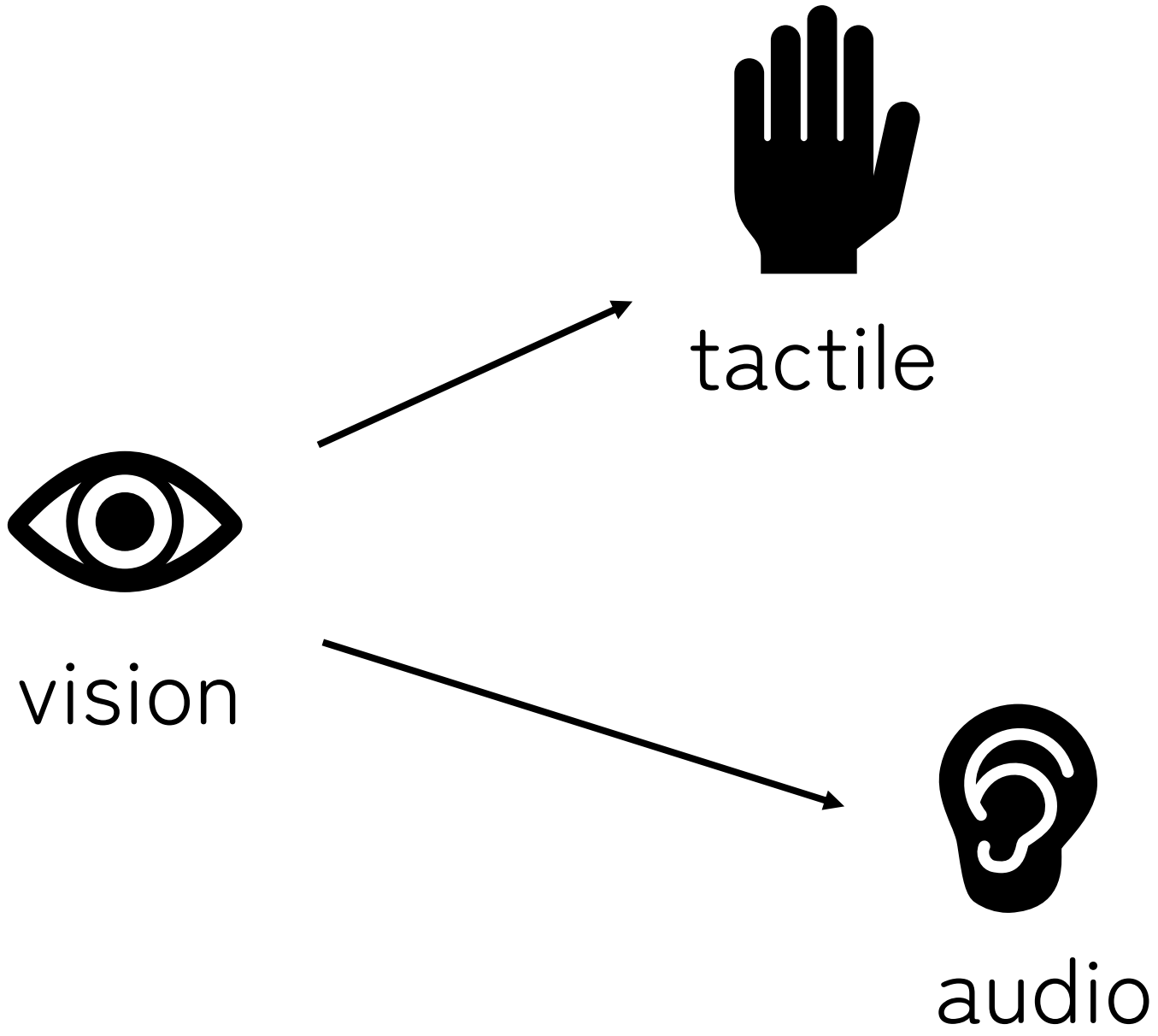
Ruei-Che Chang
Yuxuan Liu
Anhong Guo

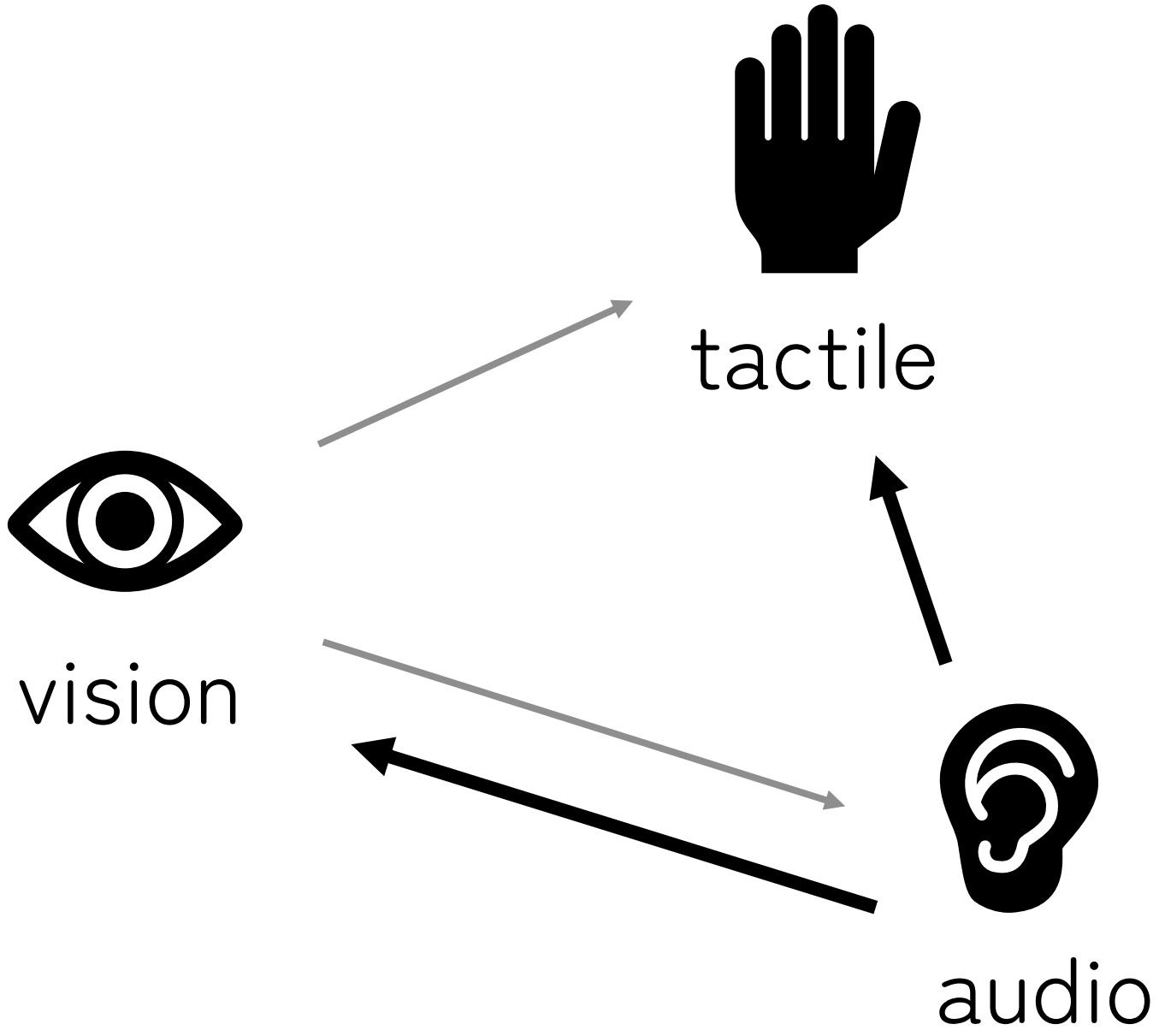
UNIVERSITY OF MICHIGAN

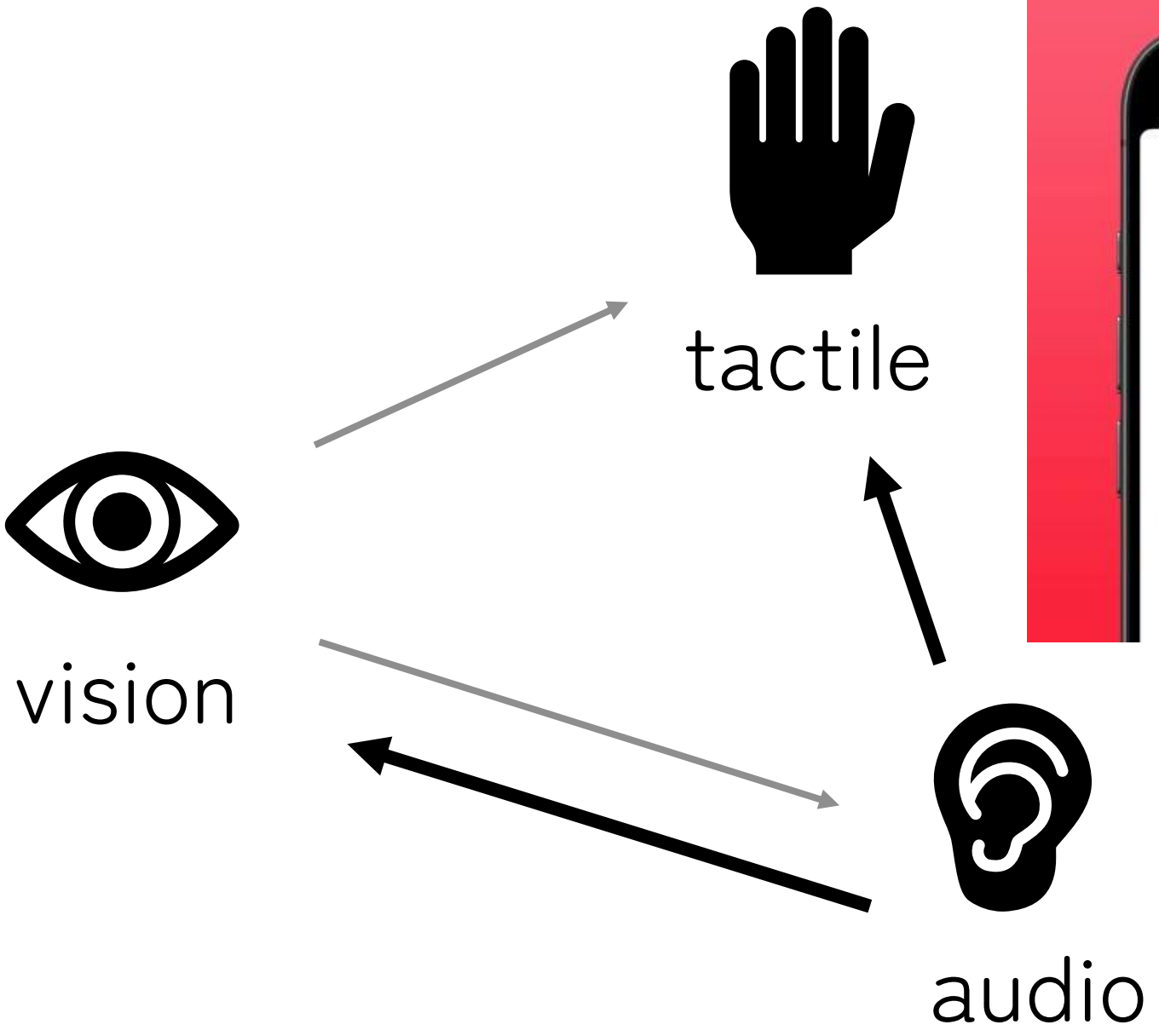

Glasses with
hidden camera
Headphones
Image-to-sound
conversion by
notebook PC
Time & Stereo Panning
Left
Right
Pitch
High
Low
Loudness
Low
High

