Decoding multisensory attention from electroencephalography for use in brain-computer interface

Wenkang (Winko) An

Microsoft Research, Audio and Acoustics Group

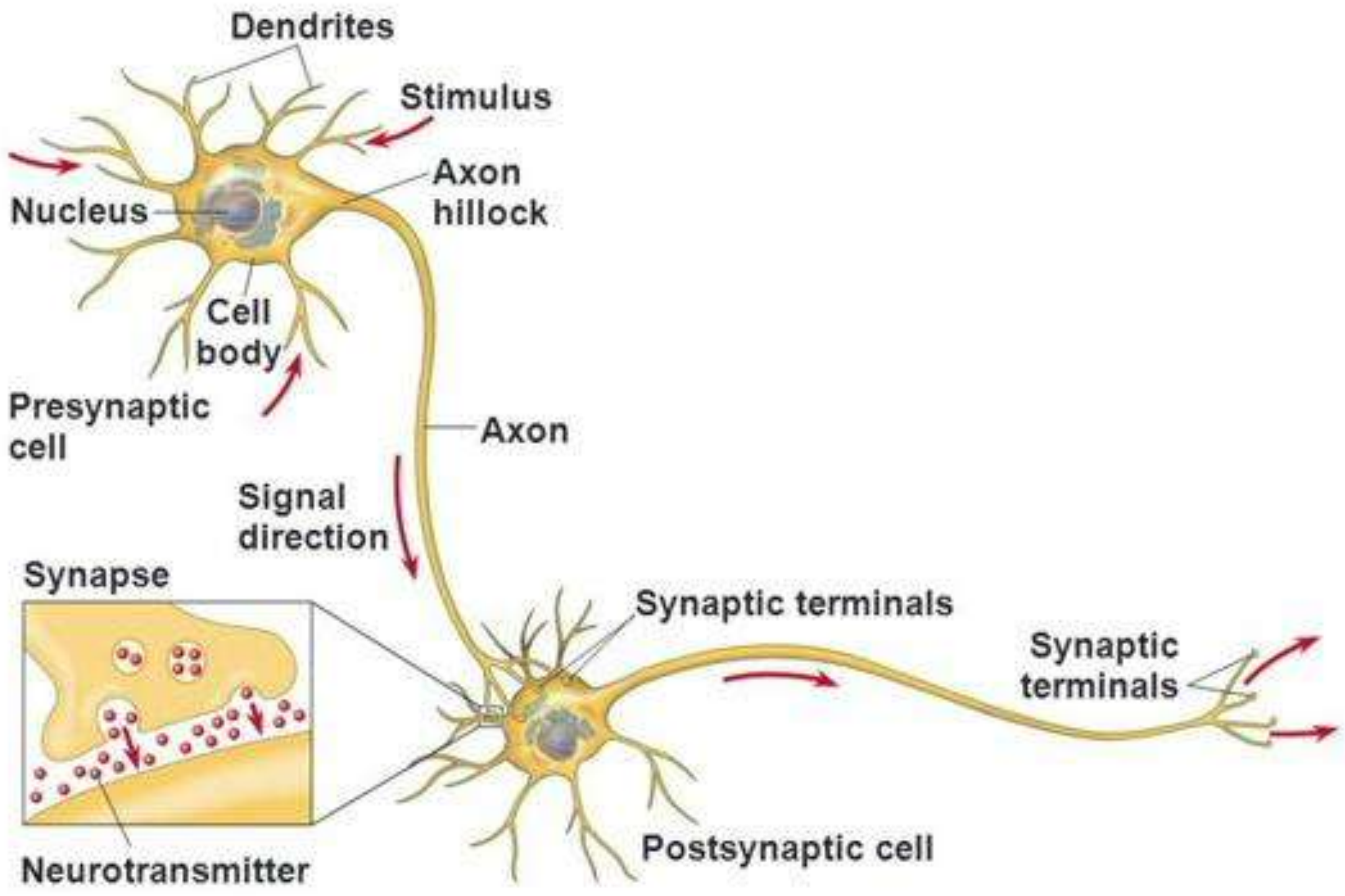
Mentor: Hannes Gamper

Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

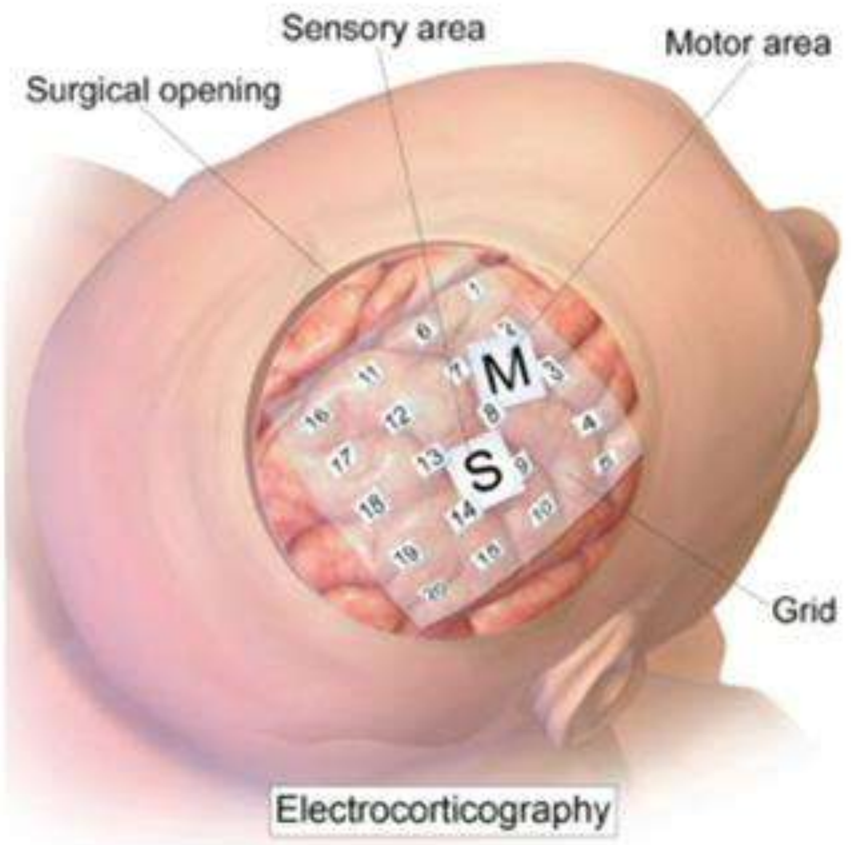
Neurons

- ~100 million – 100 billion neurons in a brain
- Chemical signal at synapses
- Electrical signal for long-distance communication



Electrocorticography (ECoG)

- Record electrical activity from the cortex
  - Invasive
- High temporal resolution
- High spatial resolution
- High signal-to-noise ratio

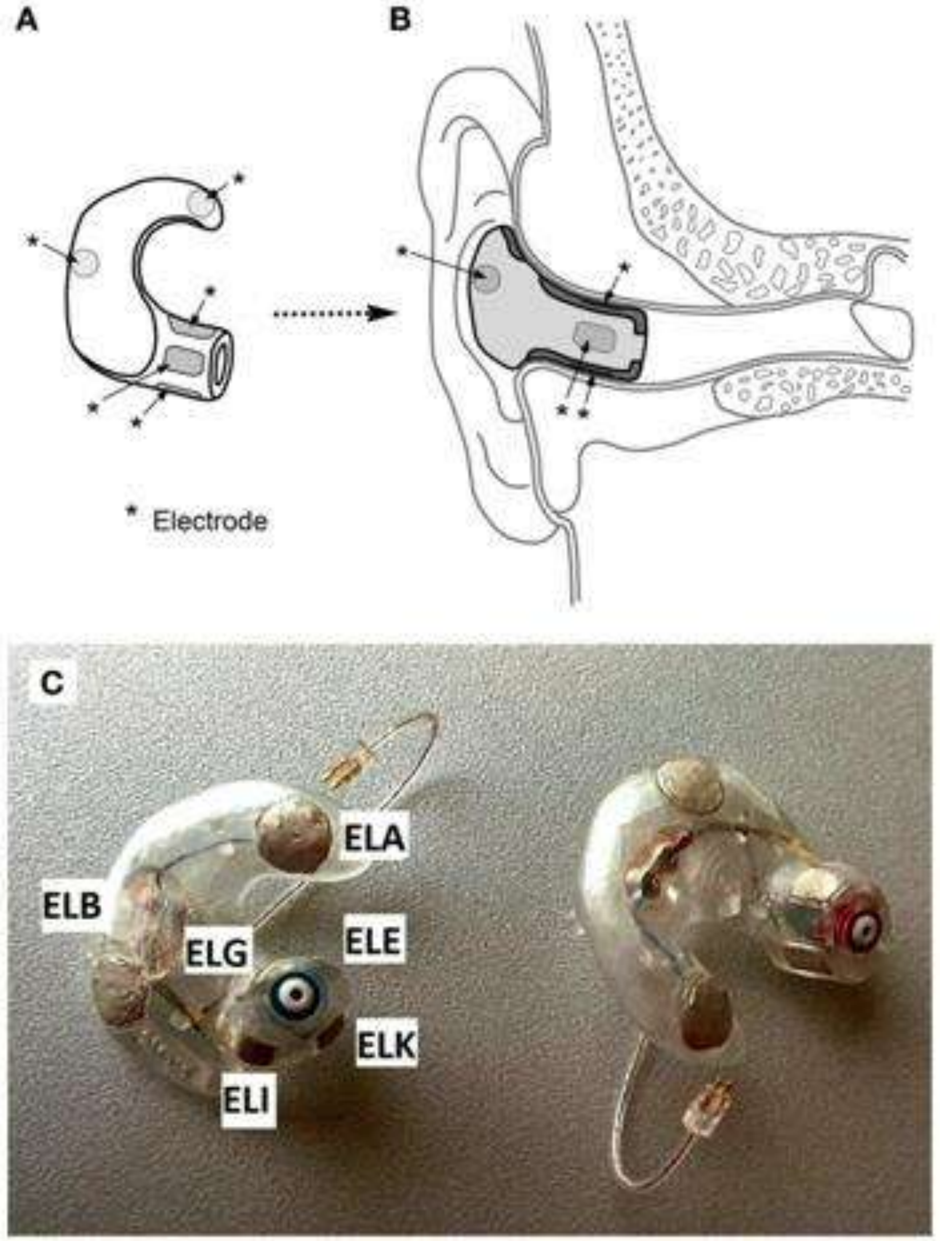


Surface electroencephalography (EEG)

- Electrical potential along the scalp (μV)
  - Non-invasive
- High temporal resolution (>1kHz)
- Low spatial resolution (cm)
- Low signal-to-noise ratio (<0dB)

In-ear EEG

- Measure EEG signal in/around the ear
- Pros
  - Unobtrusive
  - Consistent placement
- Cons
  - Even lower signal-to-noise ratio
  - No spatial information



[1]

Surface electroencephalography (EEG)

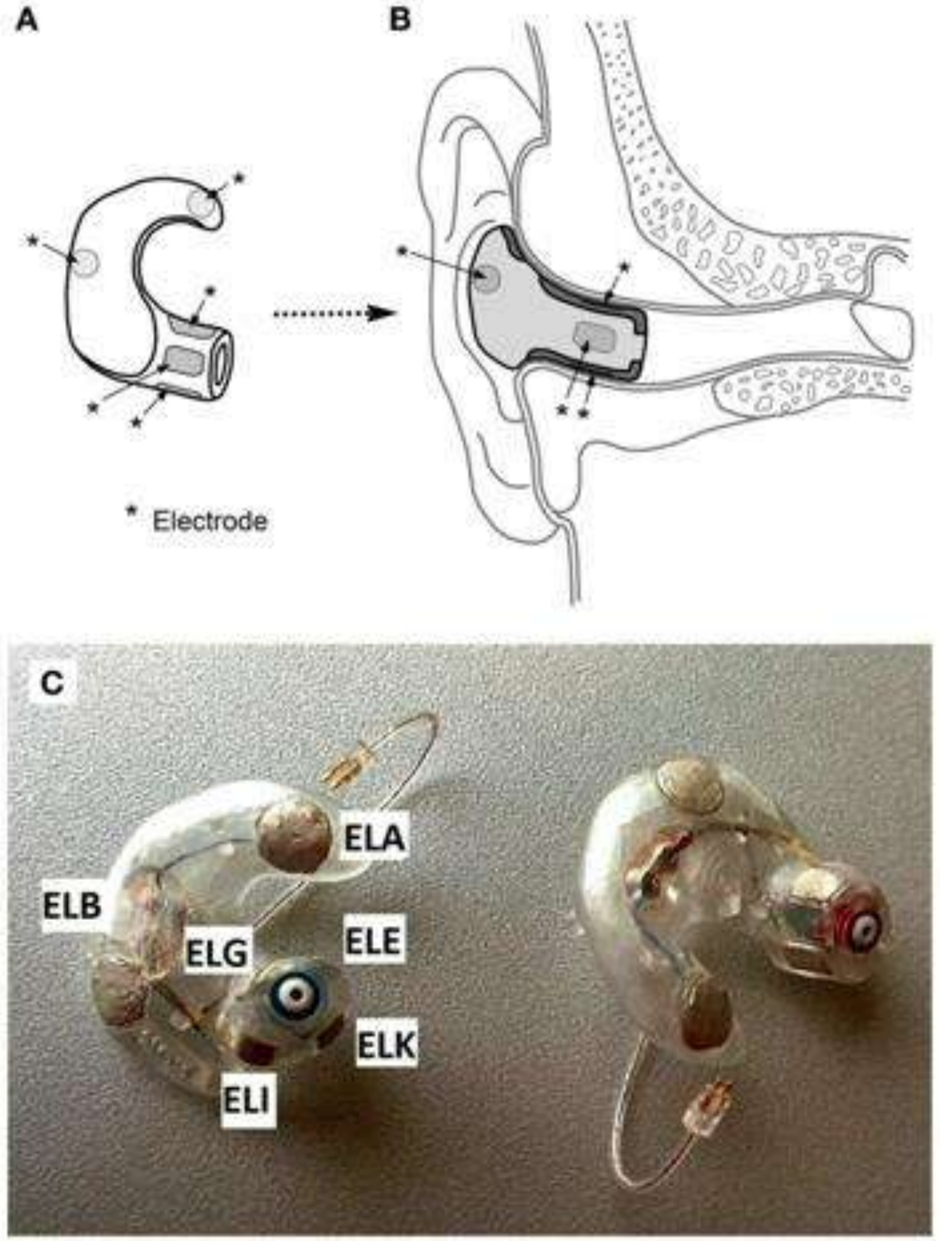
- Electrical potential along the scalp (μV)
  - Non-invasive
- High temporal resolution (>1kHz)
- Low spatial resolution (cm)
- Low signal-to-noise ratio (<0dB)

In-ear EEG

- Measure EEG signal in/around the ear

- Pros
  - Unobtrusive
  - Consistent placement

- Cons
  - Even lower signal-to-noise ratio
  - No spatial information



[1]

Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

Brain-computer Interface (BCI)

- Definition: A communication or control system that allows **real-time** interaction between the human brain and the external devices[2].
- Applications
  - Medical: ECoG during surgery
  - Psychological: stress and emotion monitoring
  - Interactive: translate brain state to a command or an action

BCI paradigms

Paradigms		Task	Cons	Information transfer rate (ITR)
Motor Imagery	Sensorimotor Rhythms	Imagining body parts movement	Weeks/Months of training[5]	~20 bits/min[3]
	Imagined Body Kinematics	Imagining continuous movement of only one body part	Usually used in ECoG, poor decoding in EEG[4]	NA
External Stimulation	Visual P300	Focus on a visual object with infrequently presented events	Visual focus required[5]	15-25 bits/min[5]
	Steady-state Visual Evoked Potential (SSVEP)	Focus on a flickering object	Fatigue due to flickering[5]	~20 bits/min[6]
	Auditory Steady-state Response (ASSR)	Attention to pure tone with a constant modulation frequencies	Low performance[5]	~1.5 bits/min[6]
	Steady-state Somatosensory Evoked Potential (SSSEP)	Attention to vibration with a constant (modulation) frequency	Low performance	~1.5 bits/min [6]

BCI paradigms

Paradigms		Task	Cons	Information transfer rate (ITR)
Motor Imagery	Sensorimotor Rhythms	Imagining body parts movement	Weeks/Months of training[5]	~20 bits/min[3]
	Imagined Body Kinematics	Imagining continuous movement of only one body part	Usually used in ECoG, poor decoding in EEG[4]	NA
External Stimulation	Visual P300	Focus on a visual object with infrequently presented events	Visual focus required[5]	15-25 bits/min[5]
	Steady-state Visual Evoked Potential (SSVEP)	Focus on a flickering object	Fatigue due to flickering[5]	~20 bits/min[6]
	Auditory Steady-state Response (ASSR)	Attention to pure tone with a constant modulation frequencies	Low performance[5]	~1.5 bits/min[6]
	Steady-state Somatosensory Evoked Potential (SSSEP)	Attention to vibration with a constant (modulation) frequency	Low performance	~1.5 bits/min [6]

Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

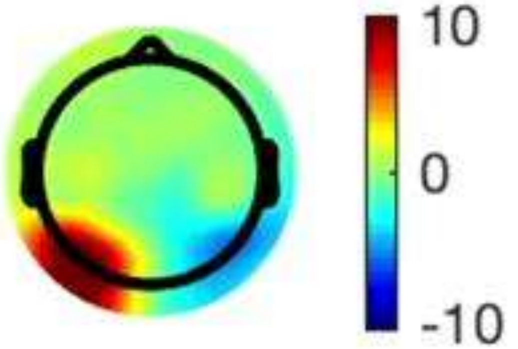
Research objectives

- Investigate the feasibility of using audio and tactile stimuli to build a functional BCI system
- Investigate the feasibility of using a interactive task-based paradigm for BCI system
- Compare the attention decoding between using audio and tactile stimuli

