Translating neurophysiological recordings into dynamic estimates of conceptual knowledge and learning

Kaitlyn Peng, Daniel Carstensen, Sarah Parigela, Om Shah, Alex Wingo, Angelyn Liu, Joy Maina, Keene Dampal, Jeremy R. Manning

Department of Psychological and Brain Sciences, Dartmouth College

tl;dr

- We used **EEG data** and **text embeddings** to estimate people's moment-by-moment acquisition of **conceptual knowledge** while watching Khan Academy lectures
- We built and validated a model of conceptual content and knowledge
- We tested whether EEG can help us track learning

Background

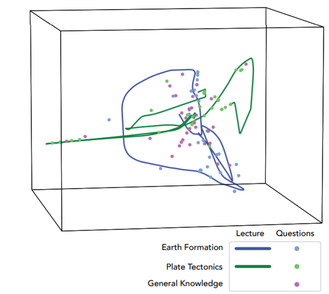
- In our prior work (Heusser et al., 2021), we developed an approach based on **topic models** (Blei et al., 2003) for extracting the dynamic conceptual content from video transcripts
- We previously showed that dynamic content models can accurately characterize learning from online course videos (Fitzpatrick et al., 2025)
- EEG-based **ISFCs** (Simony et al., 2016) track synchronized neural responses to common stimuli

Approach

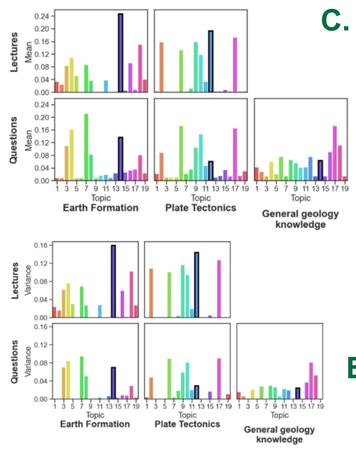
- Selected Khan Academy lectures (Earth Formation and Plate Tectonics) and made 90 quiz questions across 3 categories: Earth Formation, Plate Tectonics, and General Geology
- Collected **64-channel EEG** data from **n = 42** participants
- Computed **ISFC values in sliding windows (w_size = 5, 10, 15)**
- Used time-varying topic vectors to track conceptual content and aligned them with ISFC data to estimate knowledge acquisition

Results

A.



B.



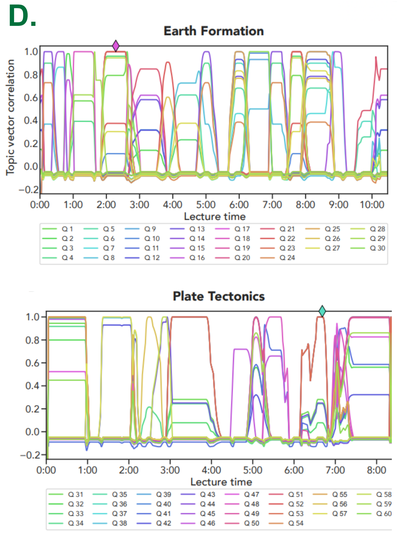
C.

Top 10 words for each topic identified by the LDA model

Topic 1	earth	year	time	time_clock	clock	ago	think	crust	probably	theia
Topic 2	plate	mantle	solid	earth	think	crust	form	glance	layer	kind
Topic 3	plate	travel	supernova	year	light	shockwave	speed	gas	mantle	speed_light
Topic 4	form	year	ago	supernova	time	period	earth	seem	billion	billion_year
Topic 5	ball	molten	magma	crust	earth	molten_ball	giant	rock	still	look
Topic 6	fluid	mantle	kind	layer	still	lithosphere	magma	part	property	asthenosphere
Topic 7	period	hadean	earth	eon	mantle	kind	hadean_eon	molten	time	distinctive
Topic 8	earth	supernova	layer	look	different	surface_earth	surface	magma	happen	formation
Topic 9	plate	move	relative	move_relative	move_away	away	new	land	pacific_plate	new_land
Topic 10	plate	rigid	move	pacific	pacific_plate	bunch	bunch_rigid	relative	move_relative	movement
Topic 11	earth	core	mantle	solid	plate	magma	kind	outer	layer	look
Topic 12	solid	plastic	kind	deformable	stuff	melt	mantle	plastic_solid	magma	little
Topic 13	plate	supernova	year	picture	rigid	form	move	happen	move_relative	relative
Topic 14	earth	happen	eventual	become	planet	supernova	little	formation	splash	probably
Topic 15	plate	form	travel	new_land	new	land	away	eon	move	light
Topic 16	earth	form	think	process	happen	modern_earth	modern	hadean	another	eventual
Topic 17	crust	mantle	layer	solid	different	earth	blue	division	form	core
Topic 18	sun	gas	around_sun	around	part	pressure	kind	little	start	happen
Topic 19	layer	earth	molten_ball	ball	molten	crust	real	lithosphere	asthenosphere	whole
Topic 20	plate	splash	earth	molten	orbit	splash_orbit	theia	relative	big	america
Topic 21	fluid	much	behave_fluid	behave	viscid	deformable	move	associate	thick	slow
Topic 22	earth	fluid	think	kind	blow	glance	think_glance	molten	layer	deformable

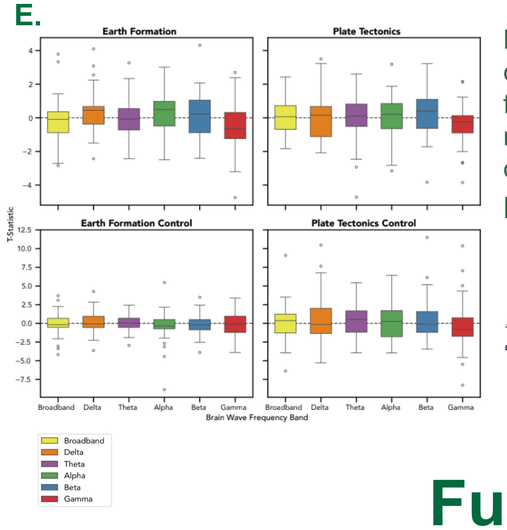
A. Lecture trajectories (lines) and question embeddings (dots).
B. Average topic weight in each lecture and question set.
C. Most common words in each topic.

D.



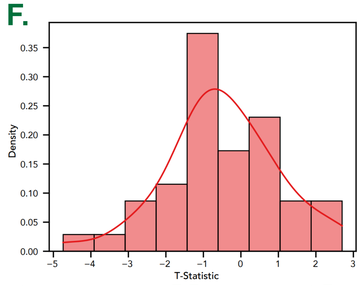
D. Correlation between lecture trajectory and question embeddings. This gives us an estimate of how relevant any moment in a lecture is for a given question.

E.



E. Distribution of t-statistics by brain wave frequency band. Each t-statistic measures the difference of means between ISFC values of correctly and incorrectly answered quiz questions.

F.



F. Distribution of t-statistics in gamma band.

Discussion

- We aligned neural activity with topic model-derived conceptual trajectories in order to explore which aspects of neural activity may be predictive of learning
- These findings pave the way for personalized, adaptive educational tools informed by real-time brain activity

Future directions

- Continue to explore which brain signals are the most predictive of learning
- Leverage these brain signals to track moment-by-moment learning
- Utilize information about predicted moment-by-moment learning to personalize teaching

Bibliography

1. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. J. Mach. Learn. Res.
2. Fitzpatrick, P. C., et al. (2025). PsyArXiv.
3. Simony, E., et al. (2016). Nature Communications.