Supplementary materials for: Text embedding models yield high-resolution insights into conceptual knowledge from short multiple-choice quizzes

Paxton C. Fitzpatrick[1], Andrew C. Heusser[1, 2], and Jeremy R. Manning[1, *]

[1]Department of Psychological and Brain Sciences
Dartmouth College, Hanover, NH 03755, USA

[2]Akili Interactive Labs
Boston, MA 02110, USA

[*]Corresponding author: Jeremy.R.Manning@Dartmouth.edu

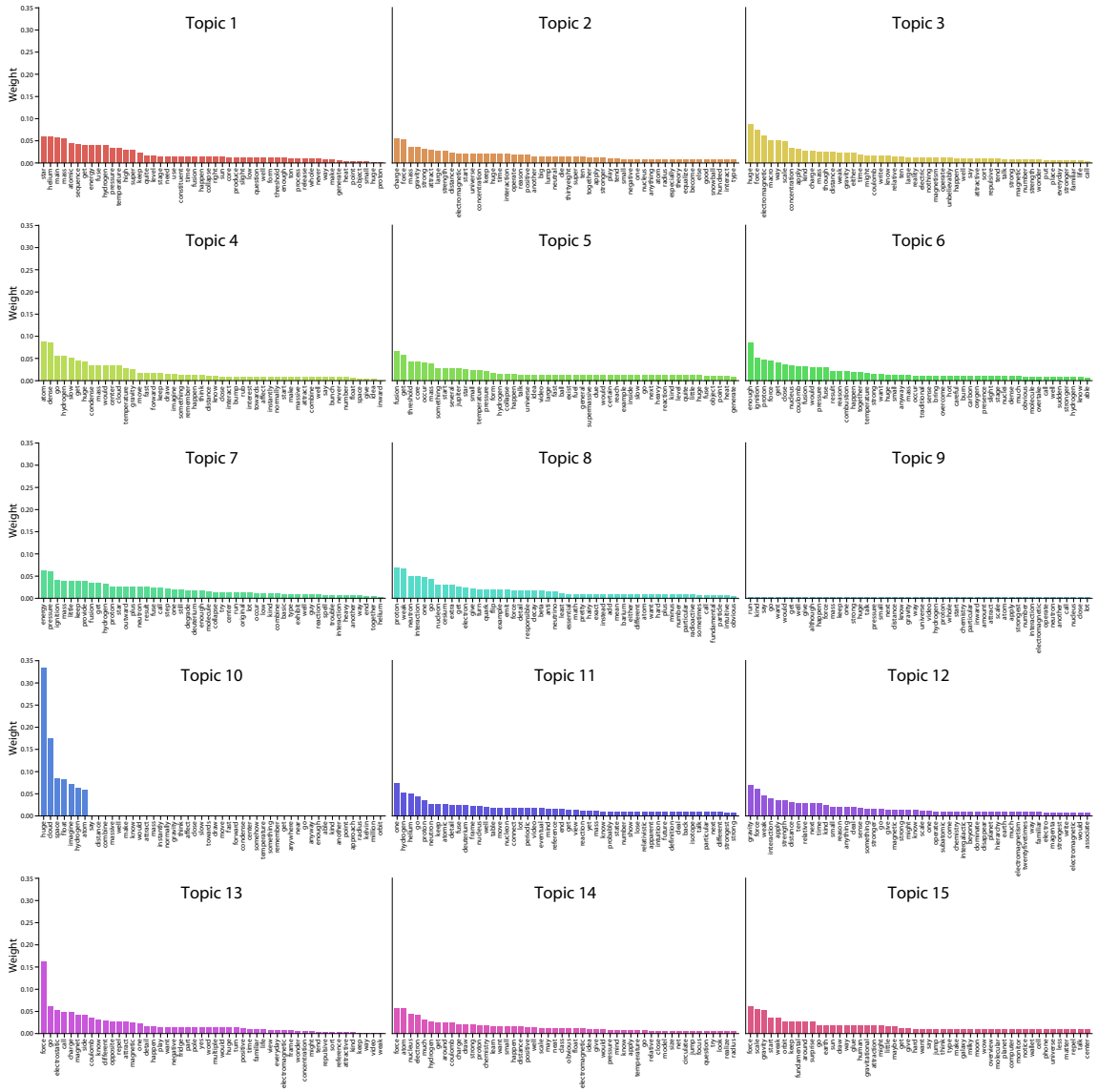
ID	Question set	Question	Correct response	Alternative 1	Alternative 2	Alternative 3
1	FFF	Why is the gravitational attraction between you and your computer too small for you to notice?	Neither you nor your computer has enough mass to cause a noticable gravitational attraction	You and your computer are too close for the gravitational attraction to be significant	Humans are too small to detect the force of gravity	The gravitational attraction between you and your computer is disrupted by the larger gravitational field generated by the earth
2	FFF	Which of the following is an example of the Weak Interaction?	A neutron in a radioactive Cesium atom is converted into a proton, leading to the release of a few particles	Light from the sun collides with a satellite orbiting Earth and exerts a small push on the satellite	Two protons bound together in a Helium nucleus resist separation despite a repulsive electromagnetic force acting on them	A distant galaxy exerts a small but detectable gravitational pull on the Earth
3	FFF	Roughly how many times stronger is the Weak Interaction than gravity?	10,000,000,000,000,000,-000,000,000	10	1,000,000	The Weak Interaction is less strong than gravity
4	FFF	Why don't you and your computer experience any attraction or repulsion due to the Weak Interaction?	The weak interaction only acts over extremely small distances	The weak interaction between you and your computer is counteracted by the other forces	You and your computer have no net charge	Neither you nor your computer has enough mass to induce a significant Weak Interaction
5	FFF	Which of the following is a difference between gravity and the electromagnetic force?	Gravity is only ever attractive while the electromagnetic force can both attract and repel	Gravity is a much more powerful force than electromagnetism	Gravity can only act over large distances while the electromagnetic force can act over large and small distances	The electromagnetic force can only act over small distances while gravity can act over small or large distances
6	FFF	Electricity and magnetism can be shown to be two cases of the same force if we:	View them in different frames of reference	Switch which charges we call positive and which charges we call negative	Consider both the effects over small distances and the effects over large distances	Consider both the attractive and repulsive properties of the two forces
7	FFF	Which of the following are the primary two fundamental forces acting in opposition between the positively-charged protons in an atom's nucleus?	The Strong Force and the Electromagnetic Force	Gravity and the Weak Interaction	Gravity and the Electromagnetic Force	The Strong Force and the Weak Interaction
8	FFF	Why does the universe have a very uneven distribution of mass but a relatively equal distribution of charge?	Positive and negative charges cancel out and become a neutral charge when they combine while masses only grow larger as they combine	Masses tend to repel while charges tend to attract	Masses tend to attract while charges tend to repel	The gravitational interaction acting between masses is stronger than the electromagnetic interaction acting between charges