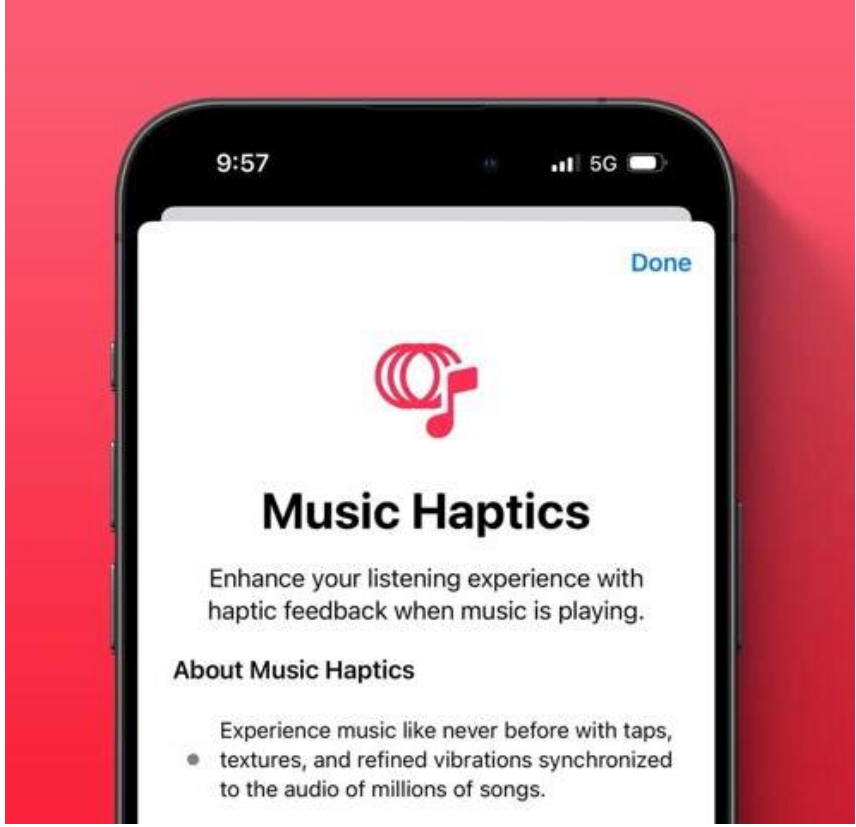


[vOICe]


tactile
vision
audio


tactile
vision
audio


tactile
vision
audio


9:57
5G
Done
Music Haptics
Enhance your listening experience with haptic feedback when music is playing.
About Music Haptics
Experience music like never before with taps, textures, and refined vibrations synchronized to the audio of millions of songs.

situational impairment

situational impairment

situational impairment

situational impairment

A. Egocentric Video&Audio

B. Multimodal Sensing

Activity: User is preparing food in a kitchen.
Environment: User is in a kitchen.
Hand: Holding a bowl
Volume: 65 dB

Activity: User is attending an outdoor concert.
Environment: User is at an outdoor concert.
Hand: Not detected
Volume: 112 dB

Activity: User is taking notes on a tablet device.
Environment: User is in an office.
Hand: Taking notes
Volume: 42 dB

Activity: User is showering a dog in a bathtub.
Environment: User is in a bathroom.
Hand: Showering a dog
Volume: 58 dB

C. Channel Availability

- **Vision / Eyes**: Slightly Affected
- **Hearing**: Available
- **Vocal System**: Available
- **Hands / Fingers**: Affected

- **Vision / Eyes**: Slightly Affected
- **Hearing**: Unavailable
- **Vocal System**: Unavailable
- **Hands / Fingers**: Available

- **Vision / Eyes**: Affected
- **Hearing**: Available
- **Vocal System**: Available
- **Hands / Fingers**: Affected

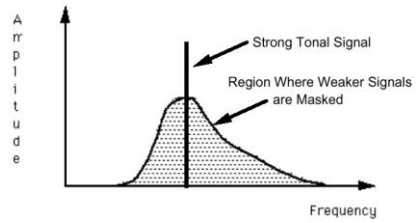
- **Vision / Eyes**: Affected
- **Hearing**: Available
- **Vocal System**: Available
- **Hands / Fingers**: Unavailable

[Human I/O]

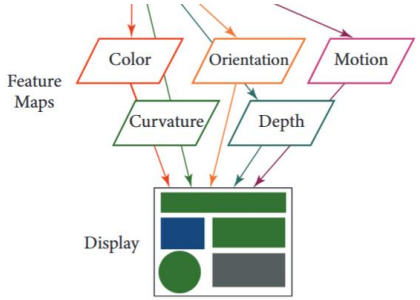
benefits of multimodal HCI:

benefits of multimodal HCI:

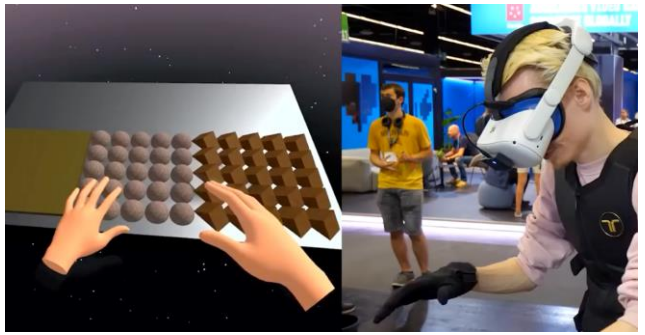
1. efficiency
   make computers more efficient



2. usability
   make computers easier to use

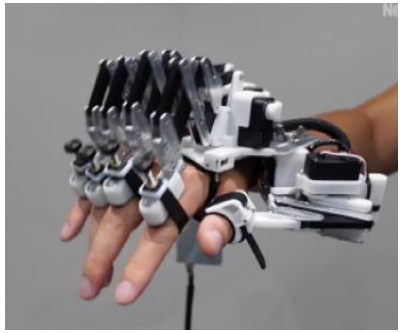


3. fidelity
   make computers feel more realistic



benefits of multimodal HCI:

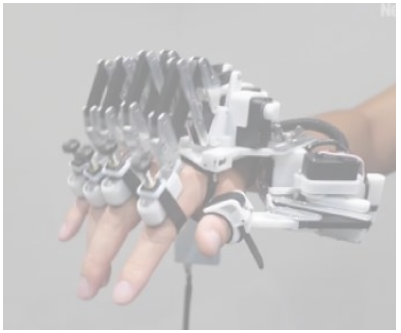
4. learnability
   make skills easier to learn