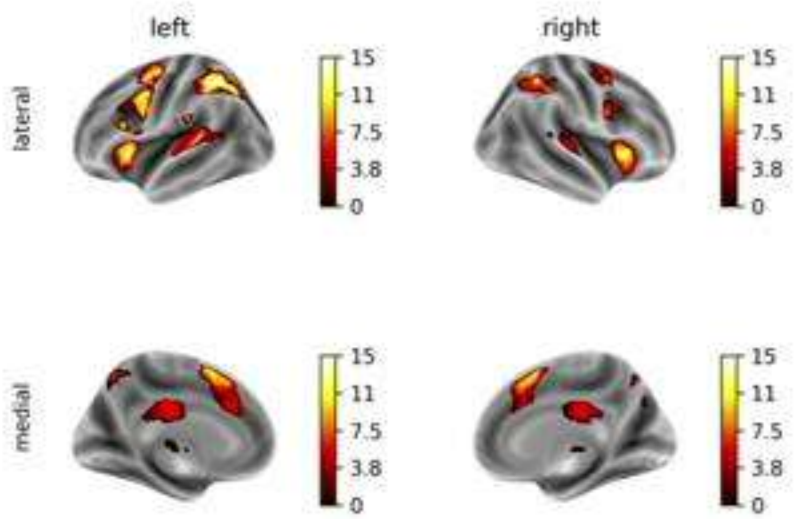
Literature review: decode auditory attention

- Auditory spatial attention
  - Power increase in the (8-14Hz) band
  - Detectable on group average level

Attend left – Attend right



Spatial > Control



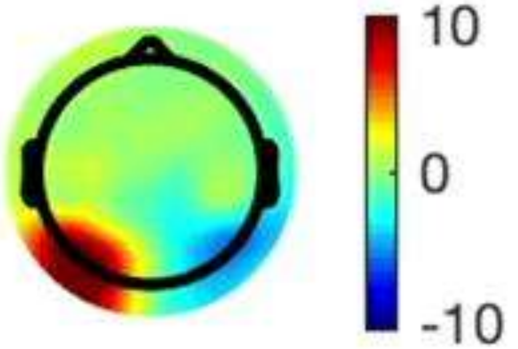
[6]

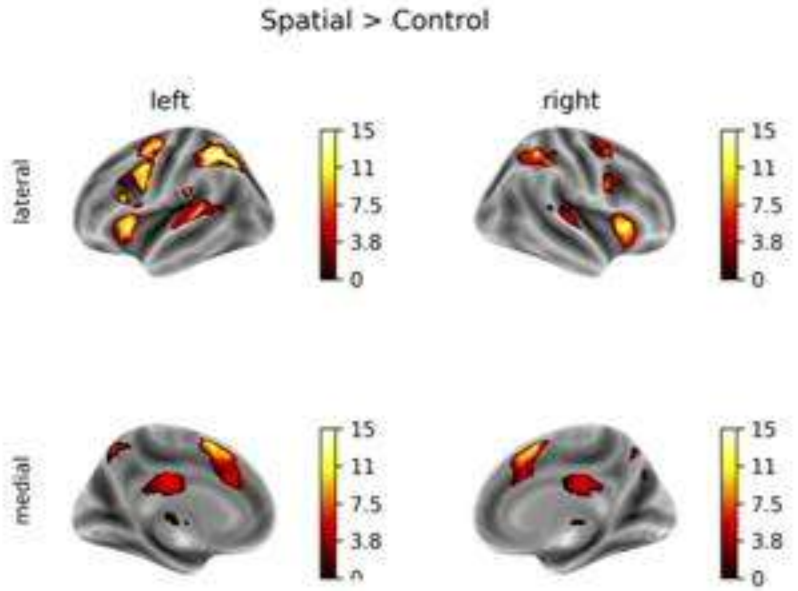
Literature review: decode auditory attention

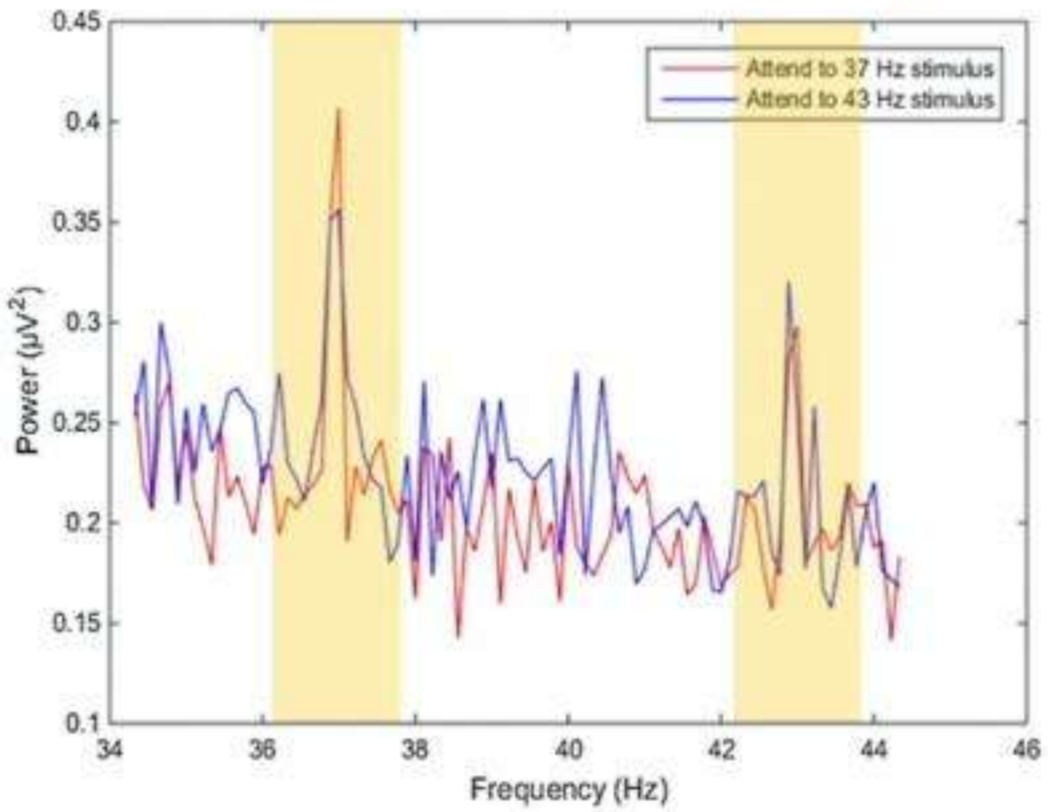
- Auditory spatial attention
  - Power increase in the (8-14Hz) band
  - Detectable on group average level

- Auditory steady-state response
  - Power increase at the modulation frequency of the attended stream

Attend left – Attend right



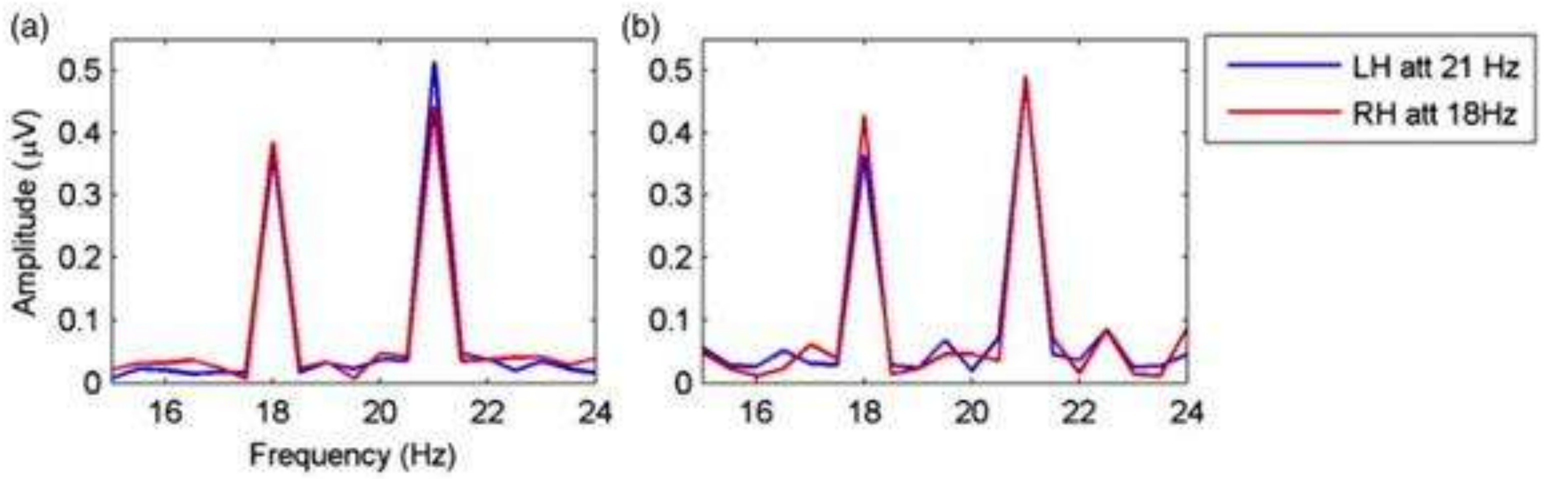


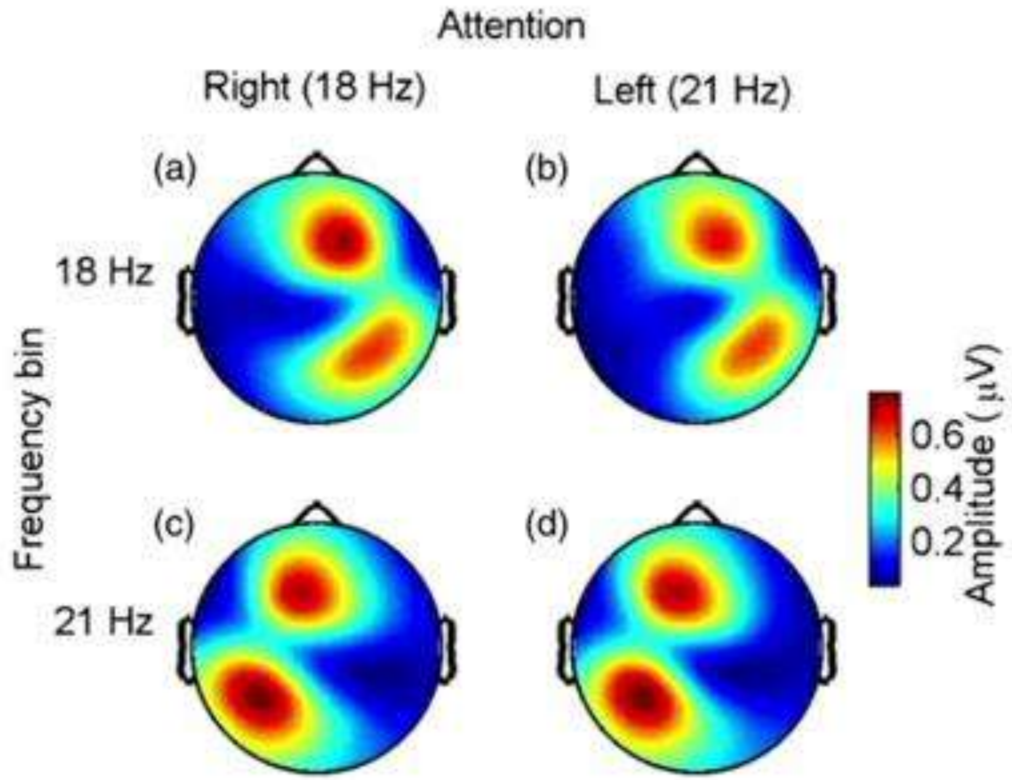


[6]

Literature review: decode tactile attention

- Steady-state somatosensory evoked potential (SSSEP)
  - Vibration at fingers
  - Power increase at the modulation frequency of the attended vibration





[5]

Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

Interactive task-based paradigm

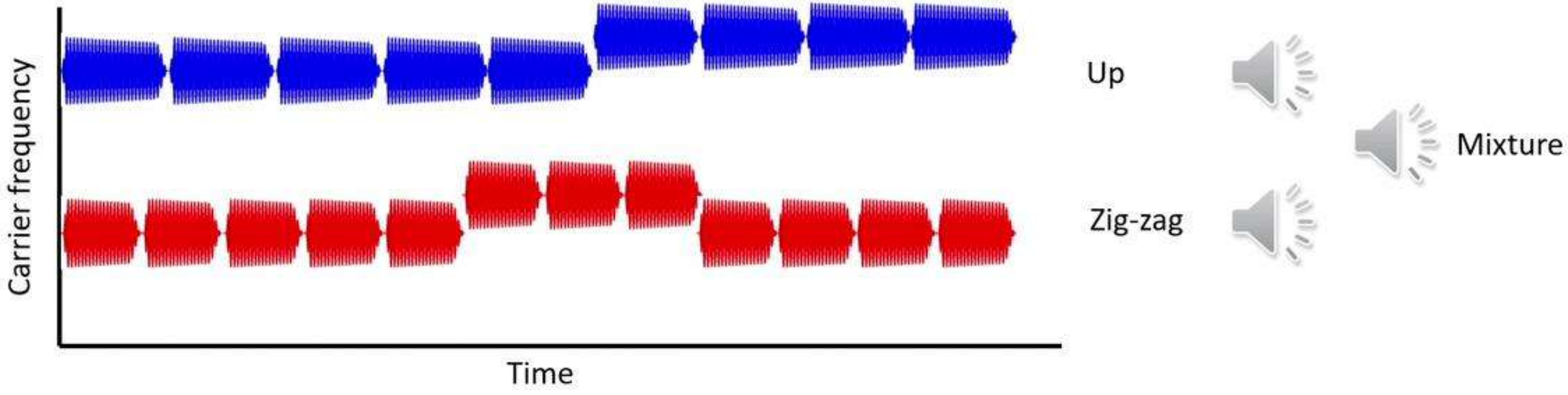
- Require a response from the subject during the experiment, and provide a real-time feedback
  - Keep the subjects engaged
- Embed an interactive task in the stimuli
  - Show subjects their behavioral performance after each session
  - More complex stimuli features
  - Easy to focus

Stimuli & experimental design: general

- Use **modulated signal** to create a stream of musical notes or vibration pulses
- The carrier frequency changes in the middle of a stream, **forming a pattern**.
- **Spatialize** the sound and vibration (left & right ear/wrist)
- Instruct the subject to **focus on one stream** using a visual cue.
- The subject **responds** with the pattern of the attended stream.

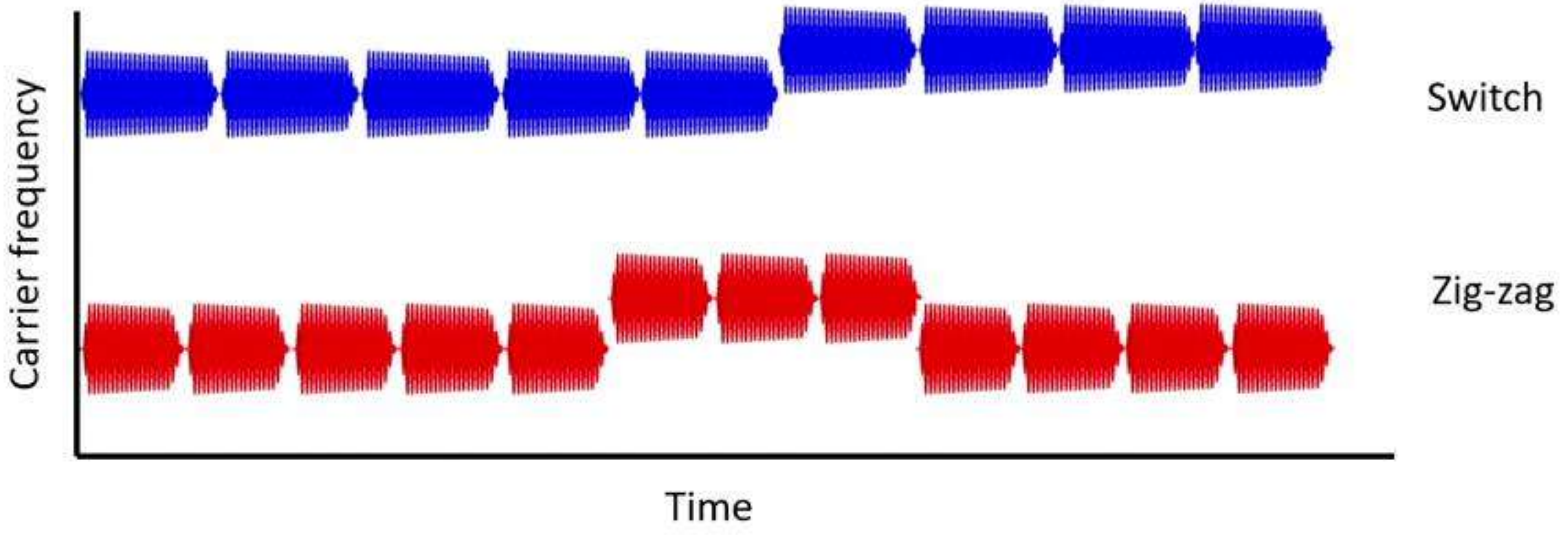
Stimuli: auditory

Channel	Modulation frequency (Hz)	Carrier frequency (Hz)			Length (ms)	Repetition	Pattern
		Low	Standard	High			
Left stream	37	703	740	777	400	9	Up/down /zig-zag
Right stream	44	396	440	484	300	12	



Stimuli: tactile

Channel	Modulation frequency (Hz)	Carrier frequency (Hz)		Length (ms)	Repetition	Pattern
		Standard	Oddball			
Left stream	27	120	210	400	9	Switch / zig-zag
Right stream	17	120	210	300	12	