9	FFF	In your body, there are a tremendous amount of negatively-charged electrons. Your computer also contains a huge number of negatively-charged electrons. We know that like charges repel, but you and your computer are not repelled apart. Which of the following explains why?	The electrons' negative charges are balanced by the positive charges of an equal number of protons	An attractive gravitational force balances out this repulsion	The Electromagnetic force only acts over very small distances	The Electromagnetic force only acts over very large distances
10	FFF	Which of the following is a similarity between the Weak Interaction and the Strong Force?	Both act only over very small distances	Both are stronger than the Electromagnetic force	Both are weaker than Gravity	Both are responsible for attractions between distant galaxies
11	FFF	Which force is stronger than the Electromagnetic Force?	Strong Force	Gravity	Weak Interaction	Electromagnetic Force is the strongest
12	FFF	Roughly how many times stronger is the Strong Force than gravity?	10^38	100	10^18	The Strong Force is weaker than gravity
13	FFF	Which of the following would have to be true for the Weak Interaction to cause repulsion or attraction between two objects?	The objects would have to be extremely close to each other	The objects would have to have the same mass	The objects would have to be extremely far away from each other	The objects would have to have different masses
14	FFF	Which force keeps us from jumping off of Earth?	Gravity	Strong Force	Weak Interaction	Electromagnetic Force
15	FFF	What does the Coulomb Force refer to?	The repulsion of objects with similar charge and the attraction of objects with different charge	The repulsion of objects with similar mass and the attraction of objects with different mass	The repulsion of objects with similar temperature and the attraction of objects with different temperature	The repulsion of objects with similar density and the attraction of objects with different density
16	BoS	Which of the following describes the effect of gravity on a cloud of atoms?	The atoms move to the center of the mass of the atoms	The atoms move away from the center of the mass of the atoms	The atoms spin around the center of the mass of the atoms	Gravity has no effect on a cloud of atoms
17	BoS	Which of the following occurs as a cloud of atoms gets more dense?	Temperature increases	Temperature decreases	Mass increases	Mass decreases
18	BoS	Which temperature does a cloud of hydrogen atoms approach as it gets denser in the process of becoming a star?	10 Million Kelvin	0 Kelvin	10,000 Kelvin	10 Billion Kelvin
19	BoS	Which of the following can overcome the Coulomb Force?	High temperature and high pressure	Low temperature and high pressure	High temperature and low pressure	Low temperature and low pressure
20	BoS	Which of the following prevents a star from collapsing as a result of gravity?	Energy released from the fusion of hydrogen atoms provides outward pressure	The fusion of hydrogen atoms decreases the temperature of the star	The gravitational pull of other stars nearby	The Weak Interaction
21	BoS	How are supermassive stars different from other stars?	Fusion occurs very fast	Fusion occurs very slow	Fusion occurs in the reverse order	Fusion does not occur at all