5. accessibility
   make computers work for more people

benefits of multimodal HCI:

4. learnability
   make skills easier to learn

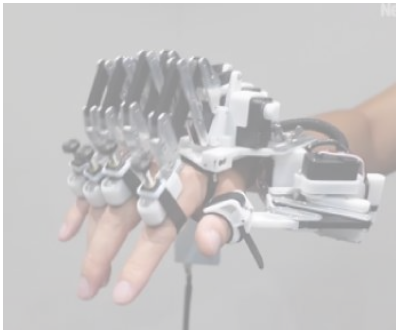


5. accessibility
   make computers work for more people

6+: communication, well-being etc.

benefits of multimodal HCI:

4. learnability
   make skills easier to learn



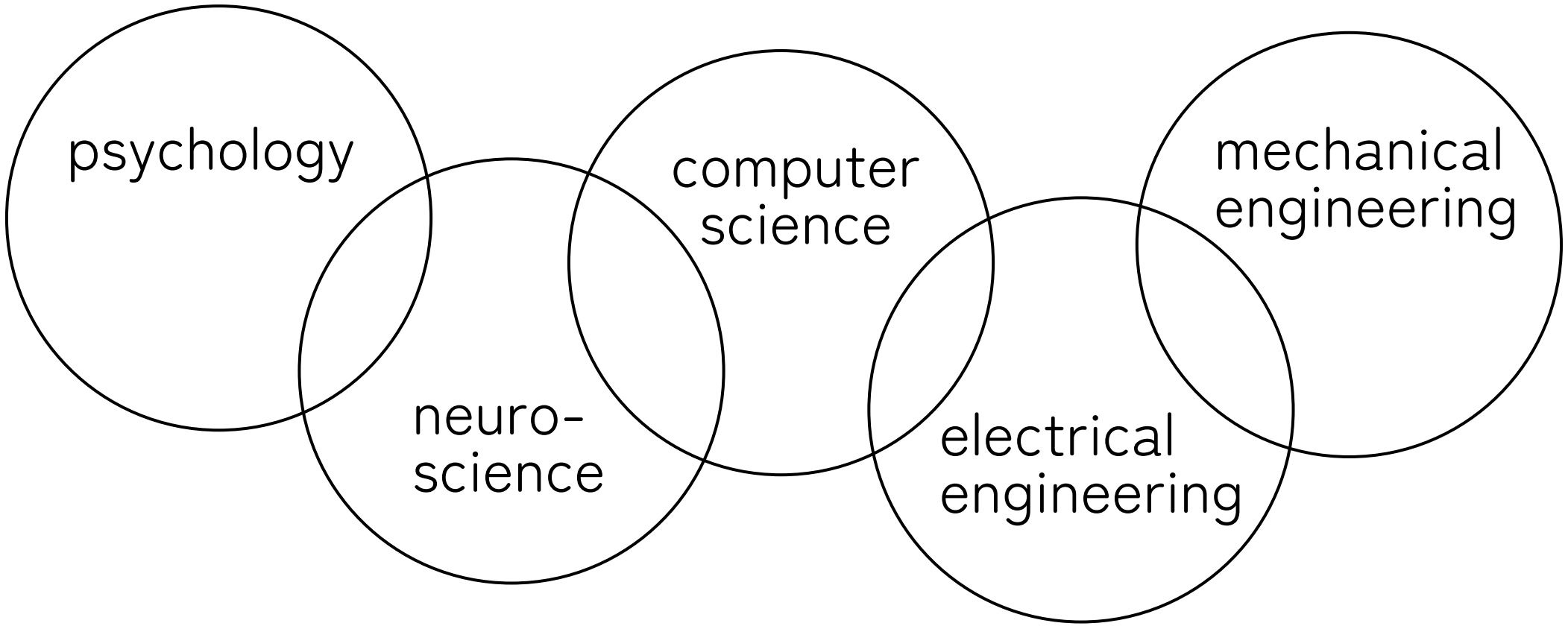
5. accessibility
   make computers work for more people

6+: communication, well-being etc.

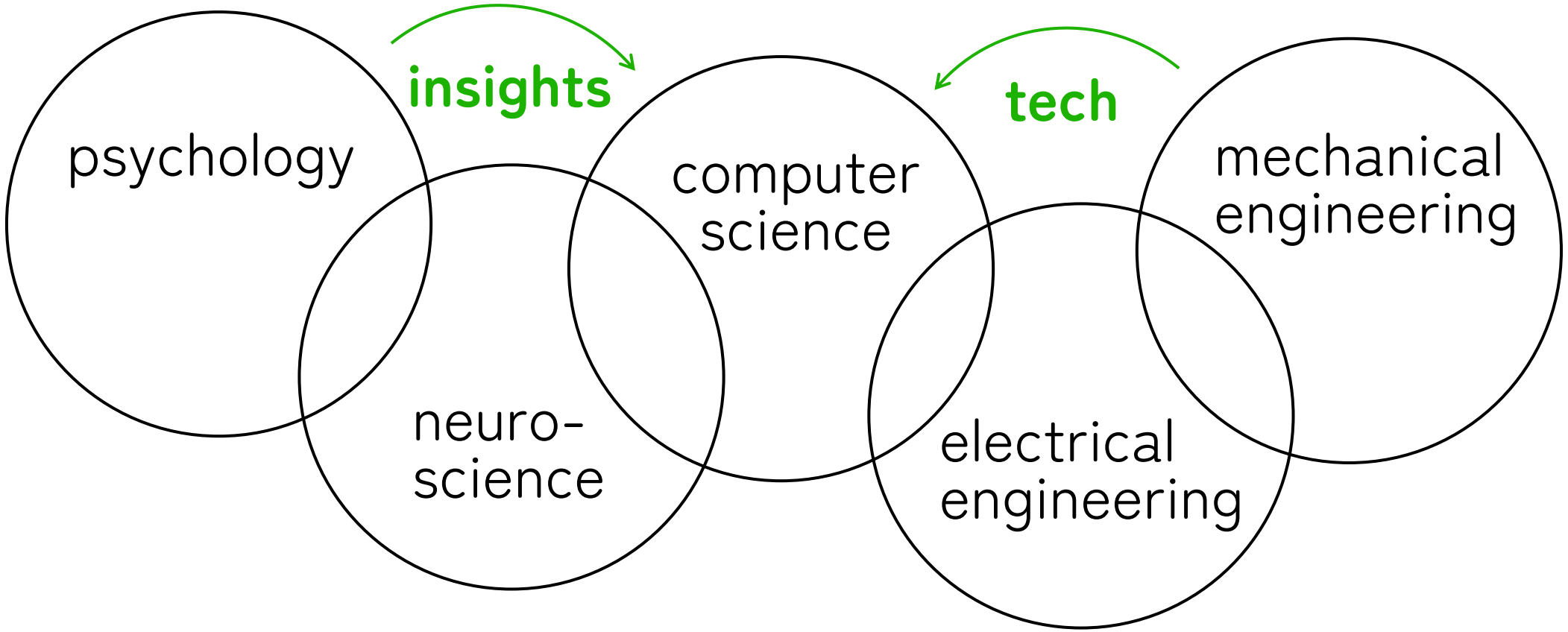
going deeper? next semester: Multimodal HCI class (in English)