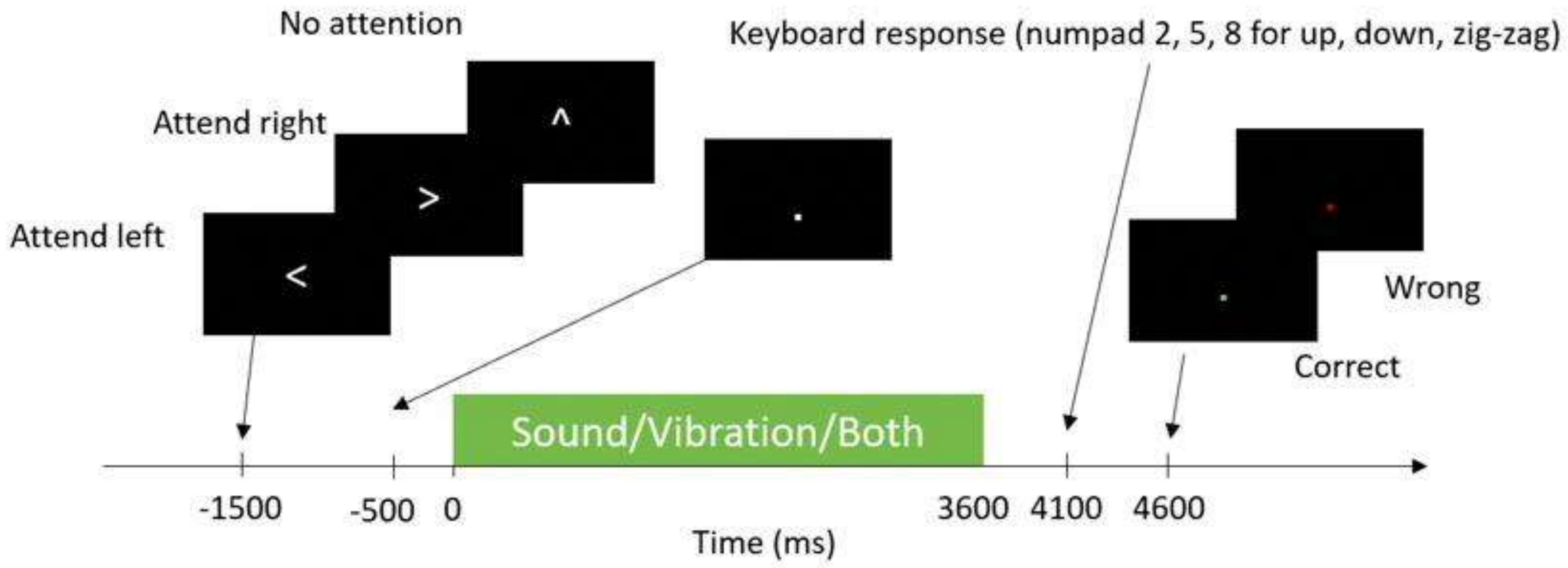


Stimuli: auditory + tactile (mixed)

- Use the same stimuli as in auditory or tactile alone conditions
- The onsets of notes (for auditory) and vibration (for tactile) are synchronized.
- Identify the pattern in the auditory stream (up/down/zig-zag)

Experiment

- 3 modalities (auditory, tactile, mixed)
- 3 conditions per modality (attend left, attend right, no attention)
- 24 trials per condition (randomized)
- ~8 seconds per trial
- ~35 min in total



Possible neural signatures

- Spatial attention
  - Parietal alpha power increase
- Auditory steady-state response
  - Power increased at the modulation frequency of the attended sound
- Somatosensory steady-state response
  - Power increase at the modulation frequency of the attended vibration
- P300 response time-locked to the oddball onset and offset

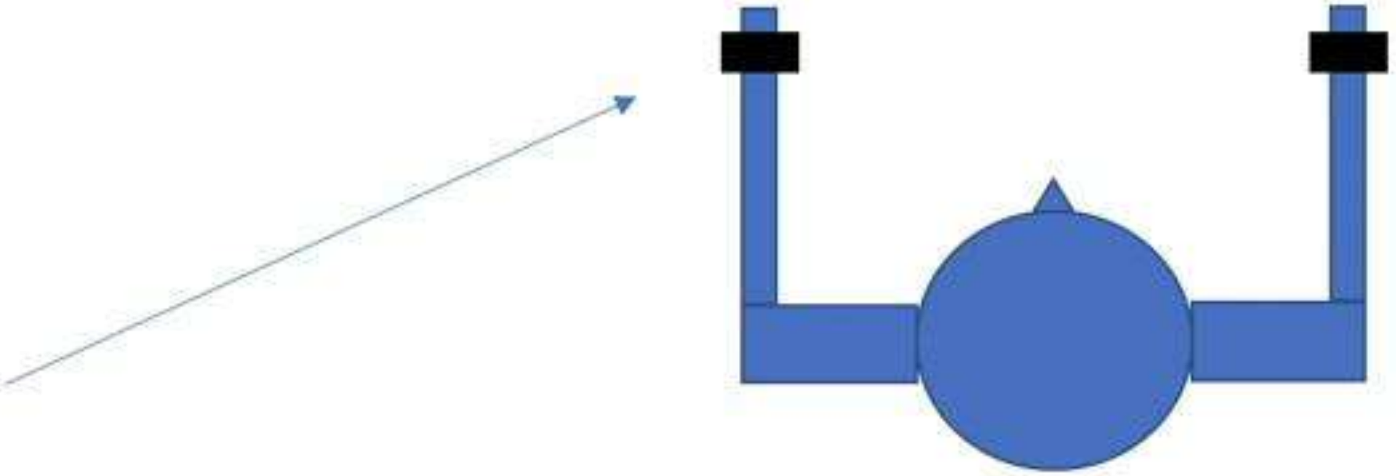
Equipment – EEG systems

- EEG
  - mBrainTrain Smarting
  - 24-channel, gel-based
  - 500Hz

- In-ear EEG
  - Made in-house
  - 2-channel, conductive cloth
  - 250Hz

Equipment – vibrotactile actuator

- Dayton coin-type audio speaker
  - Left and right wrists
  - Free the users' hands for tasks



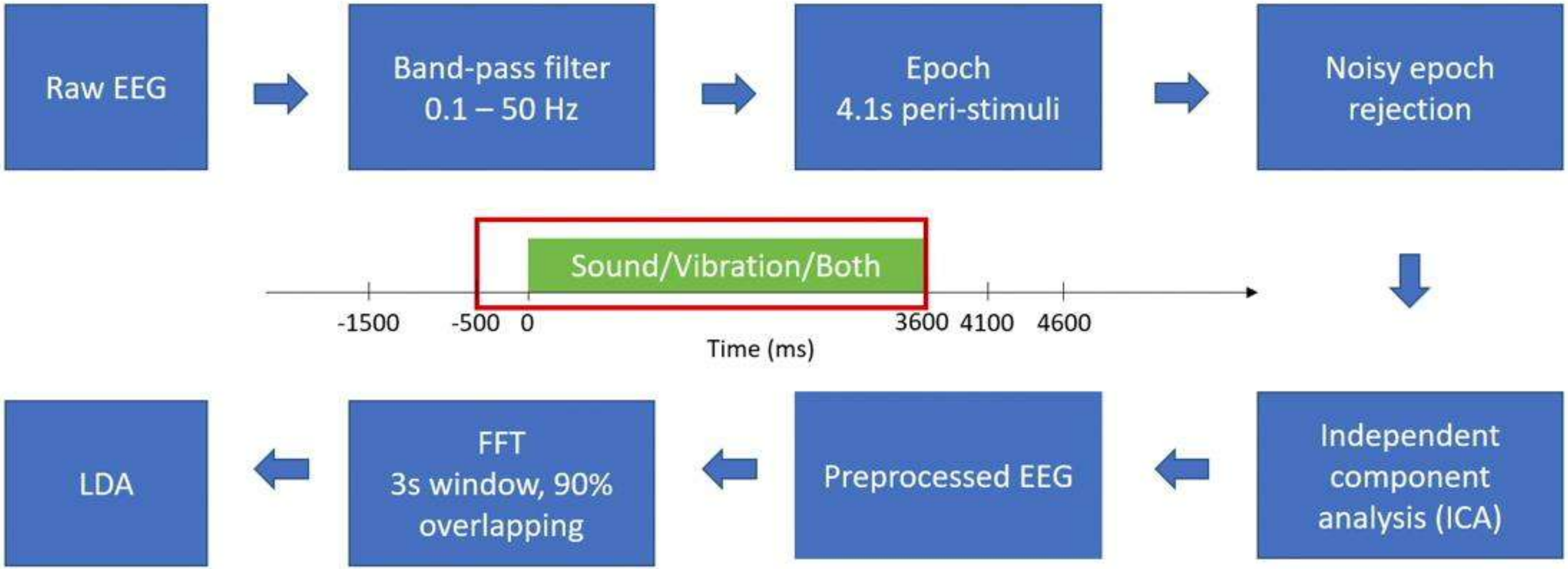
Data collection

- 12 subjects (3 female, 9 male)
- Age: 32.3 ± 7.4 years old
- 2 with previous BCI experience
- 1 left handed

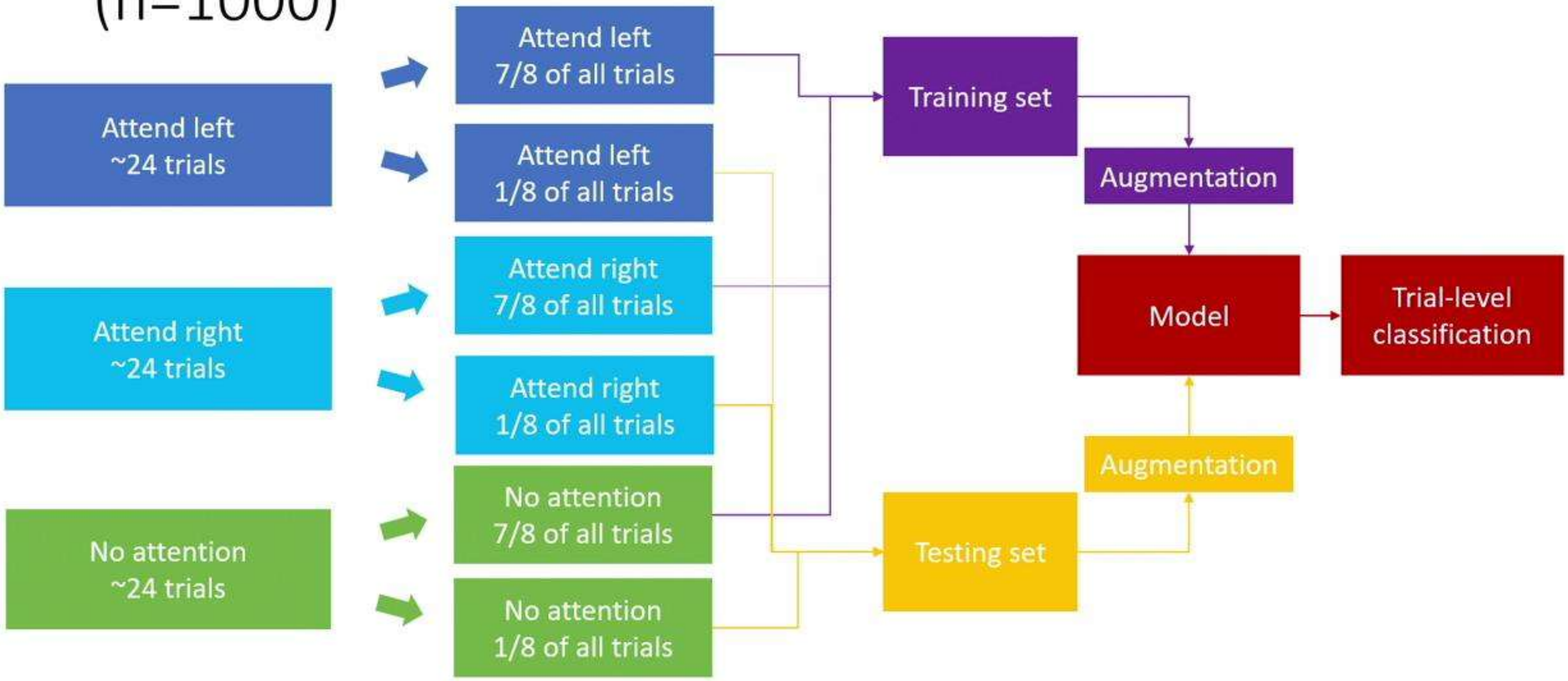
Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

Surface EEG processing and analysis pipeline



3-way classification with 8-fold cross-validation (n=1000)

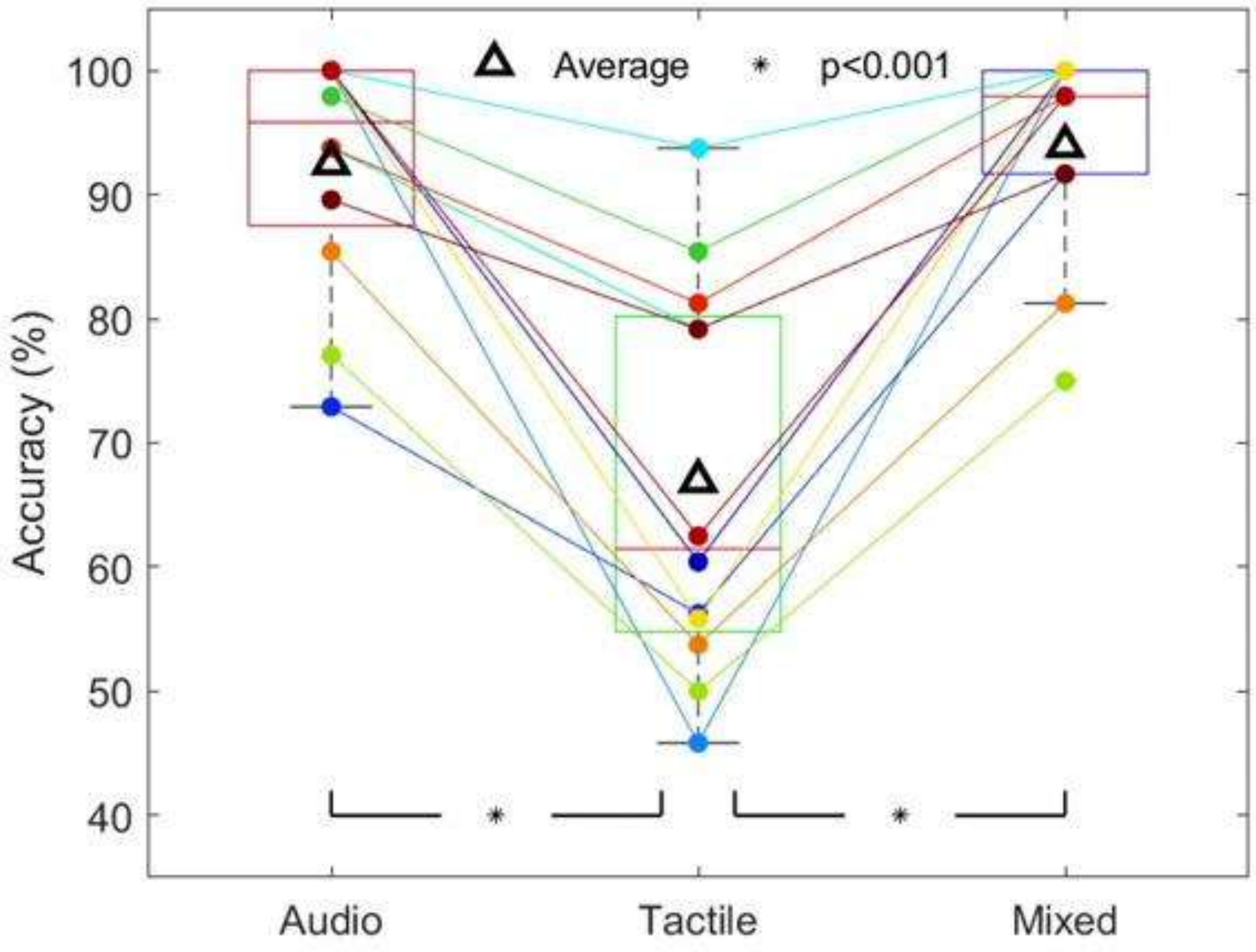


Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

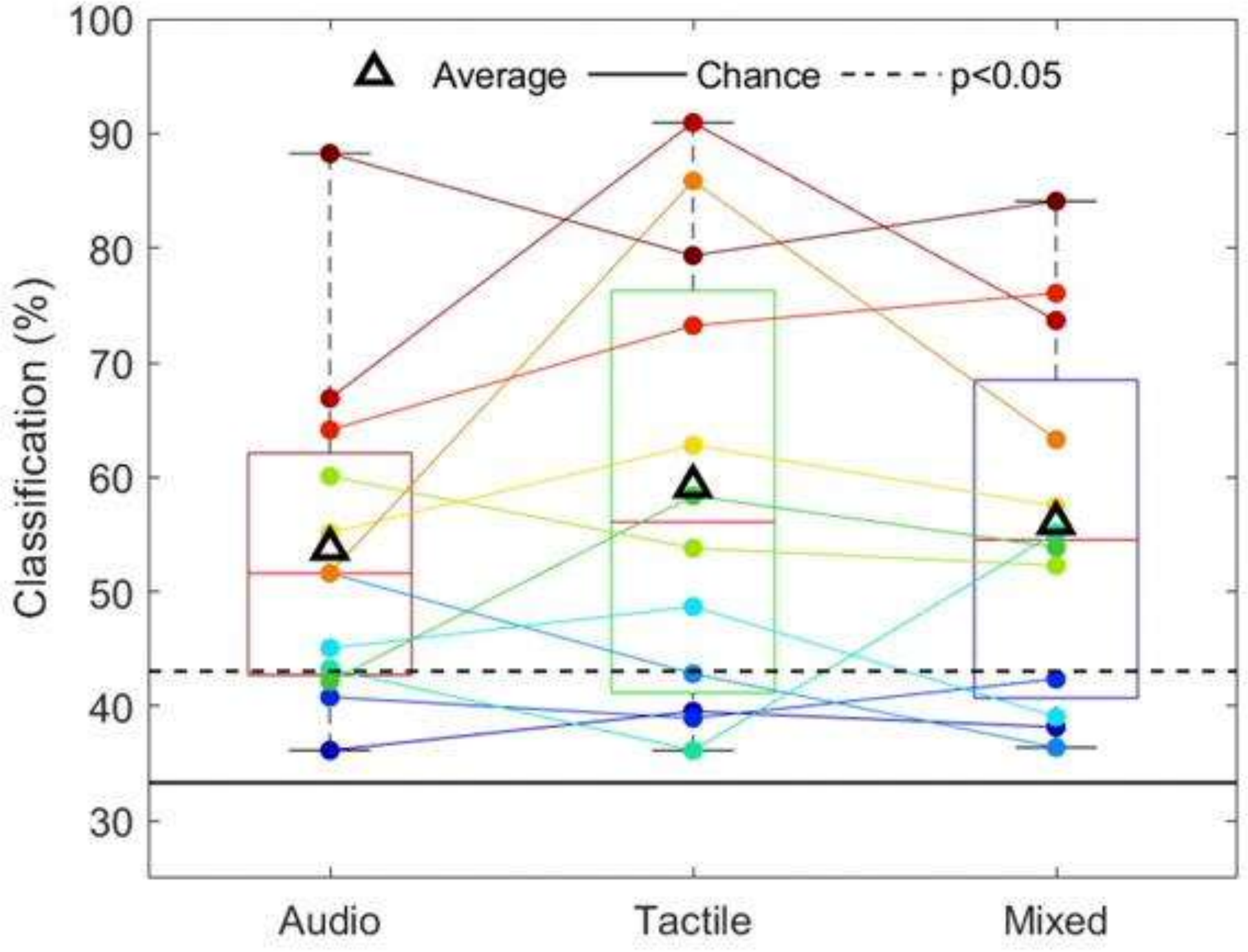
Behavioral performance

- Percent correct of the pattern identification task
- Averaged between attending to left and attending to right tasks
- Each line represents a subject