22	BoS	Which of the following is the FIRST product of two hydrogen atoms fusing together?	Deuterium	Oxygen	Helium	Beryllium
23	BoS	Once hydrogen atoms get close enough together, which of the following keeps them together?	The Strong Force	The Electromagnetic Force	Gravity	The Weak Interaction
24	BoS	When two nuclei fuse together, how does the mass of the combined nucleus compare to the mass of each of the original nucleus?	The mass of the combined nucleus is smaller	The mass of the combined nucleus is larger	The mass of the combined nucleus is the same	It is not possible for two nuclei to fuse together
25	BoS	If we say that our Sun is a main sequence star, what does that tell us about the Sun?	Hydrogen atoms in the Sun are fusing together and becoming Helium	The Sun is a supermassive star	The Sun does not experience the force of Gravity but does experience the Coulomb Force	The Sun is comprised of 10 million Hydrogen atoms
26	BoS	Which force would cause a massive cloud of hydrogen atoms to move together?	Gravity	Strong Force	Weak Interaction	Electromagnetic Force
27	BoS	Which of the following occurs as density increases?	Temperature increases	Volume increases	Mass increases	None of the above
28	BoS	Which of the following is a product of Hydrogen fusion?	Helium	Oxygen	Cesium	Carbon
29	BoS	Which of the following terms accurately describes the Sun?	Main sequence star	Supermassive star	Alternative sequence star	None of the above
30	BoS	Which of the following terms best describes a fusion reaction?	Ignition	Combustion	Decomposition	Displacement
31	GPK	Which of the following lists of particles is ordered from smallest to largest?	Electron, proton, nucleus, atom	Atom, electron, proton, nucleus	Electron, nucleus, atom, neutron	Neutron, nucleus, electron, atom
32	GPK	Which of the following defines what element an atom is?	Its number of protons	Its number of neutrons	Its number of electrons	Its mass
33	GPK	Suppose that in some atom, a proton is converted into a neutron. What changes as a result of this conversion?	The atom's element	The atom's mass (in atomic mass units)	The atom's velocity	The atom's density
34	GPK	Which of the following lists is ordered from smallest to largest?	Star, solar system, galaxy, universe	Galaxy, solar system, Milky Way, universe	Planet, galaxy, star, solar system	Earth, solar system, universe, galaxy
35	GPK	Which of the following are located in the nucleus of an atom?	Protons and neutrons	Only protons	Only electrons	Neutrons and electrons
36	GPK	Which of the following has the least mass?	An electron	A proton	A neutron	A hydrogen atom
37	GPK	What percent of an atom's space does its nucleus occupy?	Less than 1%	10%	50%	More than 90%
38	GPK	In the famous equation attributed to Albert Einstein, E = mc^2, what does the letter "m" represent?	Mass	Momentum	Moment of inertia	Moles
39	GPK	If I were to heat up an inflated balloon, which of the the following would occur?	The balloon would expand	The balloon would shrink	None of these answers are correct	The balloon could expand or shrink depending on whether it's filled with air or helium gas

