PSY1409: The Psychology of Large Language Models

📍 **Location**	TBD		👤 **Instructor**	Haley Keglovits, Ph.D.
🕐 **Time**	Wednesday 9:45–11:45 AM		✉ **Email**	hkeglovits@fas.harvard.edu
📅 **Office Hours**	TBD		🔗 **Canvas**	Link

Course Description

How was ChatGPT inspired by your brain? As the use of large language models (LLMs) grows in popularity and applicability to academics, and as we grapple with how and when to incorporate them into the research process, having a foundational understanding of LLMs is critical to ensuring we are careful interpreters of these new, powerful tools. In this course, we will study the development of LLMs and how they were inspired by findings in psychology, linguistics, and cognitive neuroscience. After delving into LLM architecture and functioning, we will debate practical and ethical questions about using LLMs in psychological research for coding, writing, summarizing papers, as well as the broader societal applications (e.g., whether using LLMs to teach content in primary schools will lead to a larger or smaller educational disparities). This course is not designed to teach you to code an LLM, but rather to help you understand the networks at a mechanistic level. We will use some code to illustrate theoretical concepts, but prior coding experience is not required. Classes will involve lectures, demonstrations, activities, and discussion. There are no additional discussion sessions associated with the course.

Course Assignments

Conversation with a Friend (15%)

For this assignment, you will record a conversation about our class content with a friend or family member who is not an expert on LLMs. Your goal is to ascertain what they know, and then help them understand LLMs more deeply than they did at the start of the conversation, using the content we have covered in class. You should record a conversation which is 10-20 minutes long, starting with the following question: Can you tell me what you understand about how ChatGPT and other large language models work? In addition to the audio recording, you will turn in a 2-page reflection on the experience detailing what was difficult to explain, where you think you need to learn more, and how you decided to convey your knowledge.

Multi-Media Project (15%)

For this project, you will create a multi-media educational tool with content from our course to (briefly) explain how LLMs work and how they can help and hinder a students' learning. You will have flexibility in the type of tool you create (visual presentation, science podcast, comic strip, etc), and we will work with the Bok Center Media Lab team to support your design and execution of your project. In whatever form you choose, you must effectively convey educational content from our course to a student audience.

Final Assignment (20%)

For your final assignment, you will propose and defend what you believe Harvard Psychology's AI policy should be. Your argument should be grounded in a description of what generative AI is, and the types of things it can and cannot do. You can bring in examples from your previous classes, stories you hear from friends, etc., but you must incorporate what we learned over the semester in the discussions, readings, and lectures. This policy should be Harvard Psychology specific, but you can use supporting evidence from other circumstances to bolster your argument. This assignment will take the form of the policy (which should be one page or less but specific), and then an 8 page explanation as to why you chose that policy and why you believe it is best for Harvard students.

Class Debates - Leader (15%)

You and a partner will be assigned one of the following prompts to lead a discussion in class and online. As the leaders on your assigned week, you will find two opposing opinion articles related to the topic, which you will send to me by Sunday night for distribution to the class. On Wednesday in class, you will give a brief introduction to the topic with main considerations for your classmates to discuss – for example, a specific case which highlights this dilemma, or how differing perspectives support their views. In preparation for the class presentation, together you will submit a 1-2-page description of what you found to me by the start of class on your assigned week. Then, you will moderate your classmates' conversation (details below) in class and on Canvas through the end of the week, and submit an individual 2-page summary of the discussion and reflection on what you learned by the following Wednesday. More details about this assignment will be provided on Canvas, and I will lead the first two discussions to model how we will approach these.

Class Debates - Commentor (15%)

On the weeks you are not leading the debates, you will be participating during in-class discussions by sharing your views on the topic posed. This engagement will support your participation grade (below). In addition, after each debate you will write a reflection on what you took from the discussion, and how you believe we as a society should approach this debate. In essence, you'll be writing your own opinion piece to follow up the ones you read. Of the 10 discussions you do not lead, you must submit 8 reflections, which should be 500-1000 words.

Debate Topics

Week 4: Should LLMs be used as conversational companions for people experiencing loneliness or depression? What are the psychological risks and benefits of people forming emotional attachments to AI chatbots?

Week 5: Should parents or educators encourage young children to interact with LLMs as a way to support language development? Does this differ for first/second languages?

Week 7: Will AI tutors enhance learning or reduce critical thinking? Will the widespread use of LLMs in primary or secondary schools increase or reduce educational disparities?

Week 9: Should LLMs replace or supplement special education services?

Week 10: Should LLMs be trained to be as human-like as possible, including human biases and failings, or should we tune them to be something different?

Week 11: Can LLM therapy tools (like AI therapists) be ethically used in mental health care? How does this differ (if at all) from physical health care?

Week 12: How should students conceptualize the work they produce with the assistance of AI? Does this apply to outside the classroom (ie, AI artists or bloggers)?