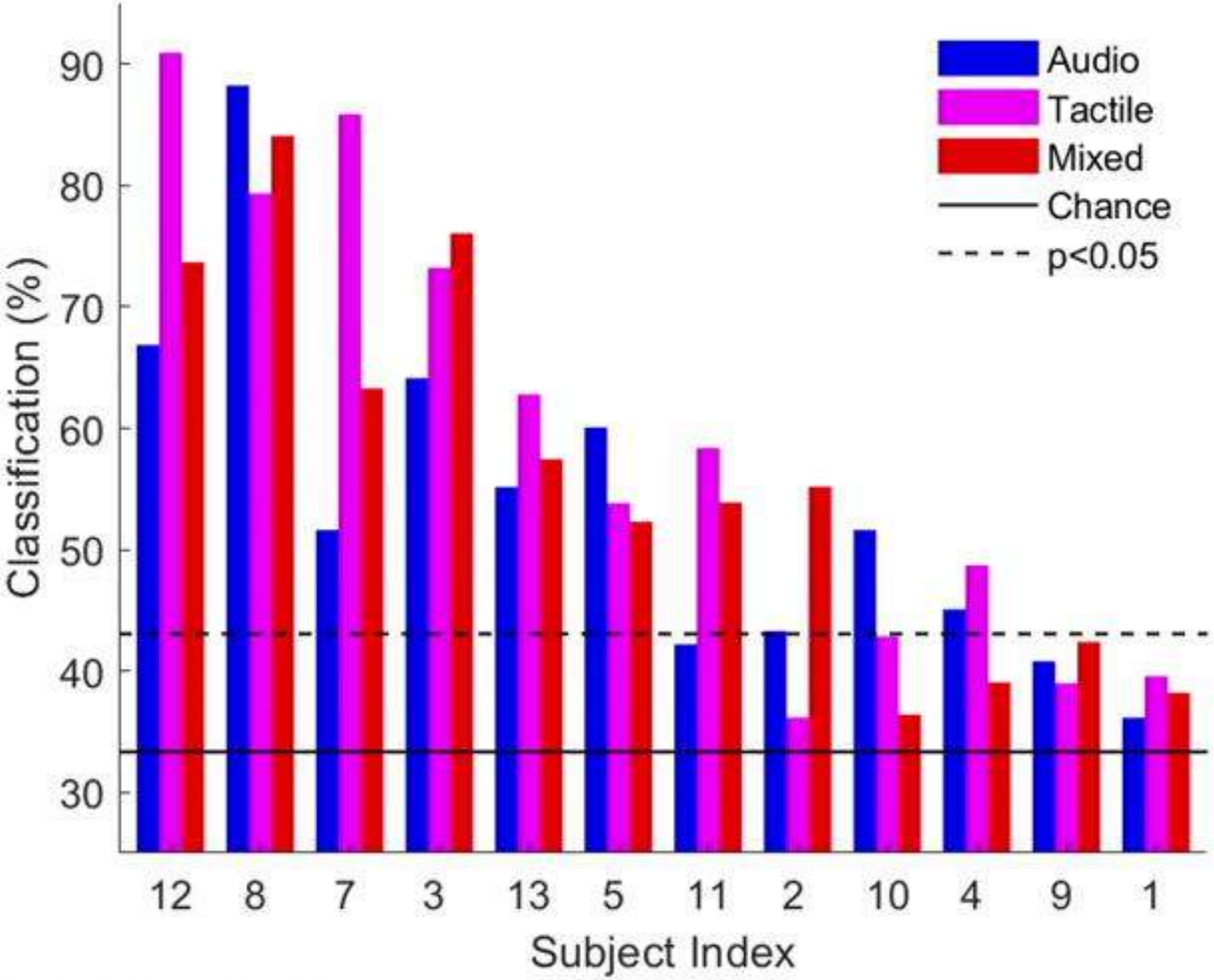
Surface EEG decoding results by modality

Modality	Mean ± Std (%)	> 33.3% (n)	> 43.1% (n)	ITR (bits/min)
Audio	53.8 ± 13.9	12	9	<1 – 14.27
Tactile	59.2 ± 18.4	12	8	<1 – 16.14
Mixed	56.0 ± 15.1	12	8	<1 – 12.07

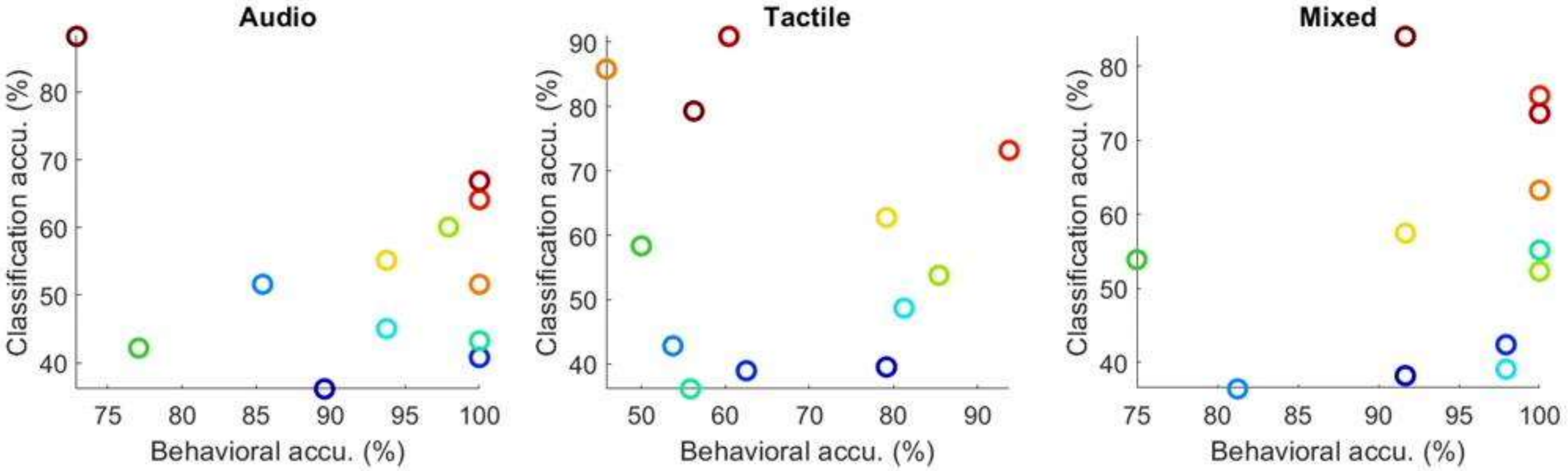


Surface EEG decoding results by subject

Modality	Mean $\pm$ Std (%)	> 33.3% (n)	> 43.1% (n)	ITR (bits/min)
Audio	53.8 ± 13.9	12	9	<1 – 14.27
Tactile	59.2 ± 18.4	12	8	<1 – 16.14
Mixed	56.0 ± 15.1	12	8	<1 – 12.07



Correlation with behavioral performance



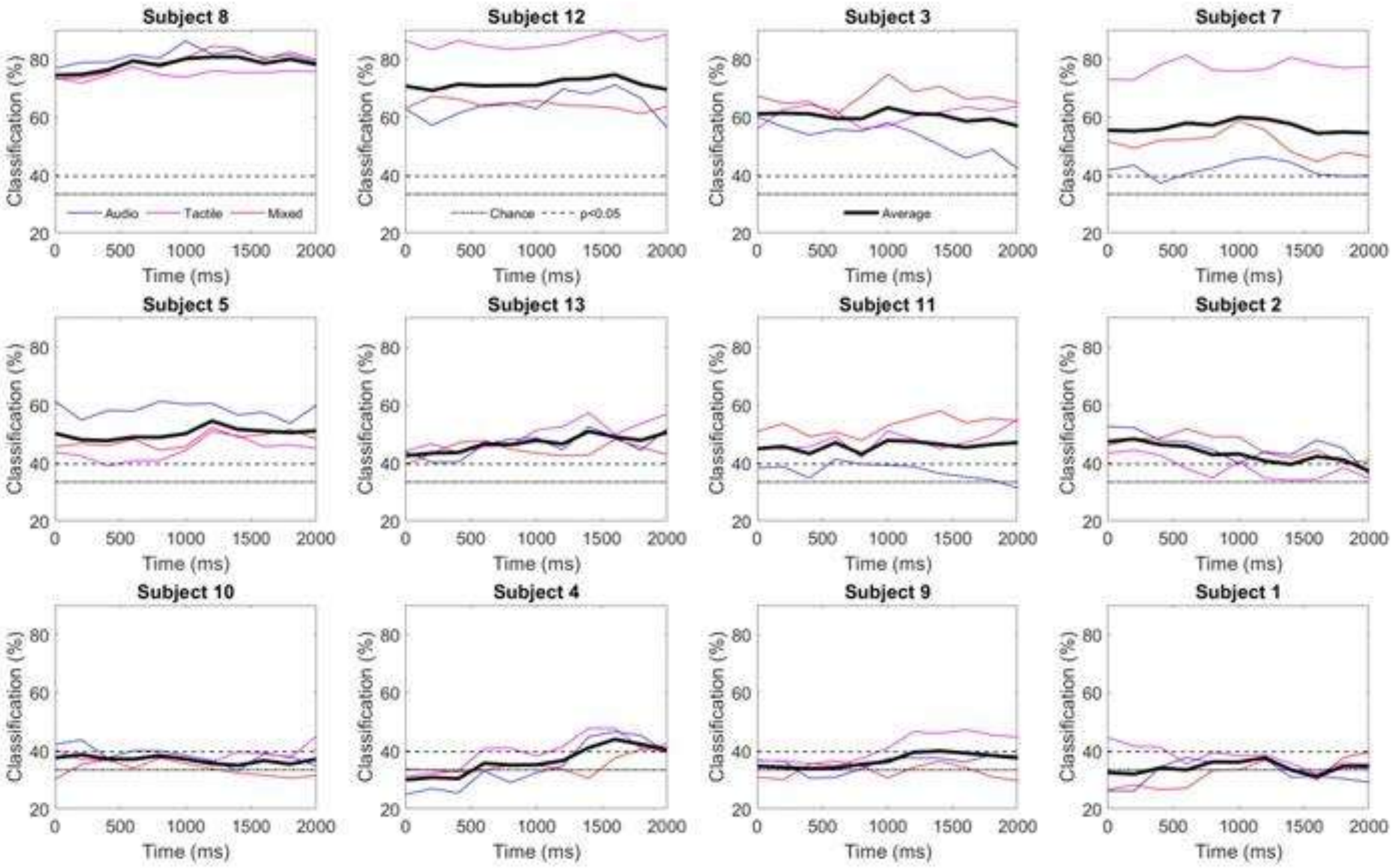
Summary I

- The tactile attention task is more difficult than the auditory and the mixed task.
- The attention decoding results vary across subjects, but are consistent across modalities within each subject.
- The relationship between behavioral accuracy and the classification accuracy seems to be unpredictable.

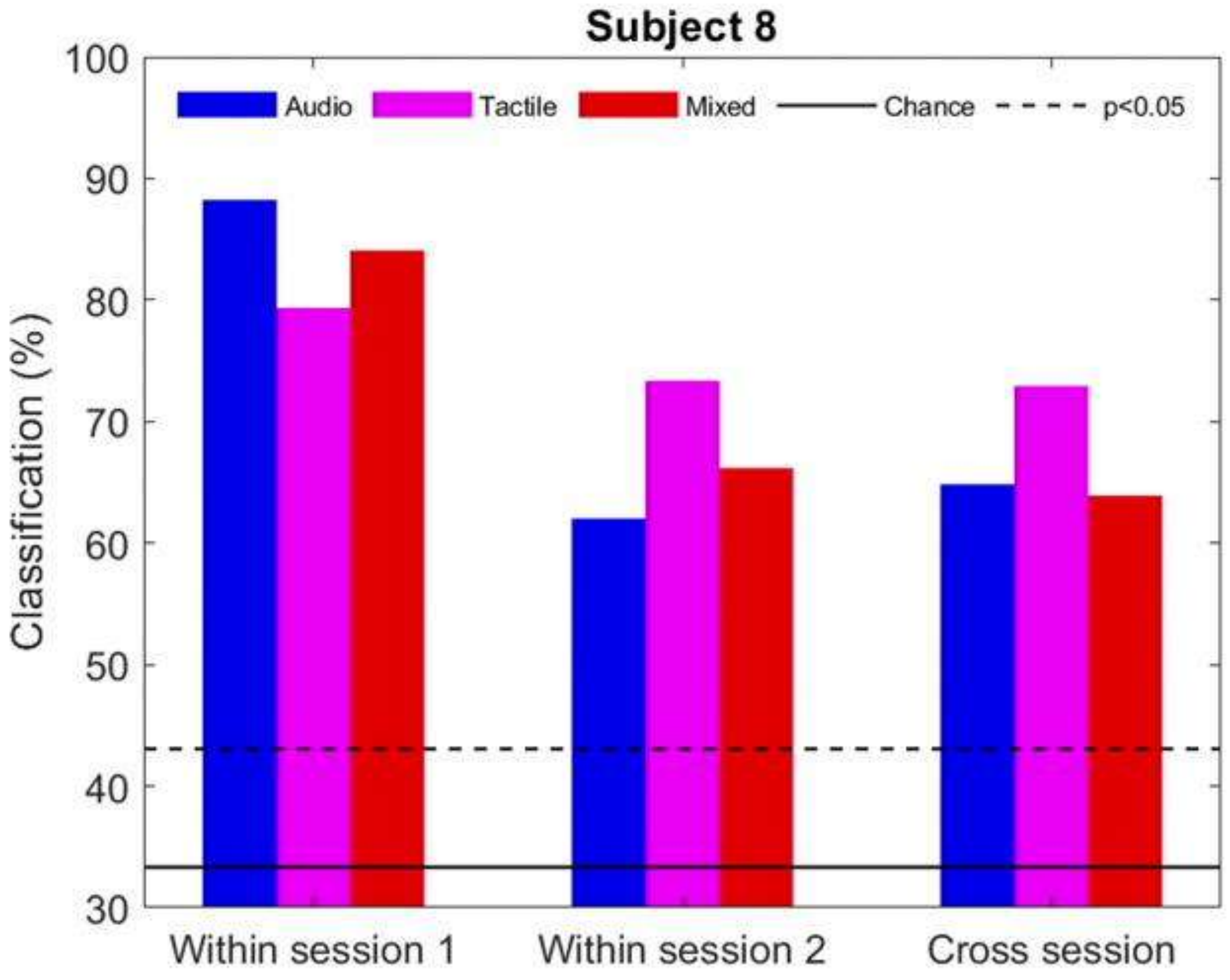
Decoding over time

- 2 second window
- 90% overlap
- Average over all the samples at the same time window

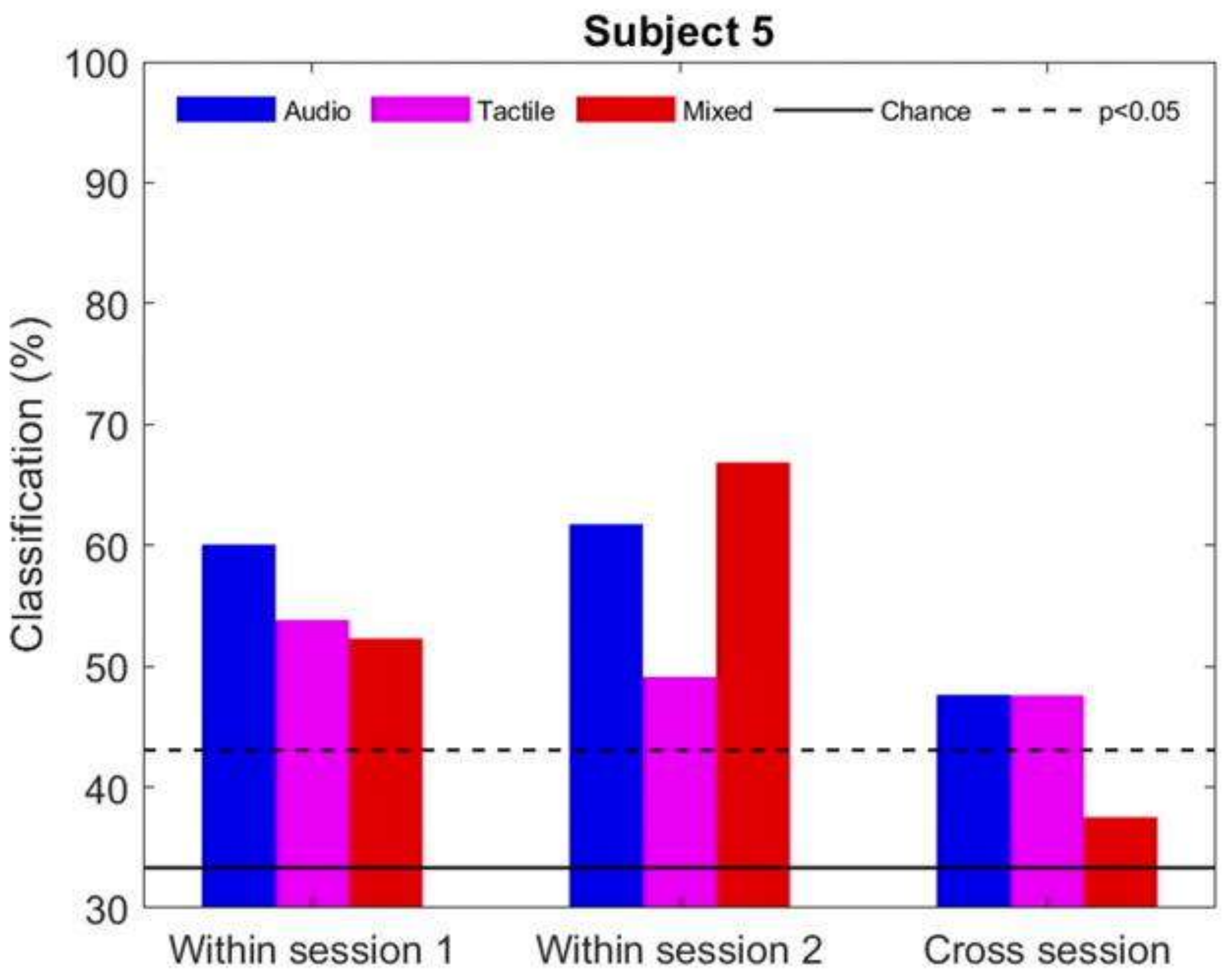
Audio Tactile Mixed
Chance p<0.05 Average