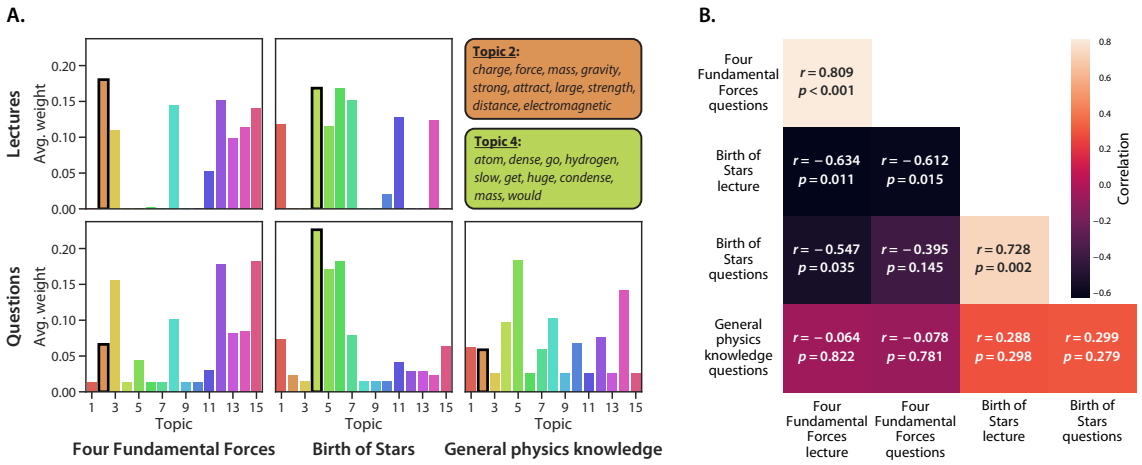
Supplementary Table 1: Question pool. Each participant completed three quizzes consisting of 13 multiple-choice questions randomly selected from the question pool. Each quiz included 5 questions about the *Four Fundamental Forces* lecture (FFF), 5 questions about the *Birth of Stars* lecture (BoS), and 3 questions about general physics knowledge (GPK). Each question appeared on at most one quiz (for each participant).

Topic	1	2	3	4	5	6	7	8	9	10
1	star	helium	main	mass	atomic	sequence	get	energy	fuse	hydrogen
2	charge	force	mass	gravity	strong	attract	large	strength	distance	electromagnetic
3	huge	force	electromagnetic	macro	way	scale	concentration	apply	kind	charge
4	atom	dense	go	hydrogen	slow	get	huge	condense	mass	would
5	fusion	get	threshold	core	occur	mass	something	start	several	jupiter
6	enough	ignition	proton	force	get	close	nucleus	coulomb	fusion	would
7	energy	pressure	ignition	mass	little	keep	provide	fusion	get	hydrogen
8	proton	weak	neutron	interaction	one	go	nucleon	cesium	extra	get
10	huge	cloud	space	float	imagine	hydrogen	atom	say	distance	combine
11	one	hydrogen	helium	go	proton	neutron	keep	atomic	detail	fuse
12	gravity	force	weak	interaction	apply	strength	distance	ten	relative	next
13	force	go	electrostatic	call	charge	magnet	side	coulomb	know	different
14	force	atom	nucleus	electron	much	hydrogen	get	around	coulomb	charge
15	force	scale	gravity	start	weak	orbit	keep	fundamental	around	surprise