Week 13: Is it ethical to create deepfakes (fine-tuned LLMs) of deceased loved ones for grief processing?

Participation (20%)

Active engagement is a key part of how we build knowledge together in this course. I expect everyone to contribute to class discussions across the semester. You don't have to speak on every topic, but you should be present, prepared, and willing to share your thoughts, questions, or reactions when you're in class. Our class time will be richer if each of us participates fully, and your voice matters in that process. During our first class, we'll co-create a set of expectations for what thoughtful, respectful discussion looks like. Your participation in the class debates are especially important for this course, as they form the basis for your reflections on the ethical dilemmas AI brings us. About halfway through the semester, you'll receive a participation check-in (on a scale from 1 to 20) to give you feedback and a chance to adjust if needed. I will take cross-semester improvements into consideration for your final grade, so it will not be set in stone at that point.

Grade Breakdown

	B+: 87–89	C+: 77–79	D+: 67–69	
A: 93–100	B: 83–86	C: 73–76	D: 63–66	E: <60
A-: 90–92	B-: 80–82	C-: 70–72	D-: 60–62	

Course Schedule

Full reading list available [here].

Week	Core Topic	Core Reading	Debate Topic	Debate Reading
Week 1: January 28	Introduction, What is an LLM	Lake et al. (2017)	No debate topic this week.	
Week 2: February 4	Biological + Artificial Neurons	McCulloch & Pitts (1943)	LLM Educators	Lu et al. (2024)
Week 3: February 11	Learning as Changing Weights, Neuroplasticity	Rosenblatt (1958)	LLMs in Psychological Research	Musslick et al. (2025)
Week 4: February 18	Learning Algorithms, Types of Learning	Crick (1989)	LLM Companions for Loneliness	Ventura et al. (2025)
Week 5: February 25	Representing Language	Bengio et al. (2003)	Child Language Development	Getman et al. (2023)
Week 6: March 4	Transformers	Vaswani et al. (2017)	No debate topic this week	
Week 7: March 11	Architecture + Brain Networks	Schrimpf et al. (2021)	LLMs in Primary Schools	Holmes et al. (2022)
Conversation with a Friend Due by March 13th at 11:59PM				
Week 8: March 18	*Spring Break - No Class*			
Week 9: March 25	Training Data	Zhong et al. (2024)	LLMs for Special Education	Voultsiou et al. (2025)
Week 10: April 1 @ Media Lab	Network and Human Biases	Bai et al. (2024)	Human Biases in LLMs	Hofmann et al. (2024)
Week 11: April 8	Self Attention + Working Memory	Wei et al. (2022)	LLMs as Therapists	Weizenbaum (1966)
Week 12: April 15	Hallucinations + Episodic Memory	Lin et al. (2022)	Ownership of LLM created content	Hamilton et al. (2024)
Multi-Media Project Due by April 17th at 11:59PM				
Week 13: April 22	Neural Representations	Sucholutsky et al. (2023)	LLMs in Grief Processing	Fabry & Alfano (2024)
Week 14: April 29	Consciousness	Turing (1950)	Is the Turing Test obsolete?	Chiang (2010)
Final Assignment: What should Harvard Psychology's AI policy be?				

Course Policies

Office Hours

Office hours are designated times that I set aside to engage with your questions or thoughts that may come up during class or on assignments. You are welcome to come talk to me about the course, the intersection of computation and psychology, or even what it is like to do a Ph.D. They are even a great place to come to do your work – you are very welcome to come and sit and read and just ask me or other classmates present clarifying questions as you go through a paper. Students in my previous classes have liked this option as a set time in their week to work on class assignments. The default format is a drop-in basis (meaning you do not have to stay for the whole hour, you can come and go as you please, and you do not need to tell me ahead of time that you will come) and are for your benefit. Because of this open-door nature, please shoot me an email ahead of time if you would like to discuss something without others present, and we can schedule a specific slot to make sure you are not interrupted. Please stop in to see me! :)

Readings

All course readings will be made available on Canvas. So that you can fully participate in class discussions, please complete the readings before the beginning of each class they are listed next to. In addition to the readings listed on the syllabus, you will be expected to read the articles shared by your classmates on weeks which have an associated discussion. These readings will be shared by the Monday of each week.

Electronics in Class

In this class, we will often do demonstrations with LLMs, so I encourage you to bring your laptop/tablet to class so you can engage with those activities. Besides these class demonstrations... Research shows that using electronic devices in class can reduce focus and lead to more superficial learning, not just for the person using the device, but for others nearby. For that reason, laptops and tablets are allowed for note-taking or class-activities only. It's surprisingly obvious when someone is multitasking during class, and doing so will negatively impact your participation grade. I know I've been guilty of thinking I'll be able to have my laptop out and not be distracted, just to find myself doing other work/email/sudoku five minutes later, so I encourage you to think carefully about how you want to take notes for this course and check in on your choice throughout the semester. If you want a paper notebook for this course, I am happy to provide one!