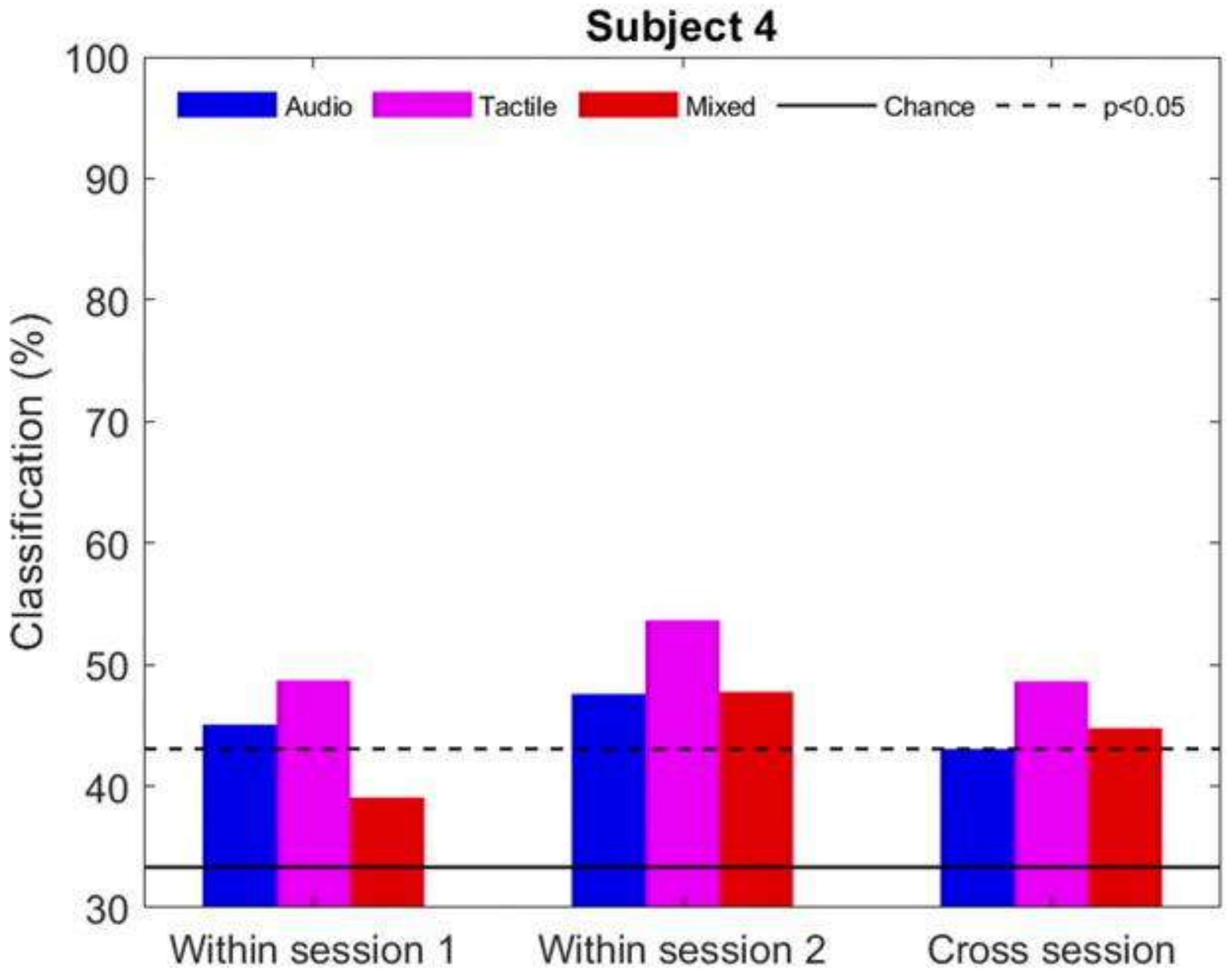
Cross-session validation: a highly classifiable case



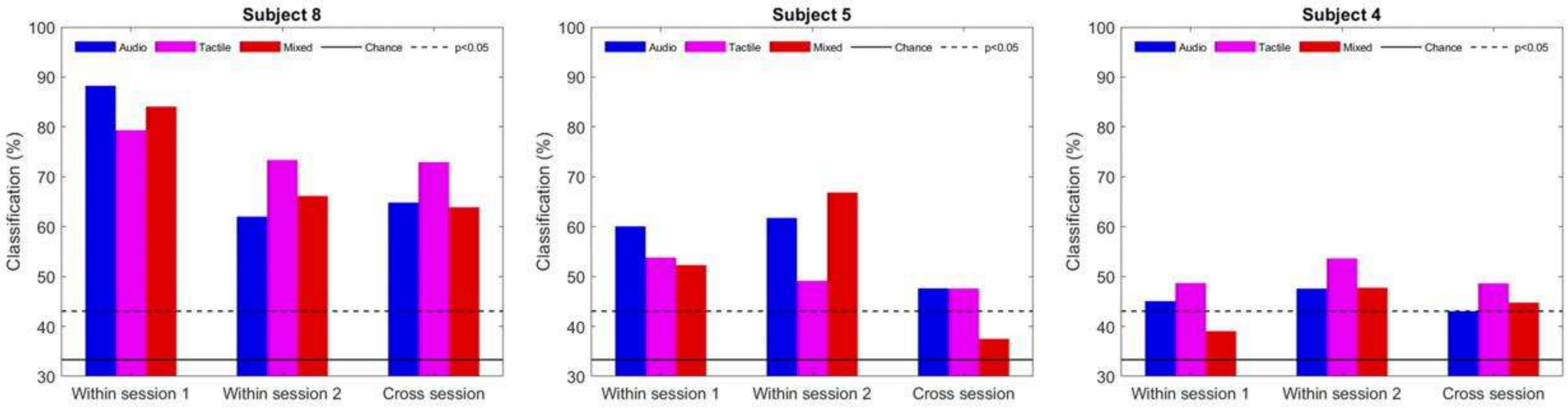
Cross-session validation: a classifiable case



Cross-session validation: a barely classifiable case



Cross-session validation: overall



Summary II

- Decoding accuracy is stable within a trial for most subjects.
  - Sustained attention

- The trained model is robust across multi-day sessions.
  - The model trained today can be used tomorrow to achieve similar results.

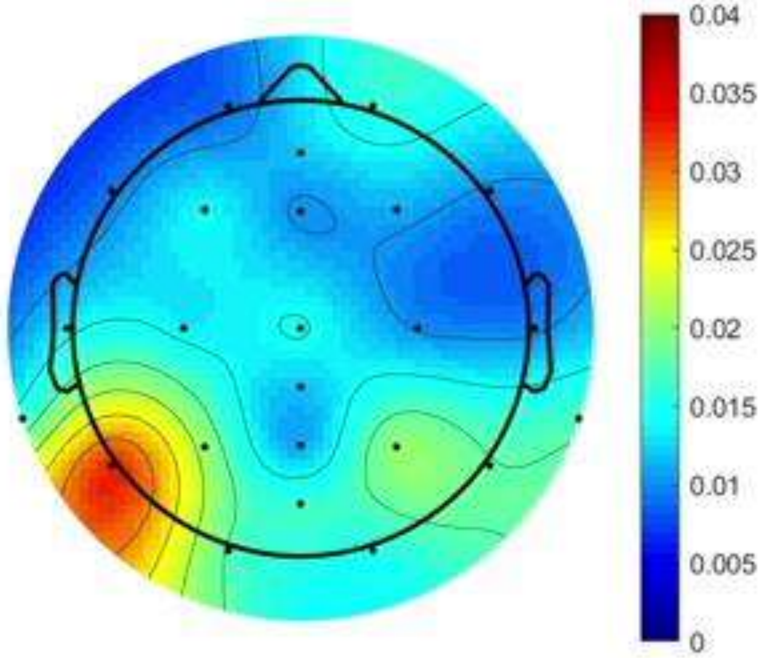
- The performance of the classifier is consistent across multi-day sessions within a subject.

Feature weight

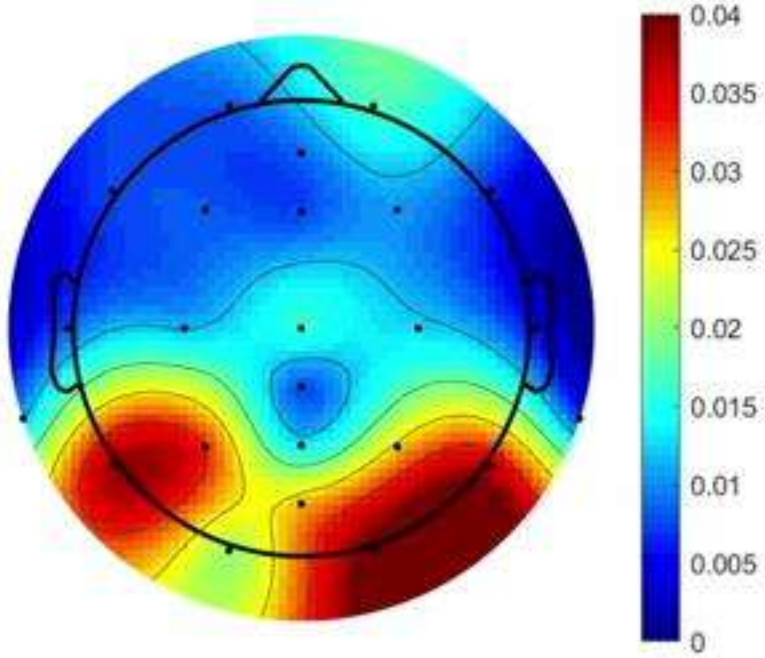
- Neighborhood component analysis (NCA)
  - Feature selection algorithm

Subject 12

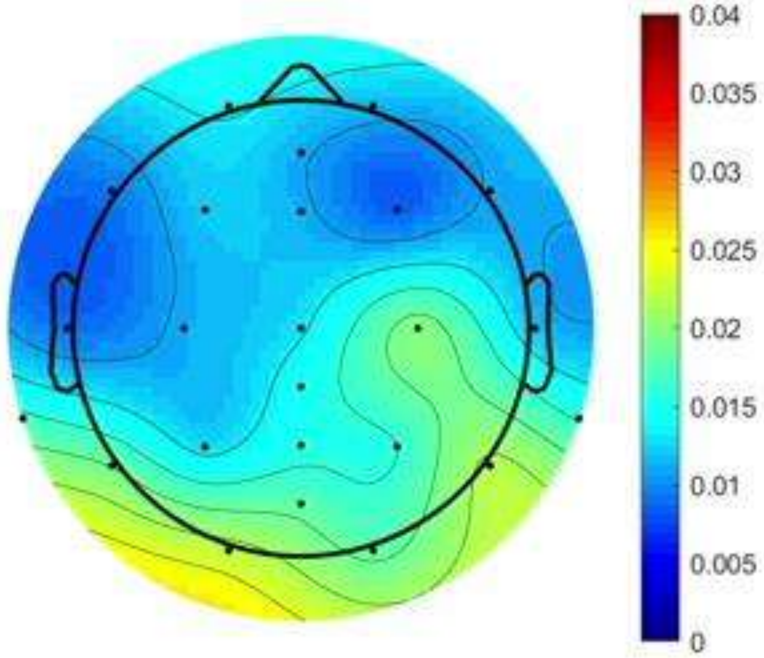
Audio - alpha



Tactile - alpha

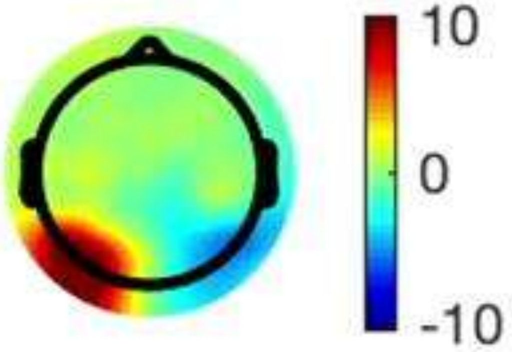


Mixed - alpha



Feature weight

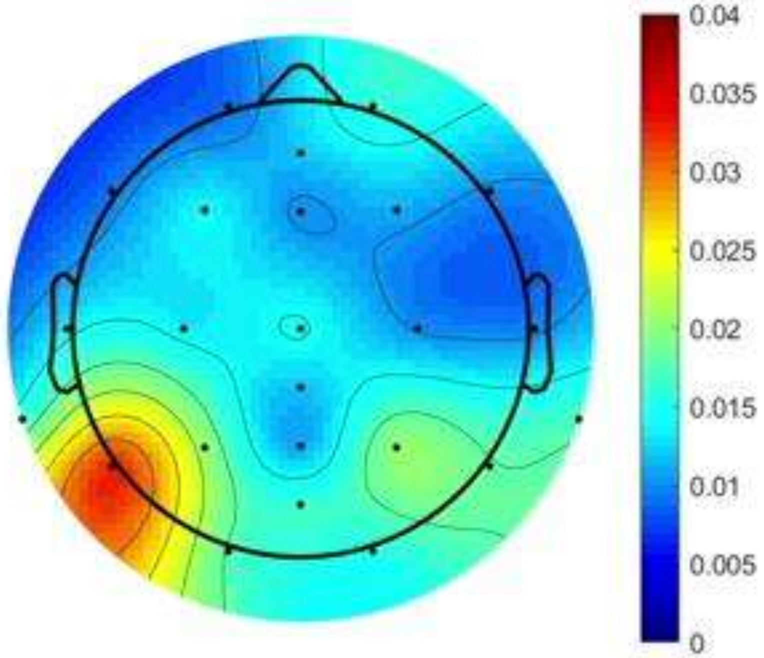
Attend left – Attend right



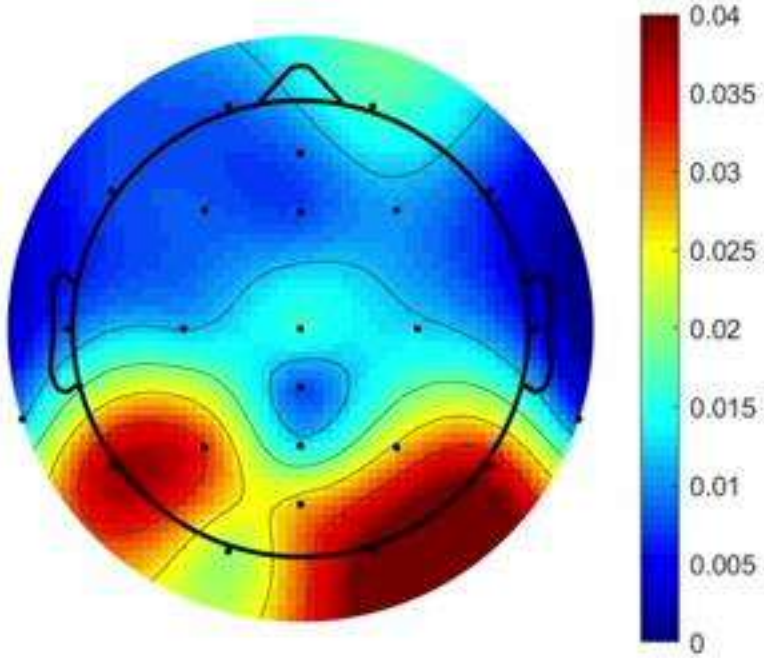
- Neighborhood component analysis (NCA)
  - Feature selection algorithm

Subject 12

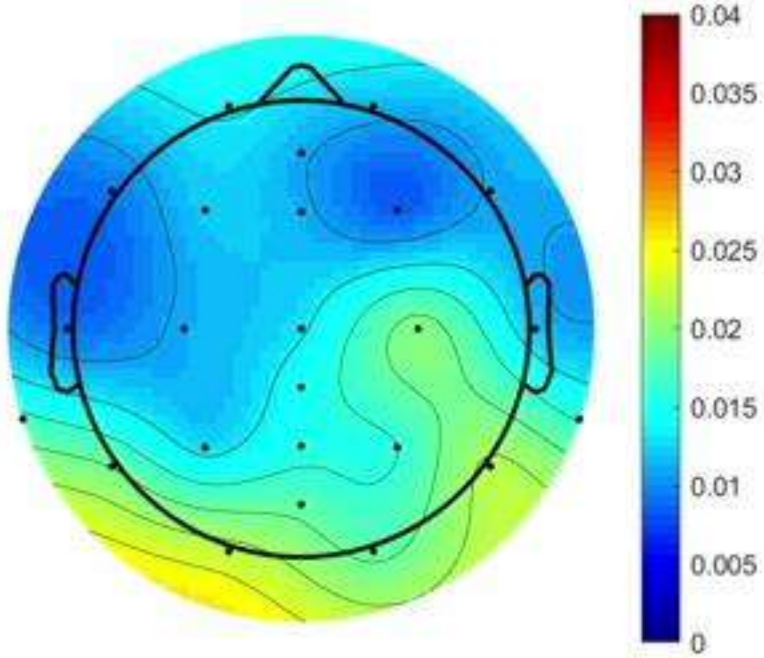
Audio - alpha



Tactile - alpha



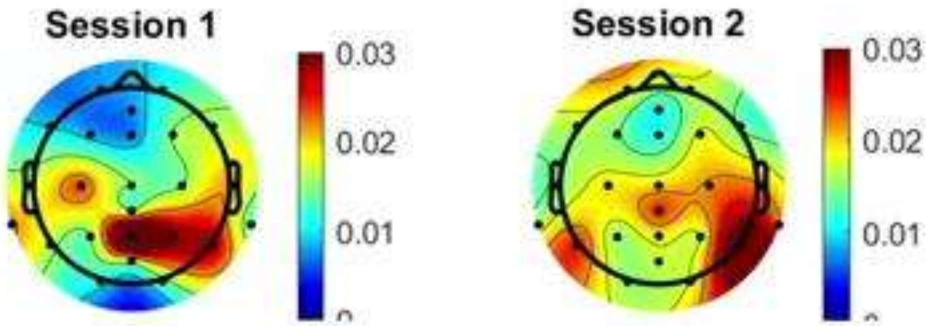
Mixed - alpha



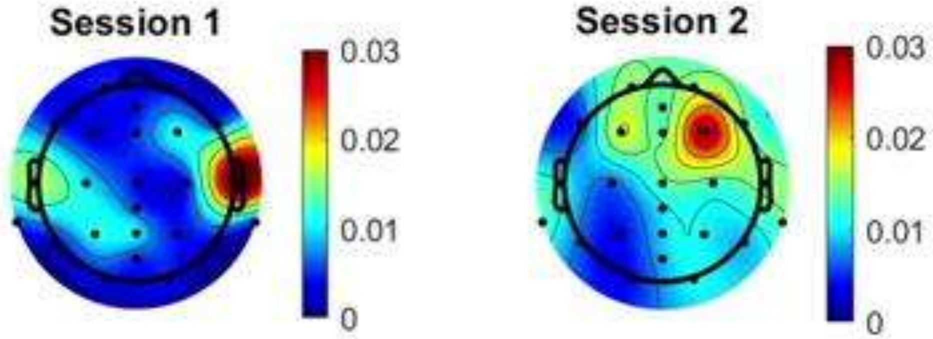
Feature weight consistency between sessions: audio & tactile