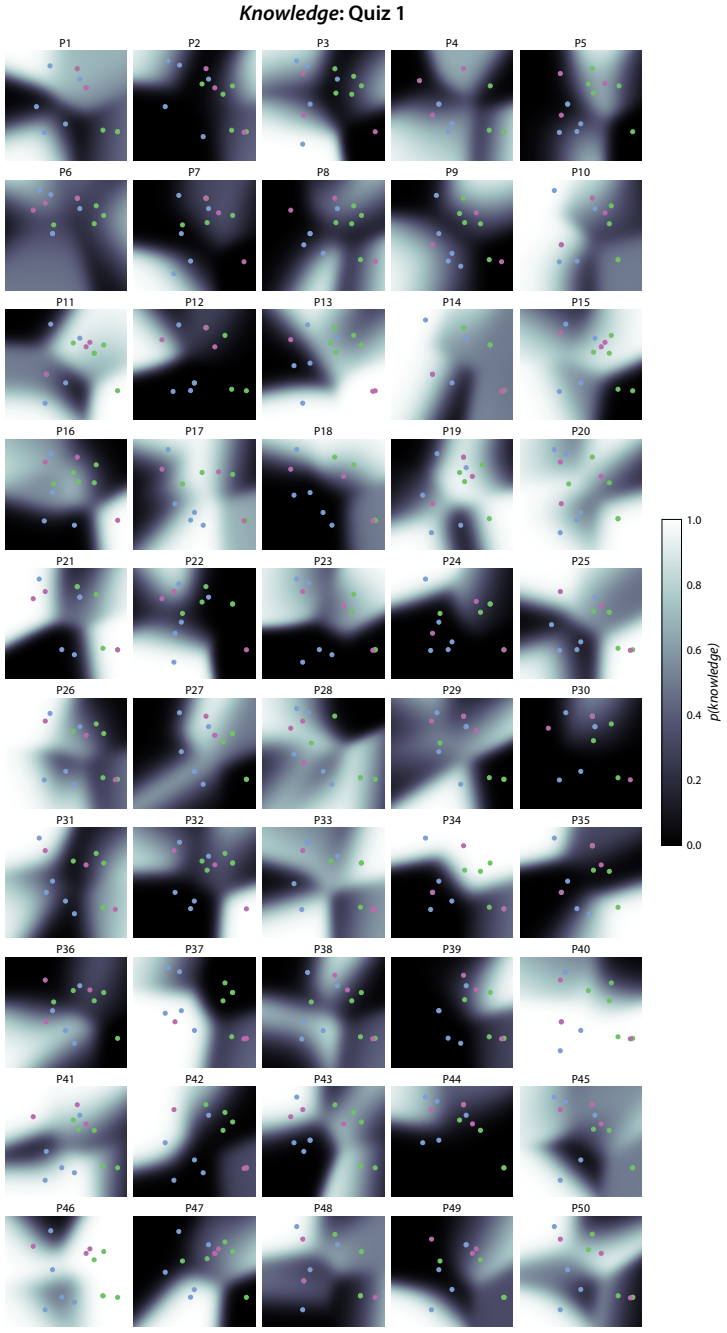
Supplementary Table 2: Topics. We fit a topic model with (up to) $k = 15$ topics to sliding windows parsed from the two course video transcripts (see *Constructing text embeddings of multiple lectures and questions*), and identified 14 topics with non-zero weights. The table displays the 10 top-weighted words (columns) from each of the topics (rows) discovered by the model. The weights of these words are shown in Supplementary Figure 1.



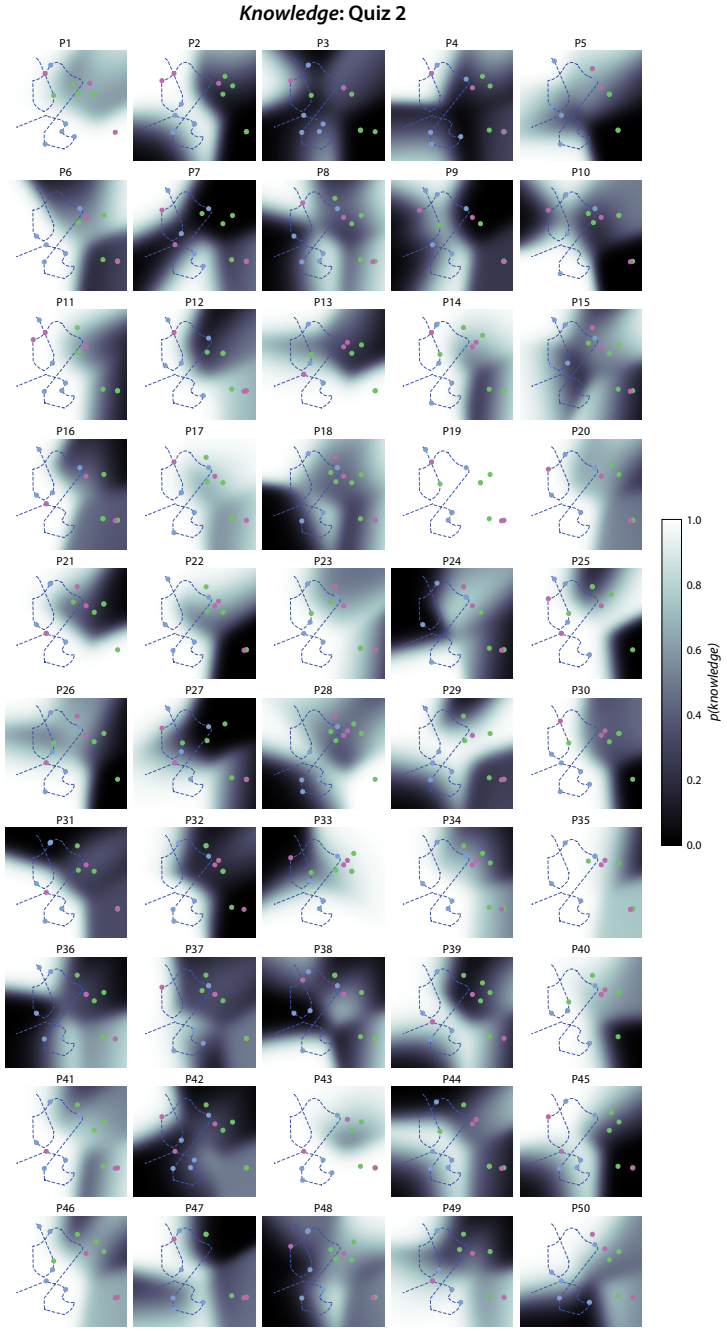
Supplementary Figure 1: Topic-word distributions. Each topic is defined by a distribution of weights over words in the topic model's vocabulary (i.e., the union of all unique words from the two lectures' transcripts, excluding stop-words, after preprocessing). Each plot above displays the weights for the 50 words weighted most heavily by an individual topic extracted from the lectures' contents (see *Constructing text embeddings of multiple lectures and questions*).