multimodal HCI is interdisciplinary

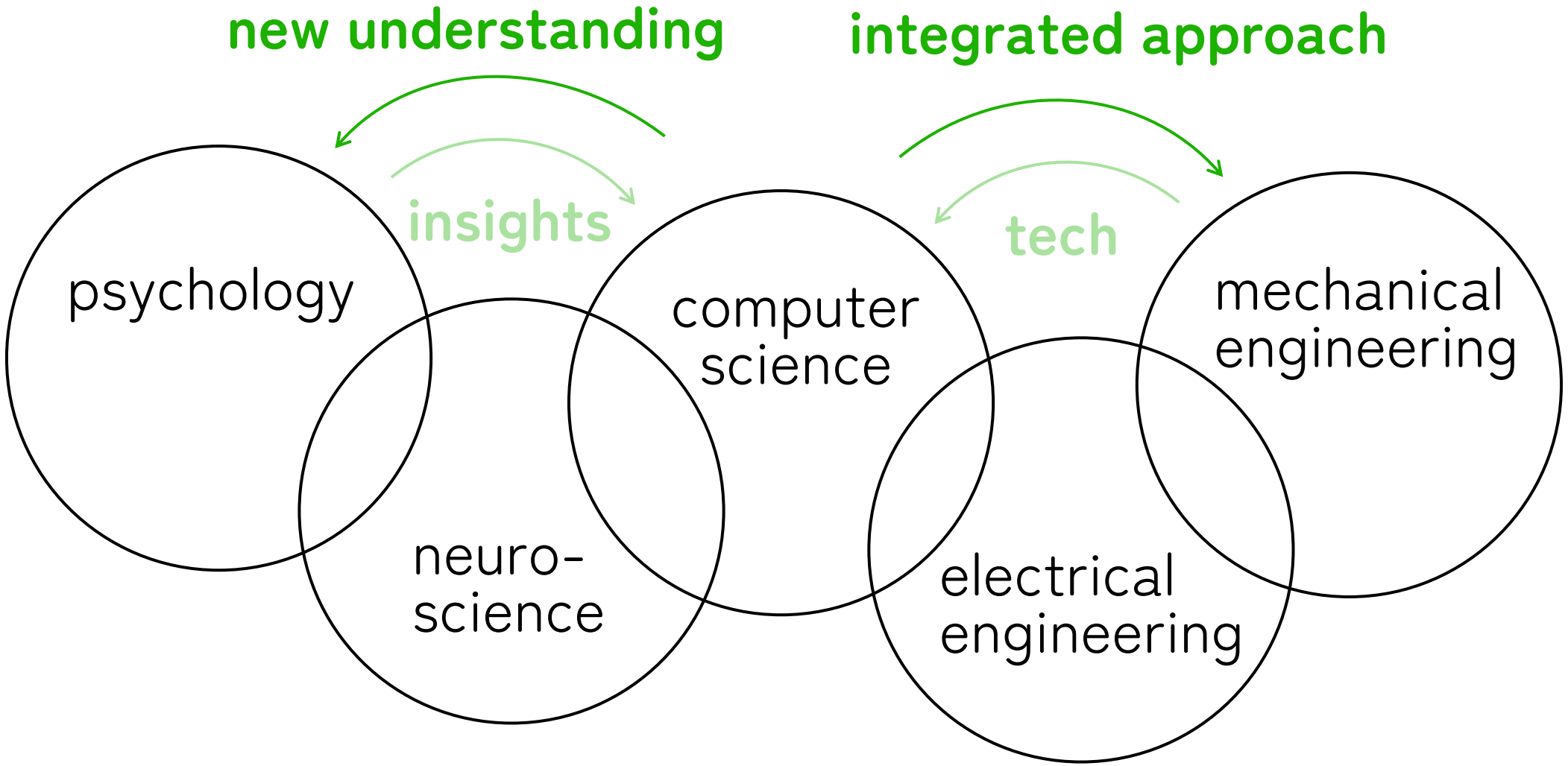
multimodal HCI is interdisciplinary



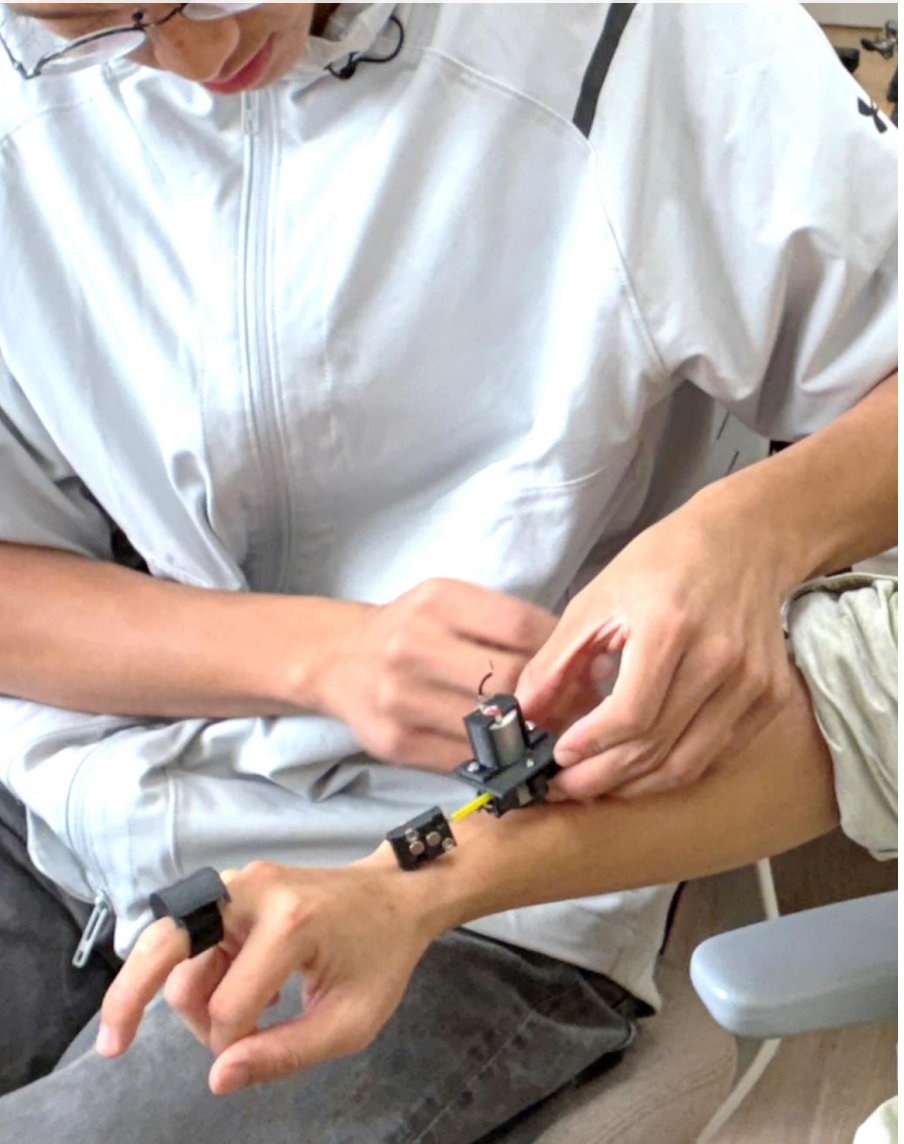