Subject 8

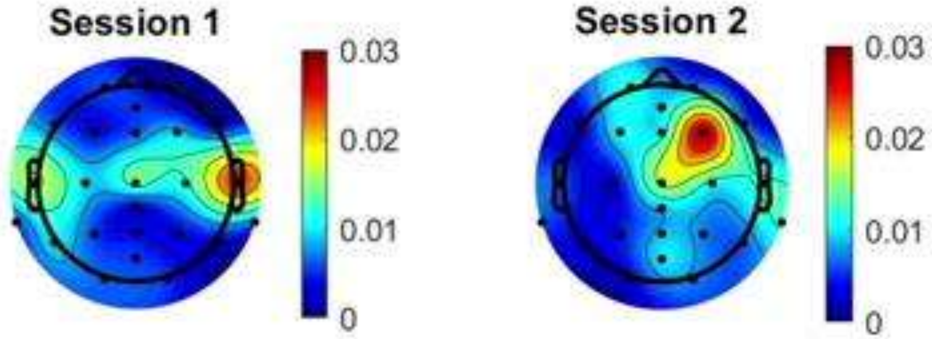
Audio – alpha band



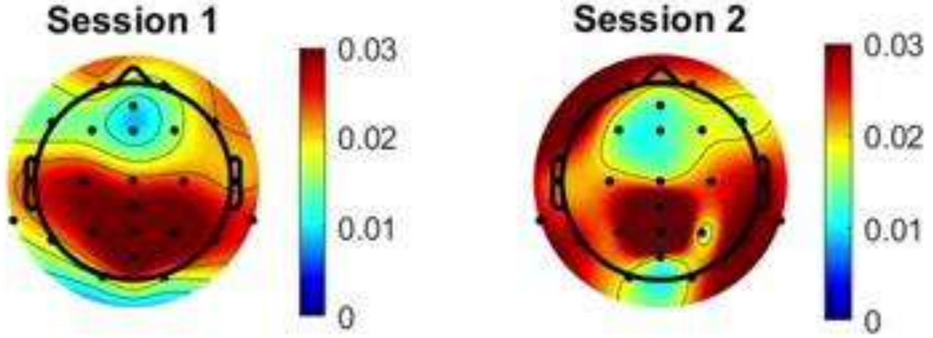
Audio – 37Hz



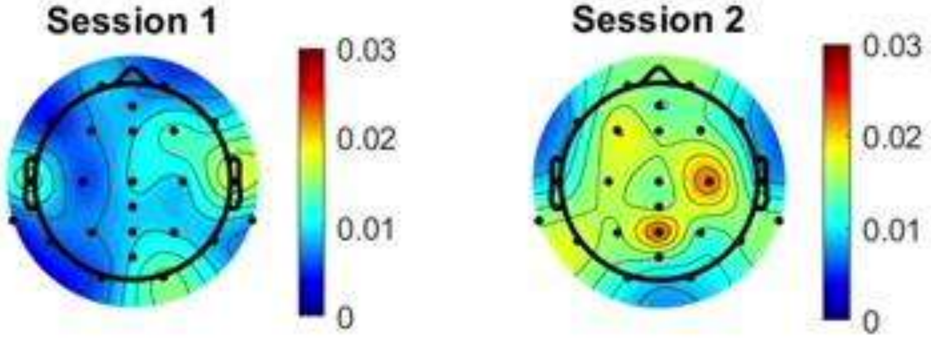
Audio – 44Hz



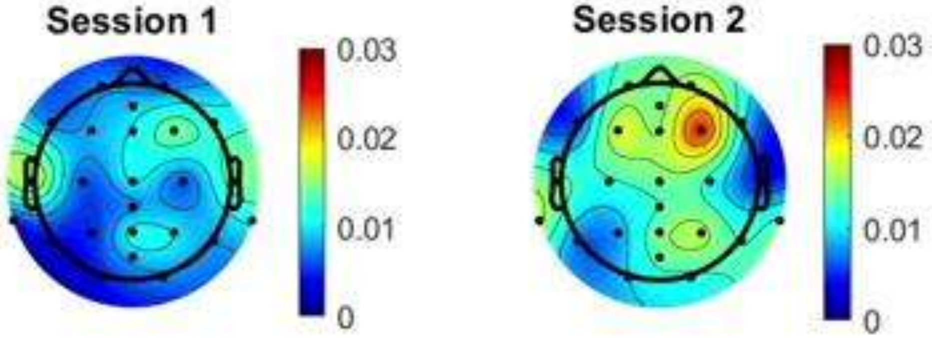
Tactile – alpha band



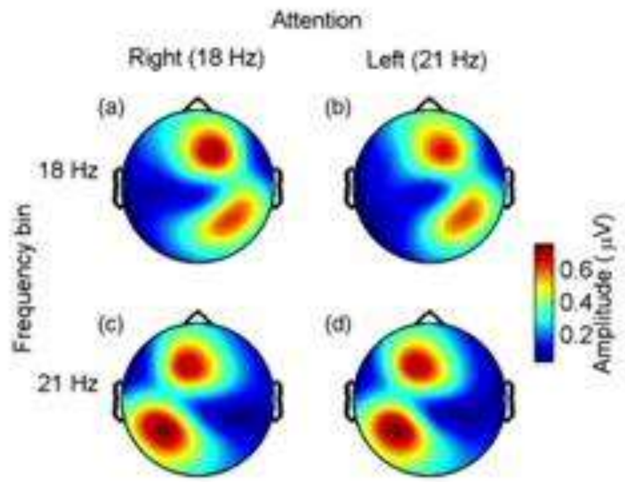
Tactile – 17Hz



Tactile – 27Hz

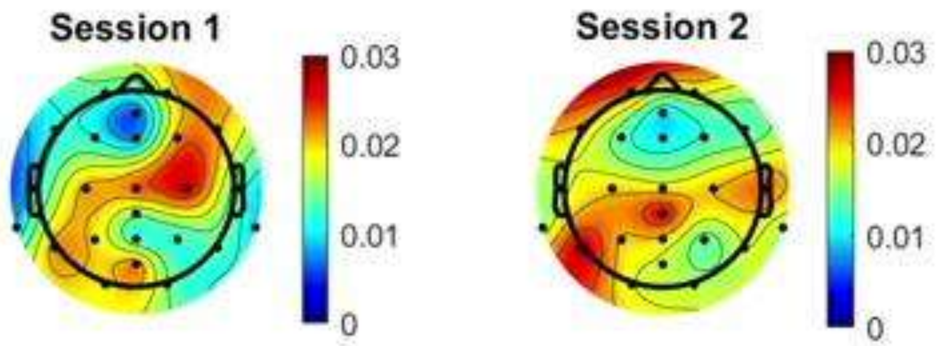


Feature weight consistency between sessions: mixed

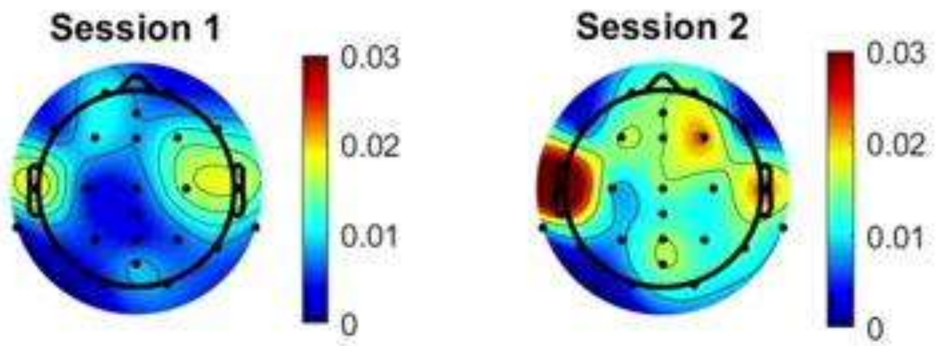


Subject 8

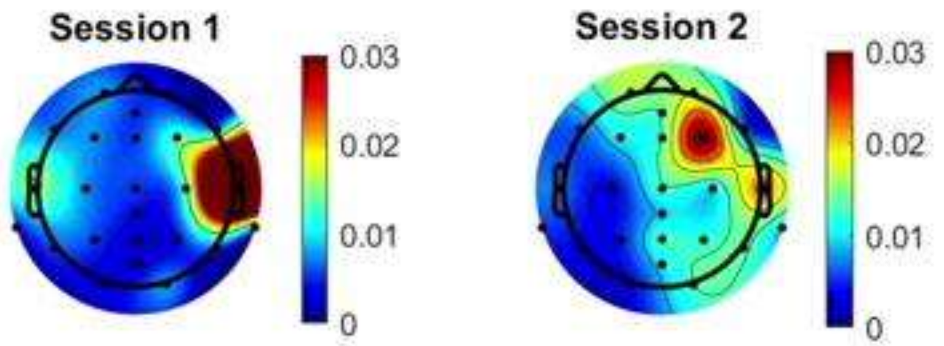
Mixed – alpha band



Mixed – 27Hz



Mixed – 37Hz

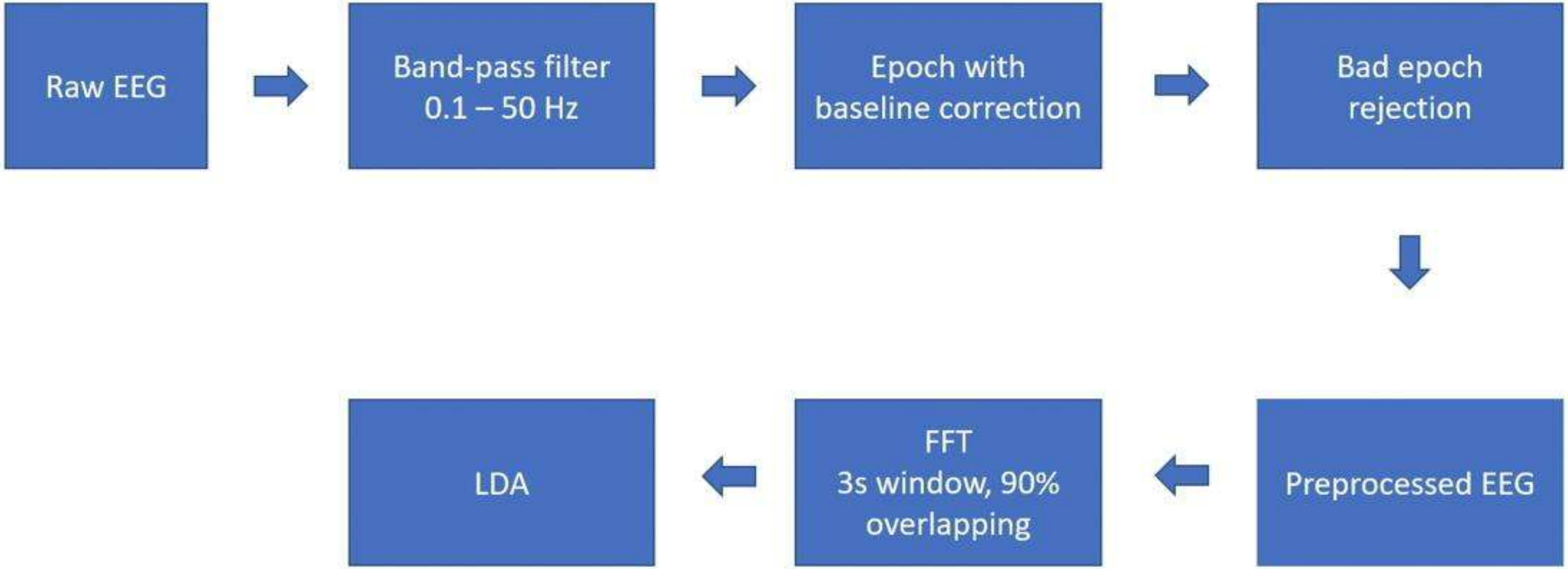


Summary III

- The trained model is decoding attention based on some neurologically relevant features.

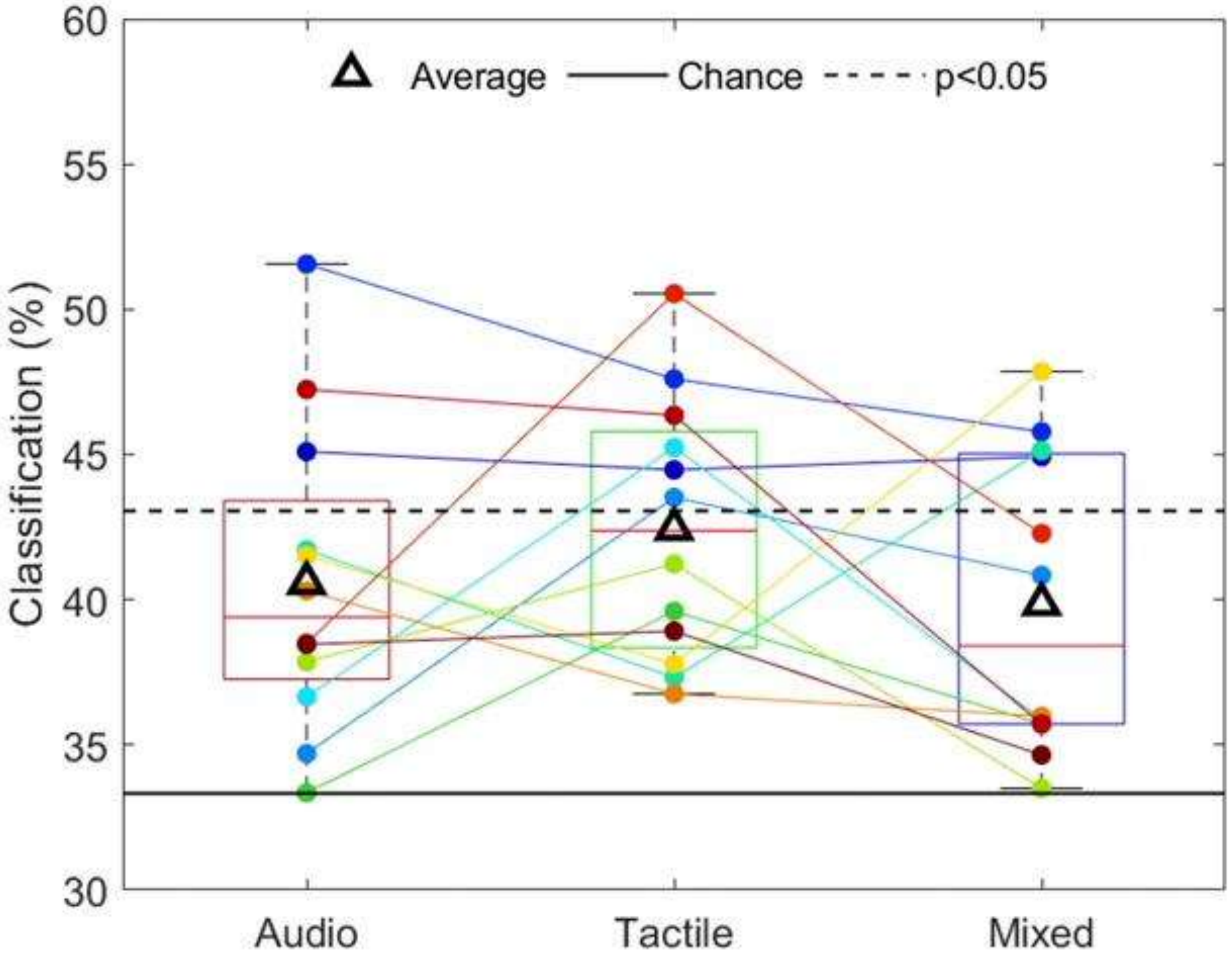
- The feature weights are similar across multi-day sessions within individual subjects.