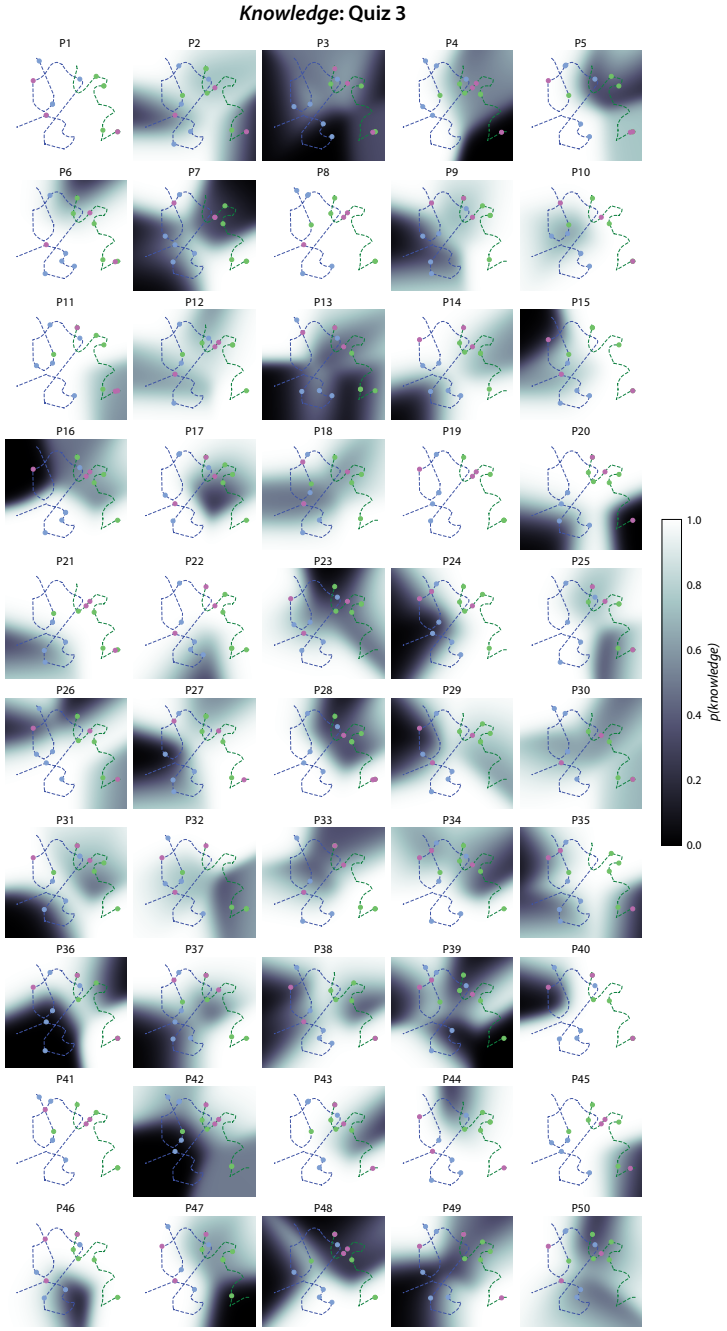
Supplementary Figure 2: Topic weights. A. Average topic weights for each lecture and question set. The bar plots display each topic's average weight across lecture timepoints (top row) and questions (bottom row); colors denote topics. The top-weighted words from the highest-weighted topic from each lecture are displayed in the upper right (orange: topic 2; yellow-green: topic 4). The top-weighted words from the full set of topics may be found in Supplementary Table 2. **B. Relationships between average topic weights.** Pairwise correlations between the distributions of average topic weights for each lecture and question set. Each row and column corresponds to a bar plot in Panel A. Also see Figure 3 in the main text.