multimodal HCI is interdisciplinary



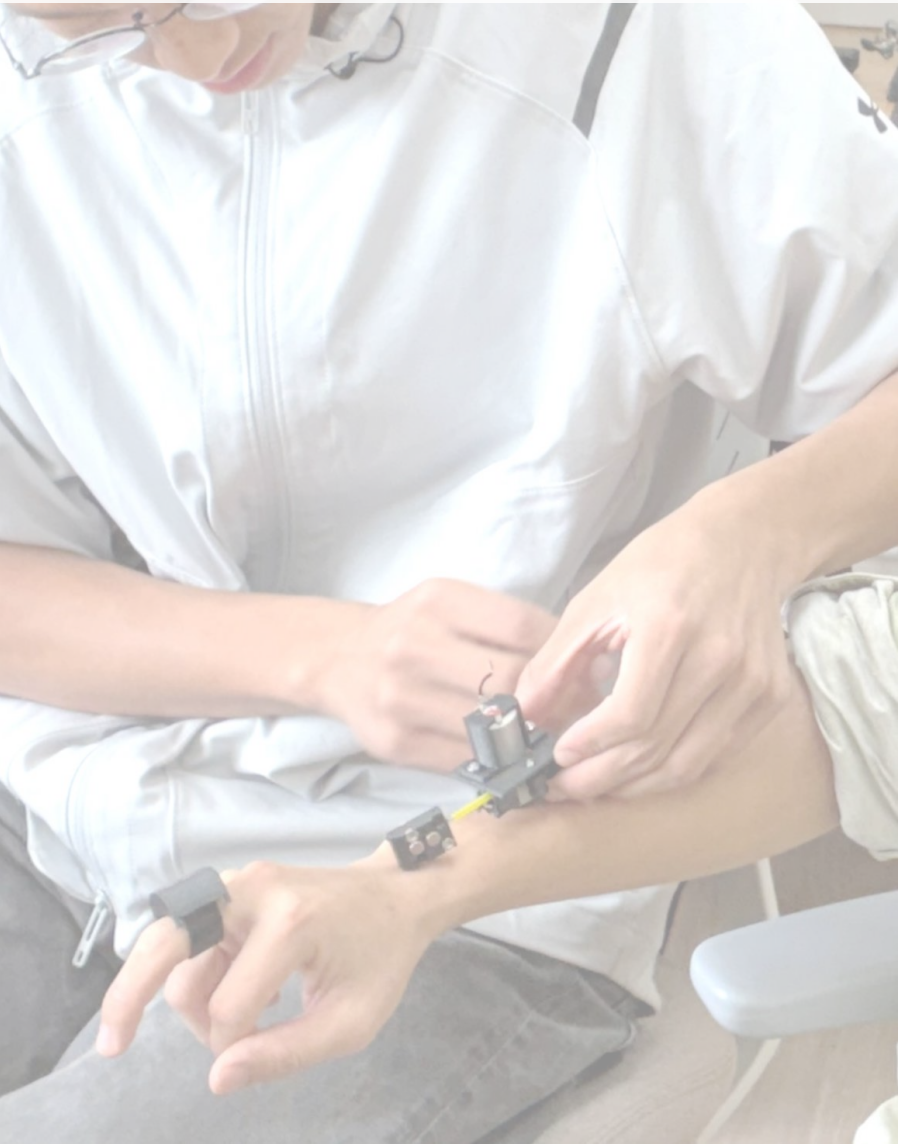
multimodal HCI is interdisciplinary



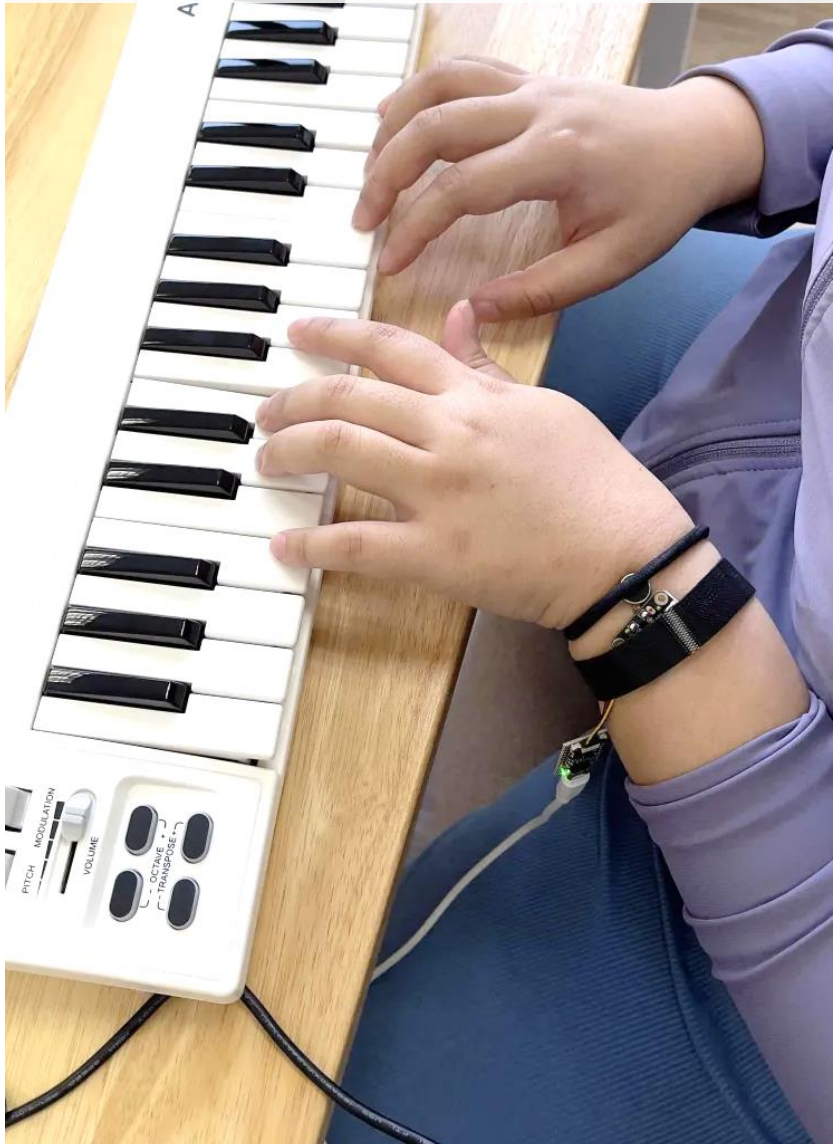
assistive robotics