In-ear EEG processing and analysis pipeline

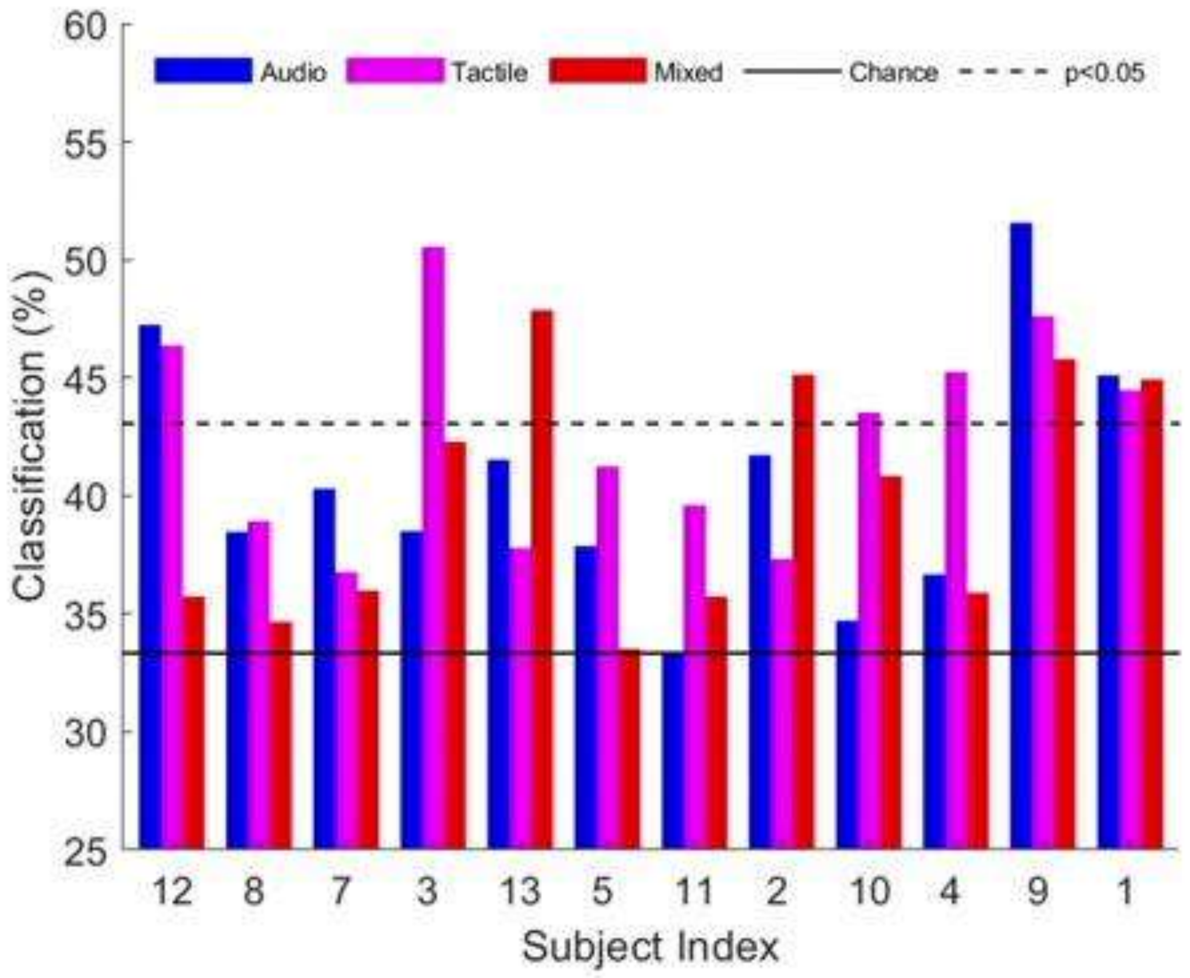


In-ear EEG decoding results by modality

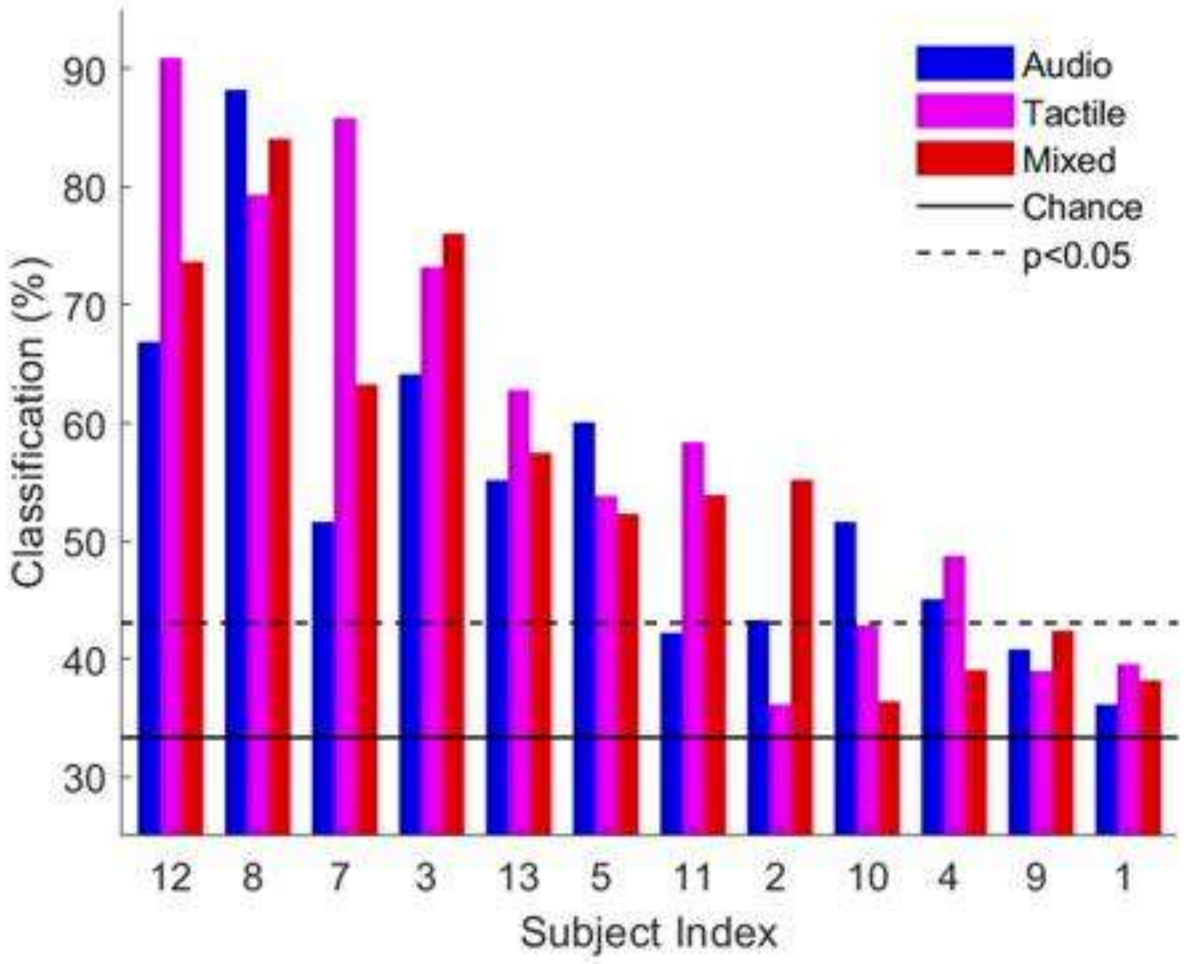
- LDA, 3-way classification
- 1000 cross-validation



In-ear EEG decoding results by subject

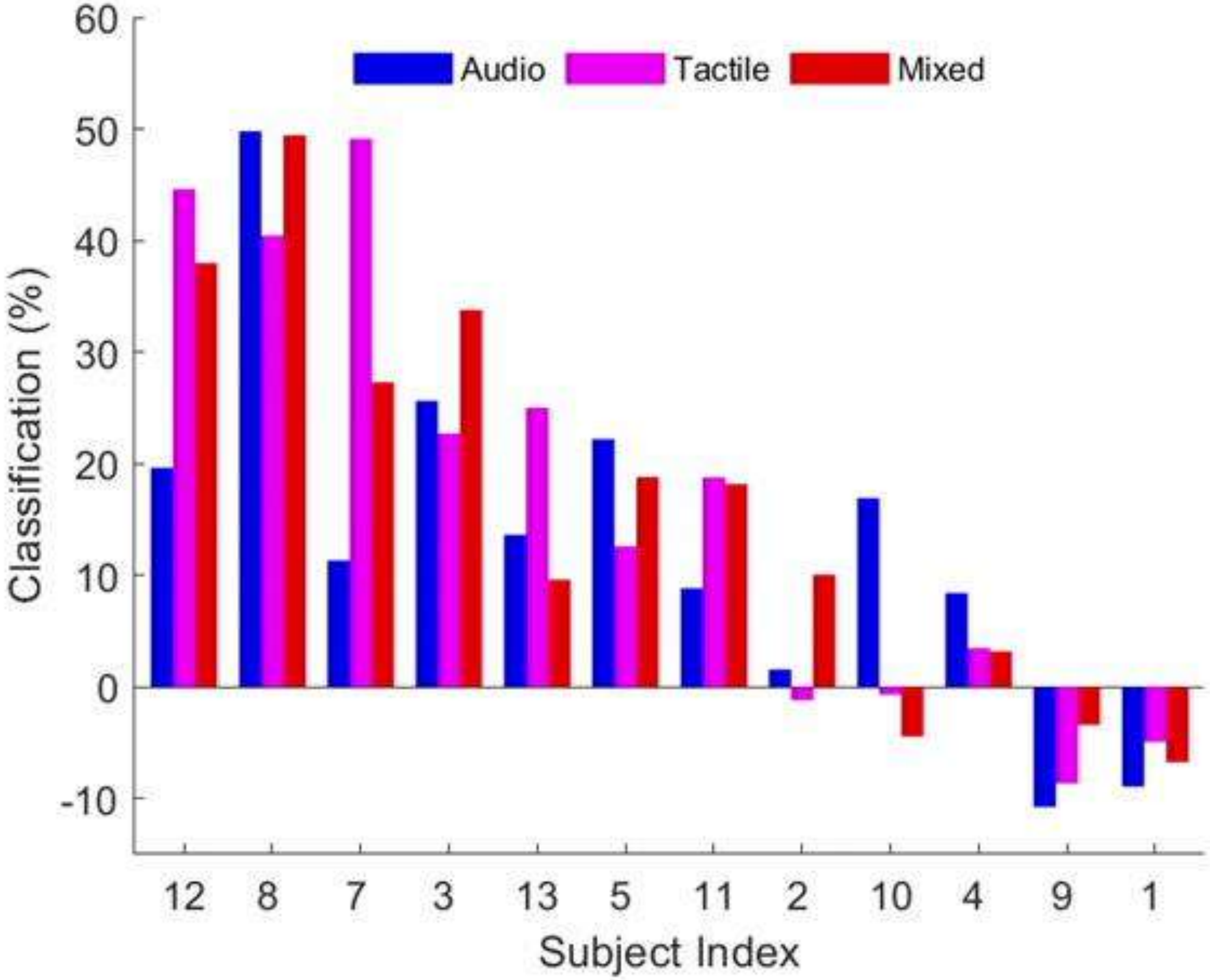


In-ear EEG



Surface EEG

Difference between surface and in-ear EEG decoding accuracy by subject



Summary IV

- The classification accuracy of in-ear EEG is generally lower than that of surface EEG.
- The classification accuracy of in-ear EEG seem to be negatively correlated with the accuracy of surface EEG across subjects.
  - Might be explained by the neural anatomy of individual subjects

Agenda

- Introduction
  - Electroencephalography (EEG)
  - Brain-computer interface (BCI)
- Research objectives & literature review
- Experiment & equipment
- Signal processing and data analysis
- Results
- Conclusions & future works

Conclusions

- Integrate auditory and tactile stimuli into an interactive task-based paradigm is a feasible way to build a functional BCI system.

- The performance of the proposed BCI system
  - outperforms previous studies using steady stimuli.
  - is comparable across sensory modalities.
  - varies across subjects.
  - is robust over time within each subject.

- In-ear EEG may be able to capture some information missed by the surface EEG, which has the potential to be used in a customized BCI system.

Future works

- Feature dimensionality reduction
  - Feature selection
- Use neural network to decode attention
  - Preliminary results show that there is a potential gain in classification accuracy, but the amount of gain varies across subjects.
- Integrate of spatial information into classification
- In-depth analysis on individual differences
- Improve task/game design and stimuli

Acknowledgement

Hannes Gamper

Adrian KC Lee

Nicholas Huang

Hakim Si Mohammed

Ivan Tashev

David Johnson

Sebastian Braun

Dimitra Emmanouilidou

Becky Gagnon

John Romualdez

Patrick Therien

SK Neang

Teresa LaScala

Todd Jurgensen

Christian Holz

Andy Wilson

Ed Cutrell

Mihai Jalobeanu

Mike Sinclair

Raymond Xia

Tamzeed Islam

Benjamin Elizalde

Fabien Brinkmann

Sahar Hashemgeloogerdi

Ana Elisa Mendez Mendez

Morayo Ogunsina

Arindam Jati

Ziqi Fan

Special thanks to Lincy Wang

Future works

- Feature dimensionality reduction
  - Feature selection
- Use neural network to decode attention
  - Preliminary results show that there is a potential gain in classification accuracy, but the amount of gain varies across subjects.
- Integrate of spatial information into classification
- In-depth analysis on individual differences
- Improve task/game design and stimuli

Acknowledgement

Hannes Gamper

Adrian KC Lee

Nicholas Huang

Hakim Si Mohammed

Ivan Tashev

David Johnson

Sebastian Braun

Dimitra Emmanouilidou

Becky Gagnon

John Romualdez

Patrick Therien

SK Neang

Teresa LaScala

Todd Jurgensen

Christian Holz

Andy Wilson

Ed Cutrell

Mihai Jalobeanu

Mike Sinclair

Raymond Xia

Tamzeed Islam

Benjamin Elizalde

Fabien Brinkmann

Sahar Hashemgeloogerdi

Ana Elisa Mendez Mendez

Morayo Ogunsina

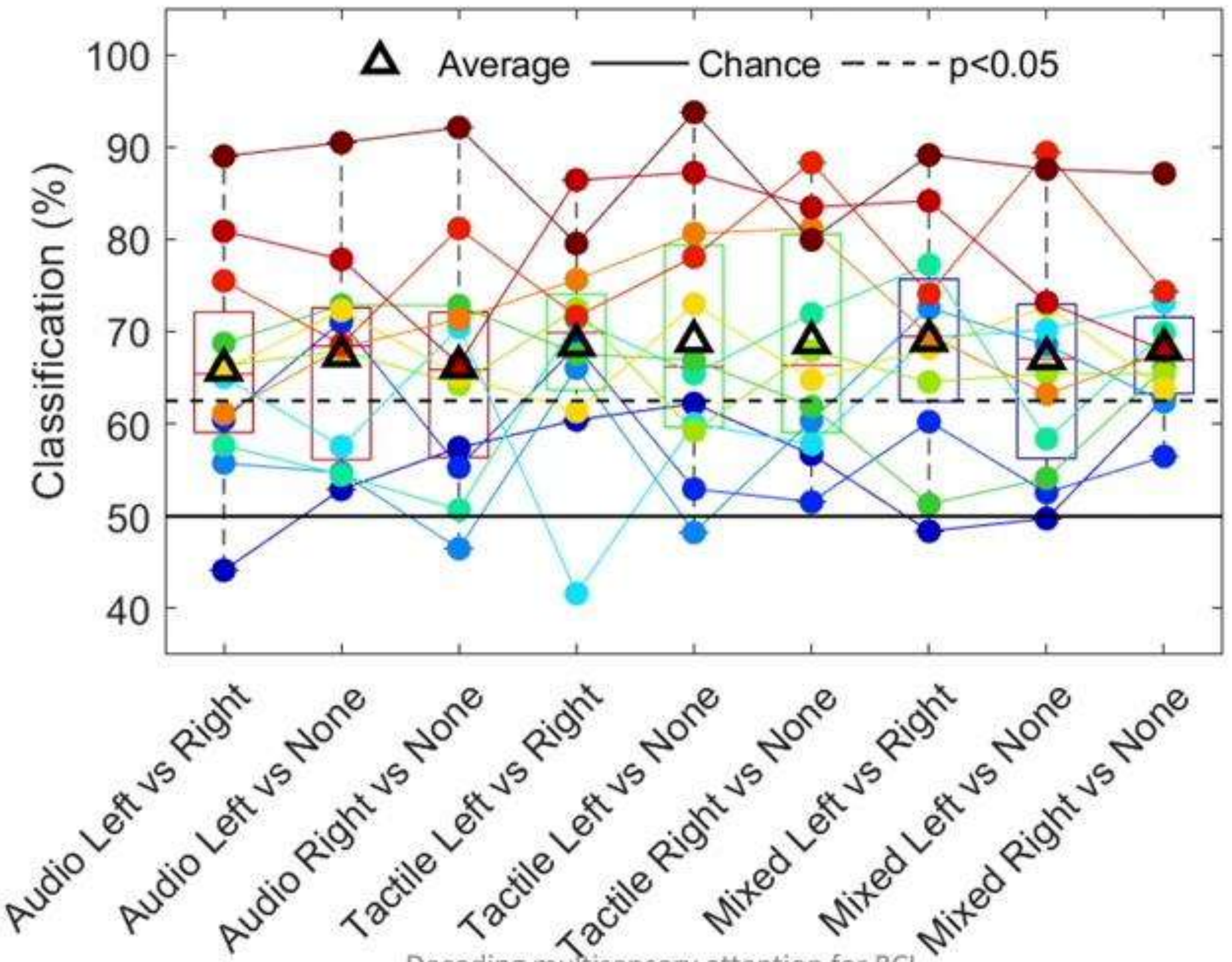
Arindam Jati

Ziqi Fan

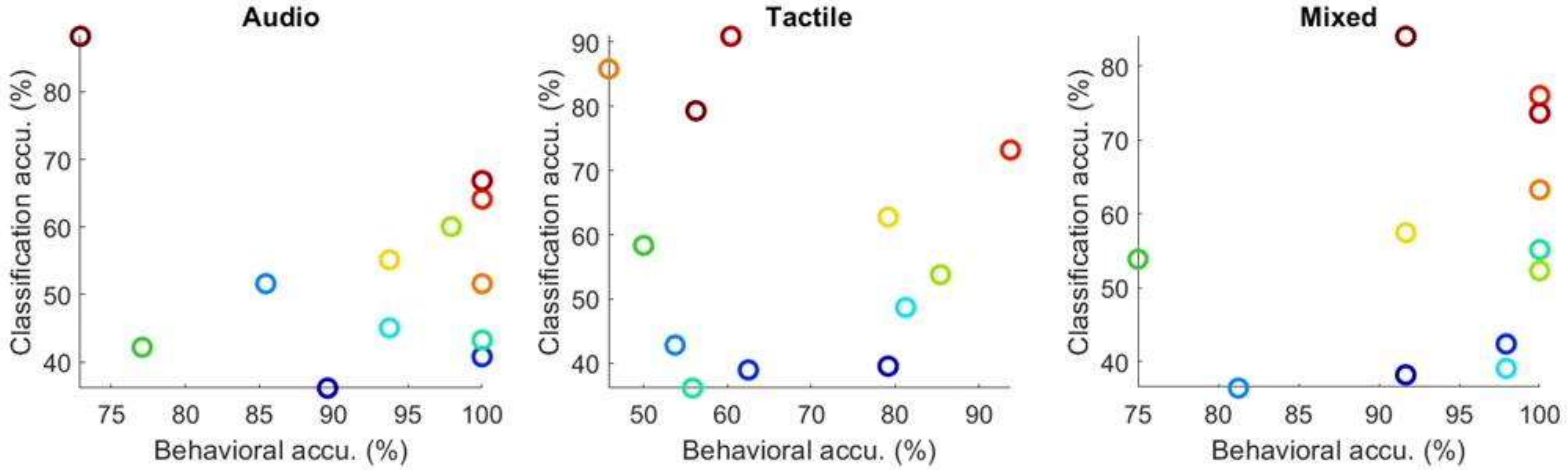
Special thanks to Lincy Wang

Thank you!