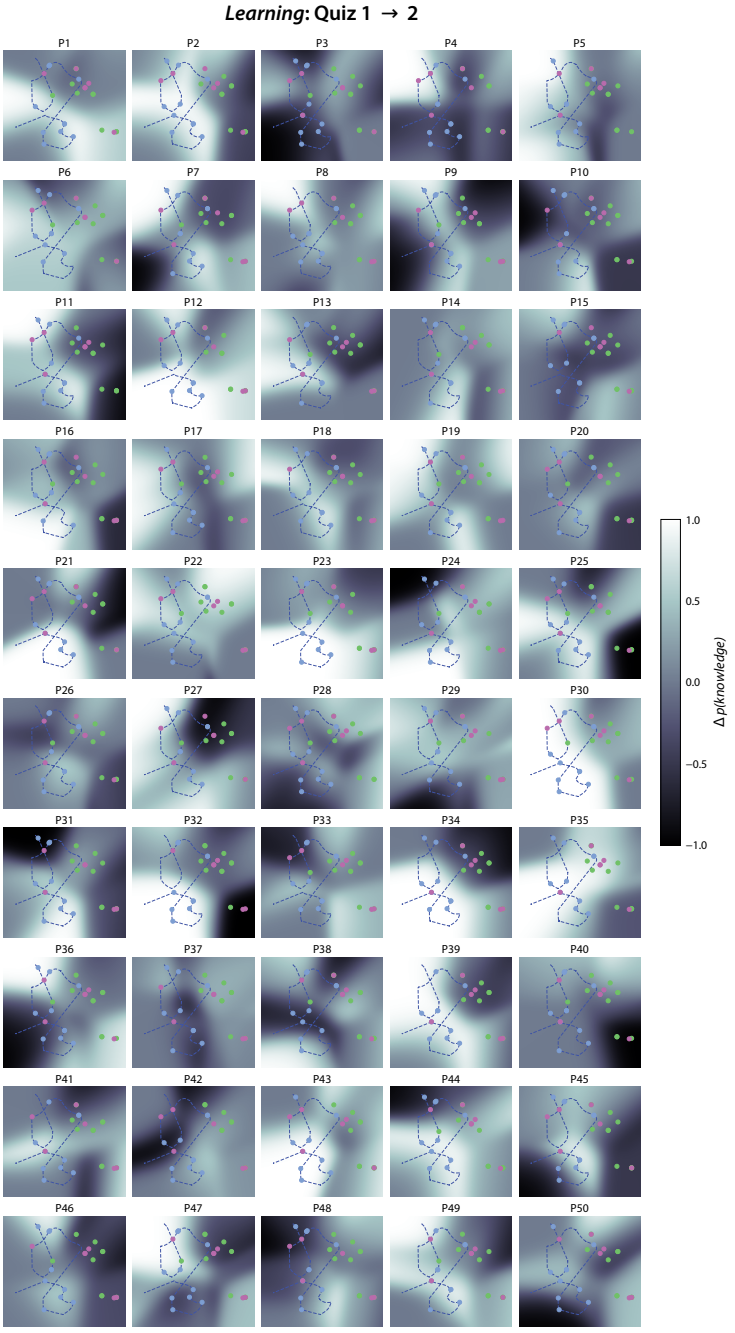
Supplementary Figure 3: Individual participants' knowledge maps estimated from Quiz 1 responses. Each panel is in the same format as the knowledge map displayed in the left panel of Figure 7A in the main text, but here the maps are shown separately for each participant.