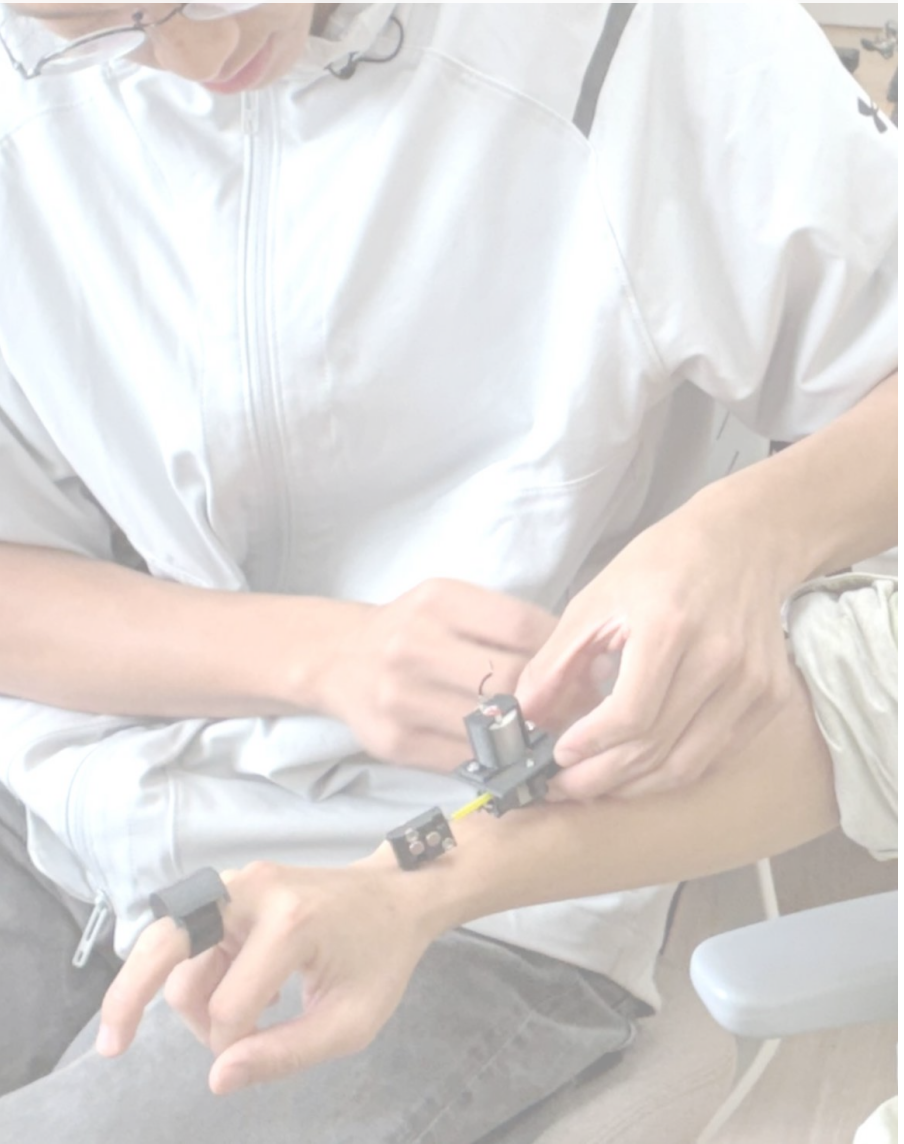
assistive robotics



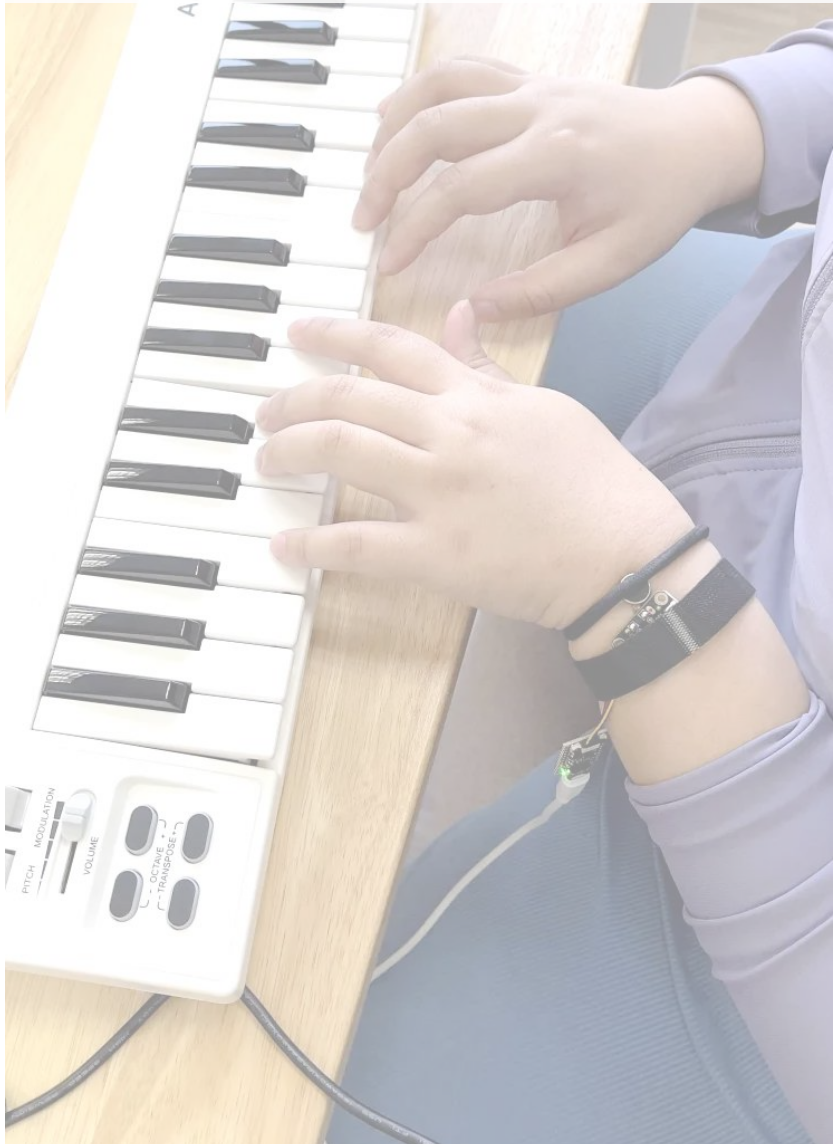
multimodal learning



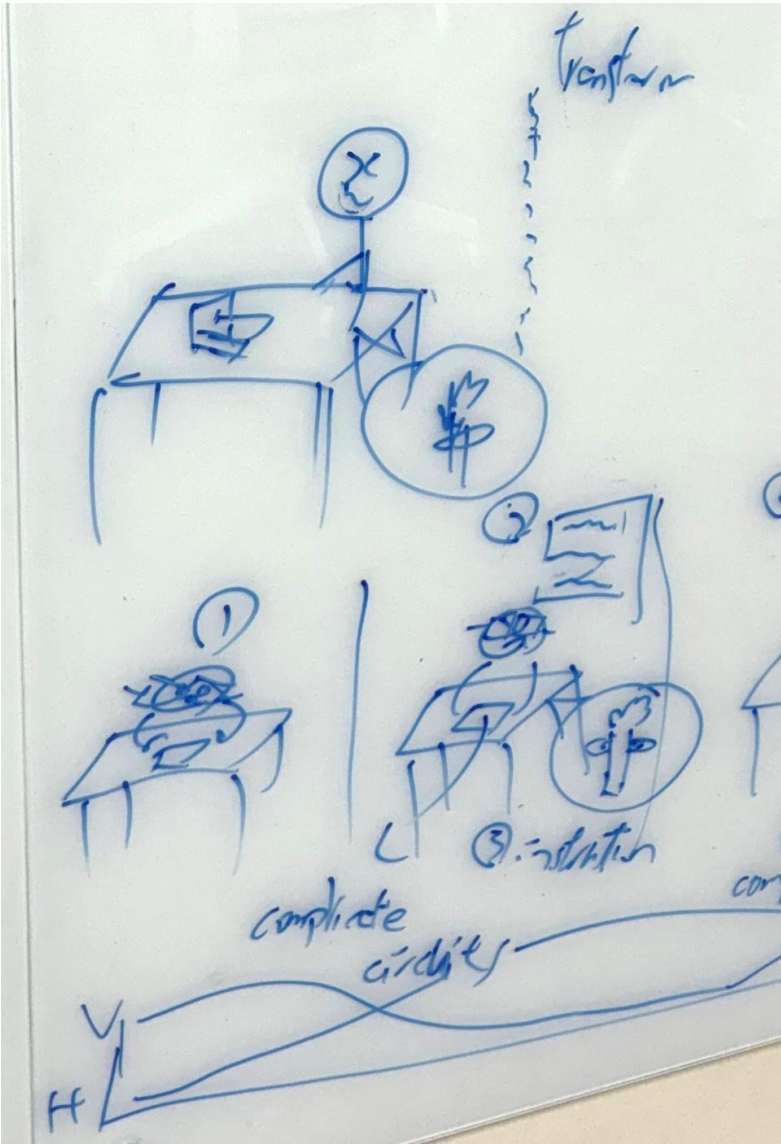
assistive robotics



multimodal learning



adaptive sensory



thank you!

dexterous interaction lab

靈 巧 互 動 實 驗 室

recruiting students!

https://lab.tengshanyuan.info/