Binary classification – surface EEG

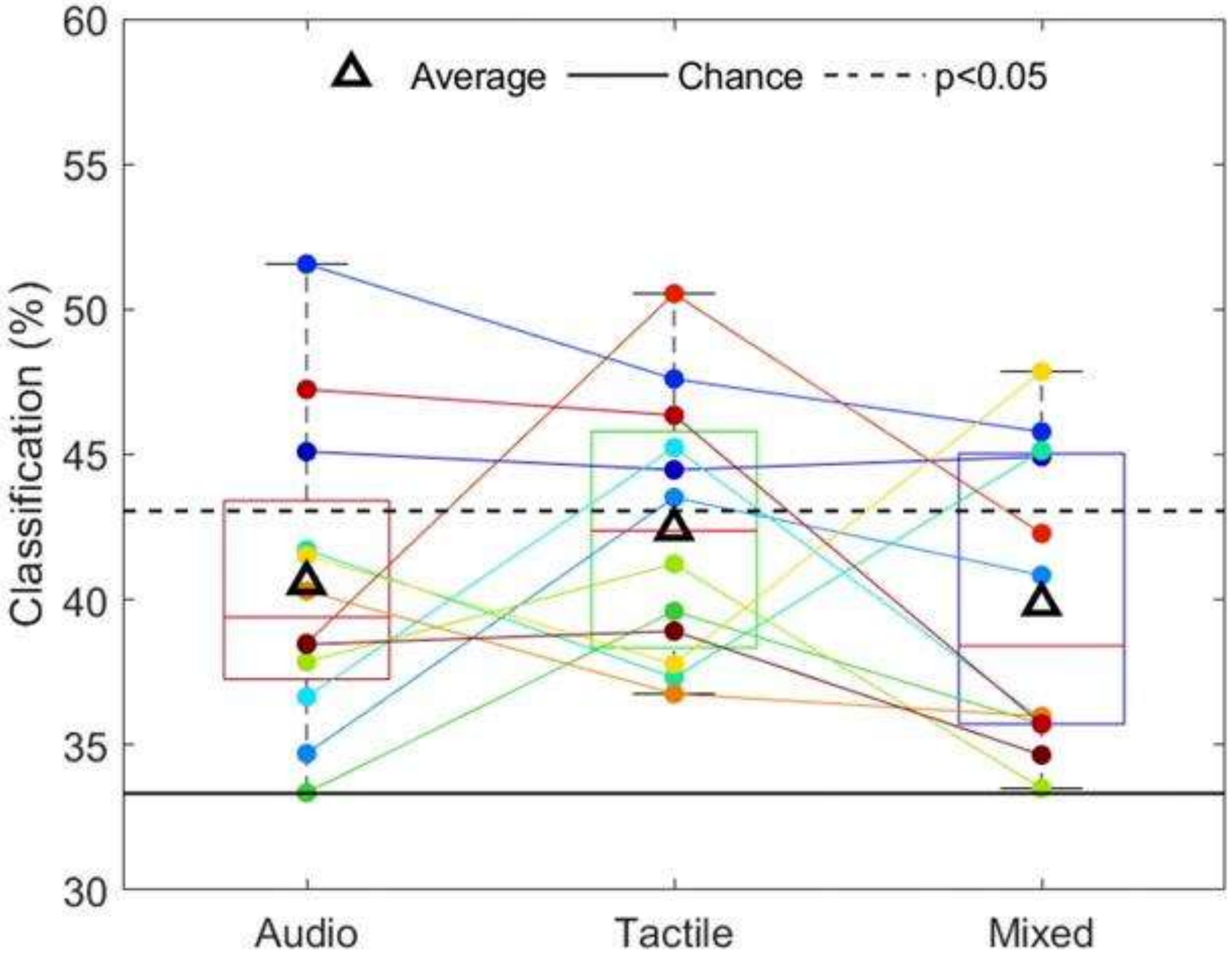


Correlation with behavioral performance



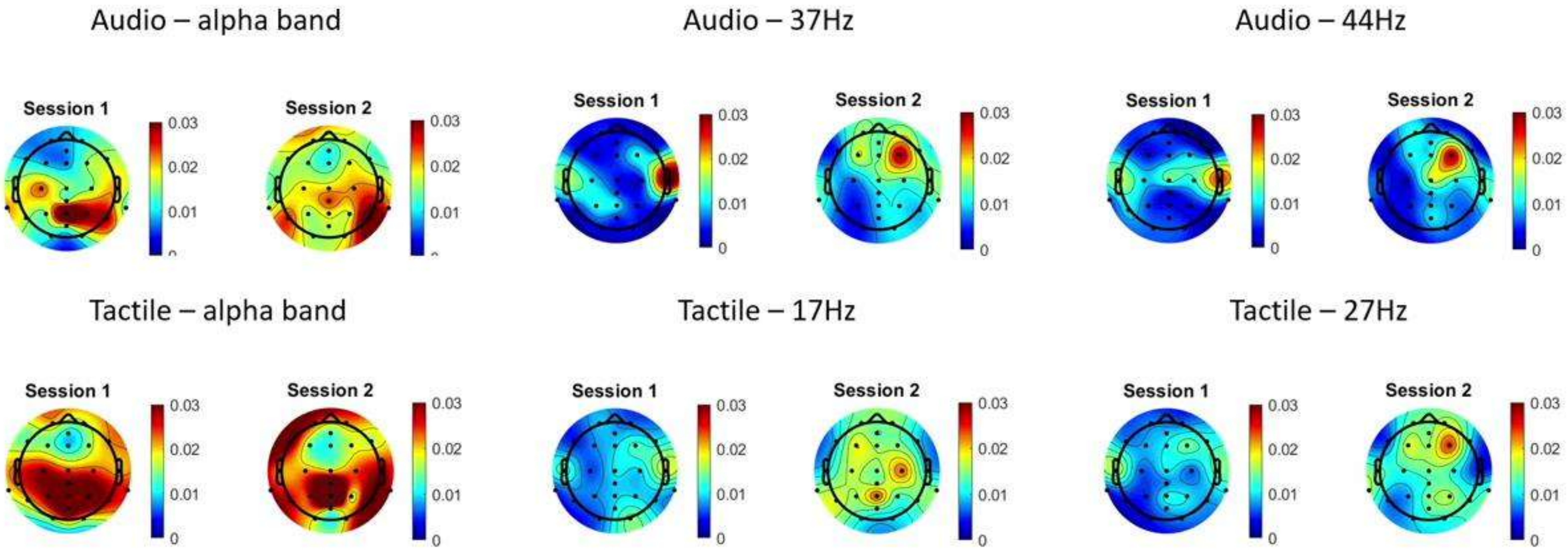
In-ear EEG decoding results by modality

- LDA, 3-way classification
- 1000 cross-validation



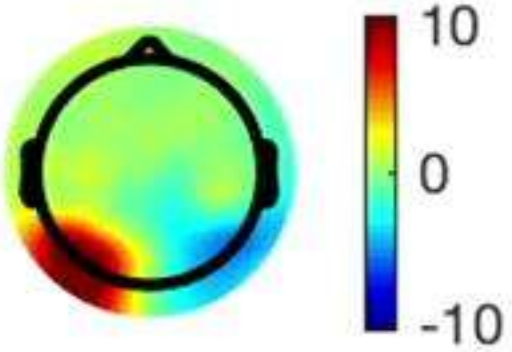
Feature weight consistency between sessions: audio & tactile

Subject 8



Feature weight

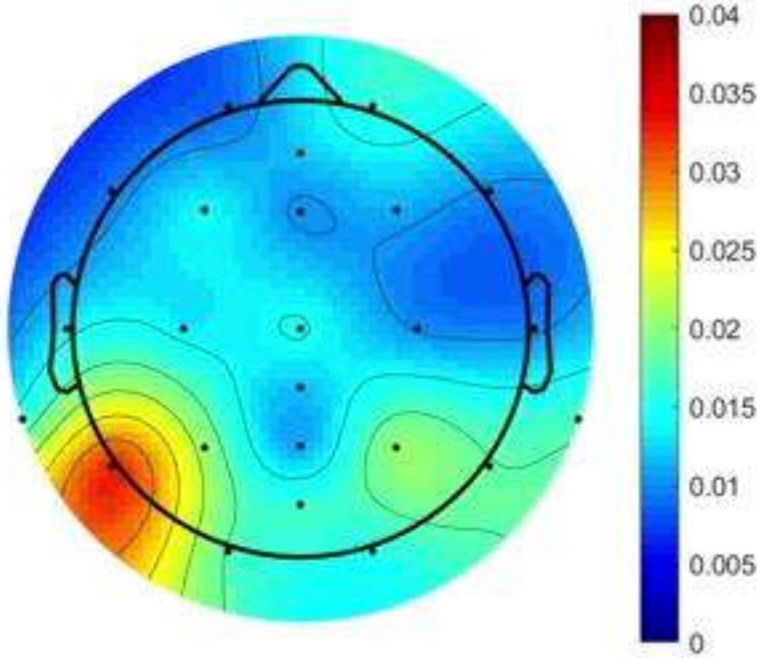
Attend left – Attend right



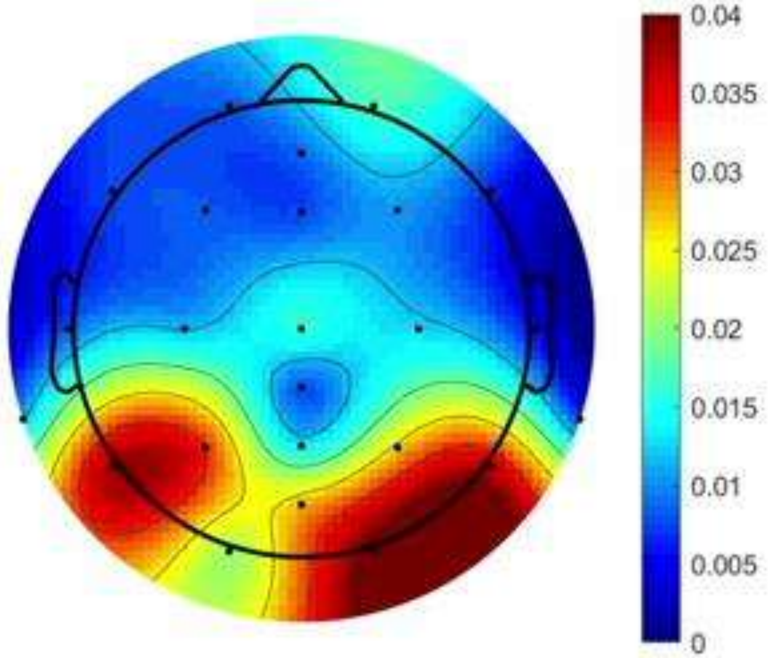
- Neighborhood component analysis (NCA)
  - Feature selection algorithm

Subject 12

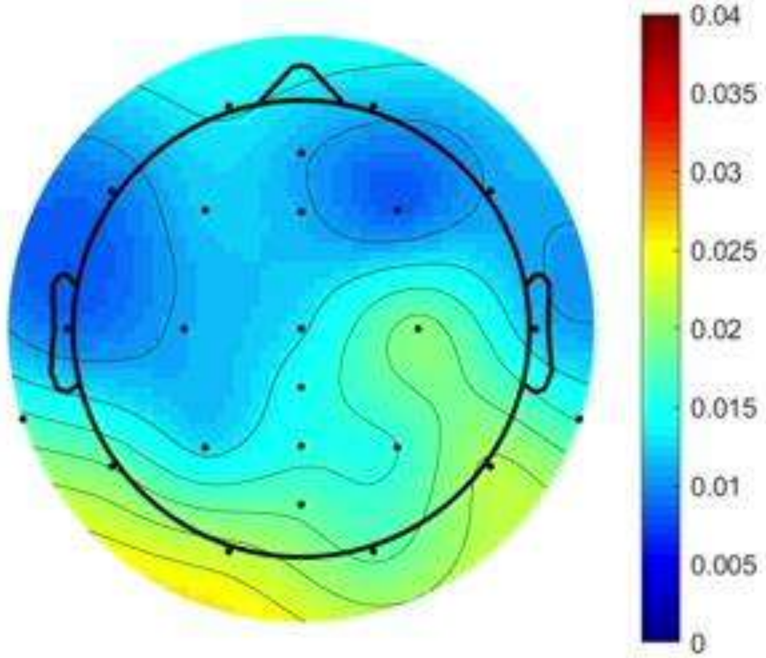
Audio - alpha



Tactile - alpha



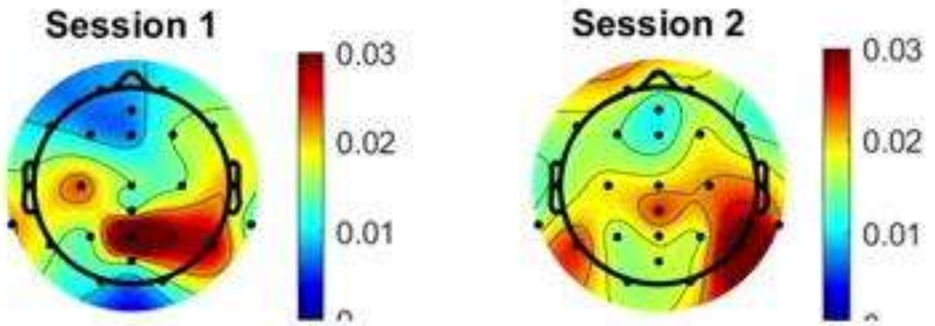
Mixed - alpha



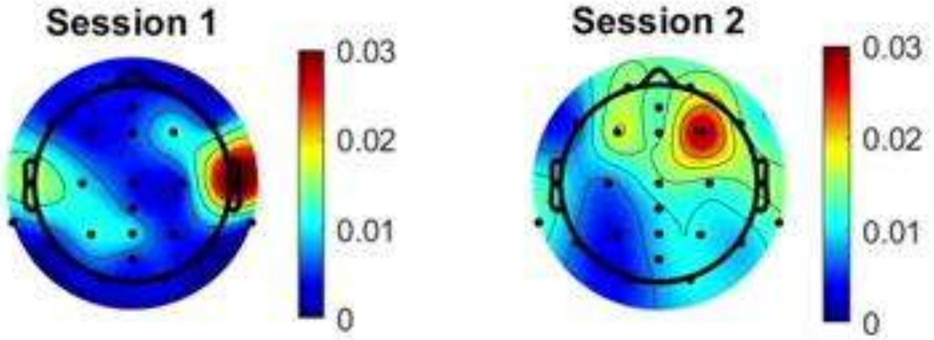
Feature weight consistency between sessions: audio & tactile

Subject 8

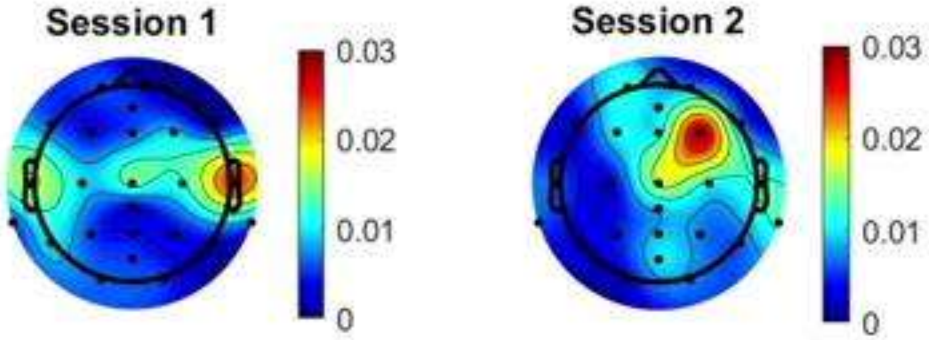
Audio – alpha band



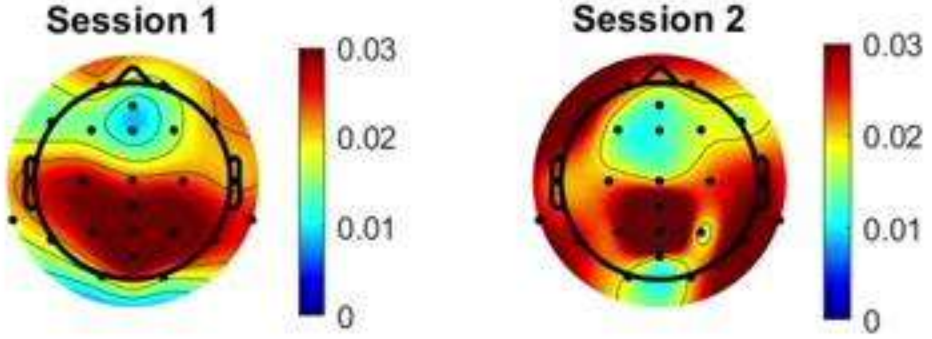
Audio – 37Hz



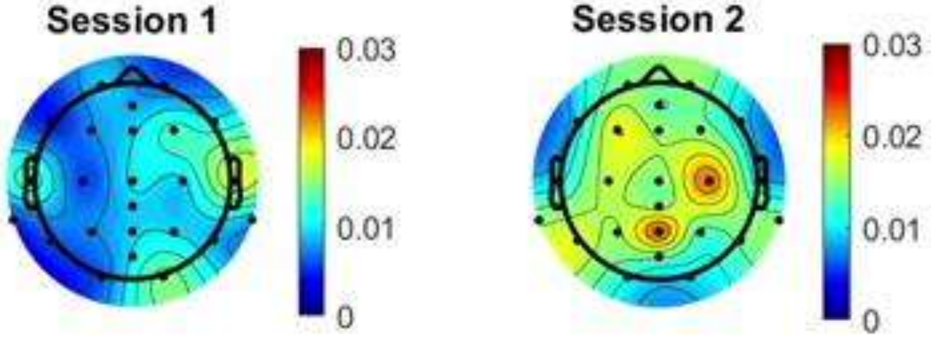
Audio – 44Hz



Tactile – alpha band



Tactile – 17Hz



Tactile – 27Hz