Collaboration

One of the great things about being in a university setting is that you get to learn through discussion and collaboration. Sharing ideas can be a valuable part of the learning process. I encourage you to talk with your peers when you're reading papers for class or thinking about what you want to write about. You will get explicit practice working with a classmate on the debate assignment. See office hours section if you want a place to work! However, all writing you submit for this course (besides the shared debate writeup) must be your own.

Accommodations & Support

If you have academic accommodations through the Disability Access Office (DAO), please share your letter with me by the end of the second week of the term so we can make sure you get support. If you do not have formal accommodations yet but have concerns about completing assignments or staying on track in this class, I encourage you to reach out to me so we can discuss campus resource options. I'm always happy to talk during regular office hours or by scheduling a one on one, and our conversations will remain private. In some cases, I may need to consult with DAO to make sure we're applying accommodations appropriately or connecting you with the right resources. My goal is always to support your learning and help you do your best. If something changes during the semester, please come talk to me as soon as possible.

Class Participation & Attendance

My goal is to create an inclusive environment where everyone feels valued, regardless of their background in psychology, programming, or personal characteristics. For students, this includes actively listening, ensuring widespread participation, and treating your peers as learners (just like you!) with the same right to contribute and grow. If you have concerns or suggestions about inclusivity in this course, I hope that you will feel comfortable bringing your ideas to me so that we can work together towards improvement. If speaking in class is a skill you're still developing, or if you feel anxious about speaking up, please come talk with me early in the semester. We can work together to create a plan that supports your learning and participation. Just like thinking critically or writing clearly, finding and using your voice is an important skill that you'll carry with you beyond this course and I want to create a safe space for you to practice it. So you can participate, regular attendance is essential for your success in this course. Therefore, you must attend at least 75% of classes to pass (and with only 13 weeks, that means 10 classes). You have one no-questions-asked excused absence to use during the semester. Absences due to illness or family emergencies are considered excused. If you plan to use one, please email me at least 30 minutes before class (9:15AM). Any additional absences will result in a deduction of 1% off your final grade. If you must miss the day you are scheduled as the discussion leader, you will be assigned an additional discussion to run. If anything unexpected comes up that might affect your ability to participate or succeed in this course, please reach out. Please don't worry that you're "bothering" me or that I don't want to know. I'm here to support your learning, and that includes helping you navigate challenges that come up along the way.

AI Use for Assignments

Here is where your syllabi will likely vary most this year; the following policy is related to my class and it is your responsibility to conform to the expectations for each course you are in. I know this can be confusing, so if you need any clarification for this course I encourage you to come to office hours to talk about it. This is a course about AI, so I want you to be able to take advantage of our subject matter and test out incorporating AI into your learning. However, I also want you to actually learn from this class, and not just turn in what the latest LLM can come up with. So, for this course, the following policies apply: You are encouraged to use generative AI to help you parse the readings assigned in this course: they are difficult, and assigned with this suggestion in mind. You do not need to report how you used generative AI for completing your readings. We will discuss strategies for doing this effectively during the course. Your responses to online discussion questions, summary and reflection as discussion leader, reflection on the conversation project, and final paper must be entirely your own writing. However, you are permitted to use generative AI in the brainstorming stage as long as you cite when and how you used it. This entails copying and pasting your prompts and responses, as well as the tool and version you used. I encourage you to only use AI for writing when you really need it! Doing the proper documentation for short debate reflections will be tedious and harder than just thinking on your own, and robs you of the opportunity to practice your persuasion skills. Any text or ideas which are not properly cited as coming from generative AI that you are submitting as your own work will be treated as plagiarism and a violation of the College Honor Code.

Academic Honesty

Unless I specifically say otherwise, all assignments must be your own independent writing. When you're drawing on someone else's ideas, whether it's a classmate, a reading, or an AI generated idea, be sure to cite your sources clearly using APA style. Plagiarism and other forms of academic dishonesty violate the Harvard College Honor Code and will be taken seriously. Improper citation, even if unintentional, can still be considered plagiarism. Submitting someone else's work (or work that isn't properly credited) not only risks academic consequences (including referral to the Honor Council), but it also cheats you out of the education you're investing your time and energy in. I have zero tolerance for plagiarism, and I want you to be thoughtful and careful about how you use sources. If you're ever unsure about what counts as appropriate collaboration or how to cite something correctly, just ask. I'm happy to help. Harvard's honor code: https://honor.fas.harvard.edu/honor-code. FAS Handbook for Students: https://handbook.college.harvard.edu/. Harvard Guide to Using Sources: https://usingsources.fas.harvard.edu/