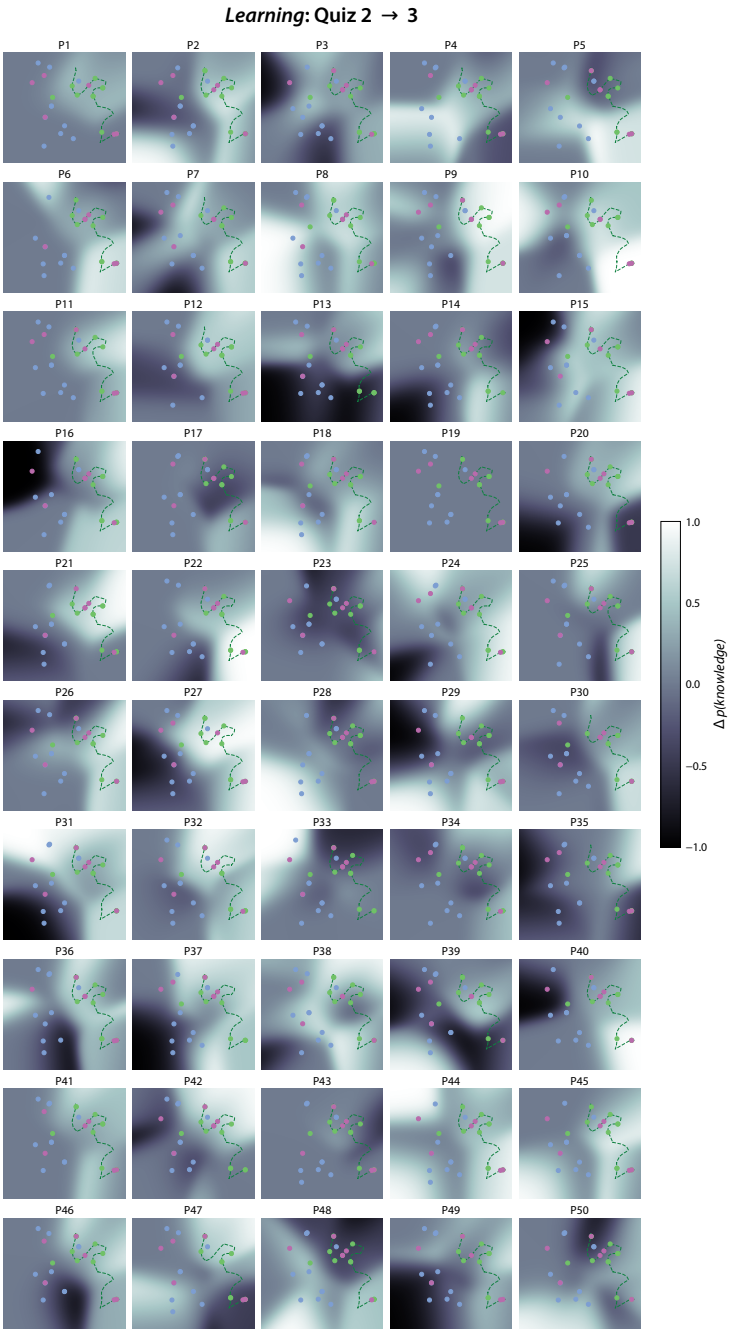
Supplementary Figure 4: Individual participants' knowledge maps estimated from Quiz 2 responses. Each panel is in the same format as the knowledge map displayed in the middle panel of Figure 7A in the main text, but here the maps are shown separately for each participant.



Supplementary Figure 5: Individual participants' knowledge maps estimated from Quiz 3 responses. Each panel is in the same format as the knowledge map displayed in the right panel of Figure 7A in the main text, but here the maps are shown separately for each participant.



Supplementary Figure 6: Individual participants' learning maps estimated from Quiz 1 and 2 responses. Each panel is in the same format as the learning map displayed in the left panel of Figure 7B in the main text, but here the maps are shown separately for each participant.



Supplementary Figure 7: Individual participants' learning maps estimated from Quiz 2 and 3 responses. Each panel is in the same format as the learning map displayed in the right panel of Figure 7B in the main text, but here the maps have been computed separately for each participant.