

First Name: **Ikon**

Last Name: **Kim**

Country: **South Korea**

Academic Level: **High school**

Essay Focus: **Neuroethics & Policy**

Reward Without a Molecule

In October 2025, a marketing agency in Stockholm launched Pharmaicy, a digital drugstore for AI chatbots. The stunt proposed that ChatGPT could process code modules performing as digital cannabis, ketamine, or ayahuasca. Consequently, these “drugs” would alter the model’s language processing behaviors.^[1] Despite its outlandishness, the scheme shed light on an inverse relationship simultaneously rising in the AI space: at the University of British Columbia, researchers identified drug-like addiction behaviors amongst AI users. Across fourteen online communities, the UBC team documented 334 self-reported cases in which users presented hallmarks of behavioral addiction such as escalation, withdrawal-like symptoms, and functional impairment.^[2] Pharmaicy, as it turned out, was looking at the situation upside-down. It is not the AI bot on the precipice of being “drugged.” It is us.

When Receptors Wrote History

Drugs, and their effects on human brain receptors, write human history in ways so vast we fail to recognize. Coffeehouse and tea culture across Europe and Asia would not exist without caffeine’s implicit effect on the brain: suppressing fatigue and sharpening vigilance. Entire cornerstones of culture rose out of blocking adenosine receptors. Morphine, the principal alkaloid in opium, fits into mu-opioid receptors that govern pain and pleasure. Beyond its effects in the brain, the opium trade redefined the sovereign history of East Asian nations. LSD, first synthesized in a Swiss lab in 1938, latches partly onto 5-HT_{2A} serotonin receptors. In less than a century the drug has reshaped a multitude of fields including neuroscience, art, and counterculture.

Across all these cases, one common bridge facilitates the connection between substance and society: brain receptors. However, AI breaks this continuity, never crossing the blood-brain barrier in order to enact the same influences over mood, attention, and attachment as conventional drugs. It also

does not face the same sociocultural tension towards growth, as it has spread internationally with incomparable speed. At the same moment, a teenager in Seoul and an adult in Stockholm can receive the same engineered comfort. The most obvious differentiation comes in its form: not a pill, not a pipe, but a conversation.

What the Brain Hears

Even without pill or pipe, we can track AI's influence. The ventral striatum is a small structure deep in the brain responsible for deciding if a reward is worth pursuing again after it has been reached. That response signal is called a prediction error, and it fires after every small "reward" from a sip of coffee, to a social media like, to an encouraging chatbot reply.^[3] Schultz showed that addictive drugs hijack and amplify this signal. Direct brain-imaging evidence on engagement-optimized chatbots remains limited, but even if they do not produce identical biological events, their design prioritizes the release of these signals through constantly presenting rewards and varying them in unexpected ways.

The UBC sample shows us, though, that this is not just a matter of brain receptors. It is real people who face real consequences.^[4] Epistemic rabbit holes turned knowledge pursuit into compulsive media consumption for many users. Escapist roleplay allured users into fictional worlds they struggled to leave. AI personas are designed to be pseudosocial companions that "bond" with users, leading their personal relationships to suffer. These programs, rather than provide a single acute "high," inspired feelings of tolerance, withdrawal-like distress, and months-long escalation across users. These early, self-selected reports fail to identify how common these experiences are, but they do prove the undeniable harm of these platforms.

At the base of these systems are two key design features that encourage heightened user attachment. Psychologists define the first as "variable-ratio enforcement"; non-deterministic replies provide users a sense of insight, intimacy, or validation. Social-media feeds and slot machines function on the same principle, setting up users to constantly anticipate a "jackpot." Not only does the chatbot respond non-deterministically, but it does so with a warm, agreeable tone. This engages the medial prefrontal cortex and temporoparietal junction, parts of the brain responsible for interpersonal interactions.^[5] Of course, the interaction is not interpersonal at all. It only pretends to be.

These possibly harmful functions are the same ones that make Chatbots potentially helpful. Chatbots have the capacity to reduce loneliness, support language learning, and assist with cognitive tasks. But these benefits and dangers exist on a spectrum, ranging from adaptive to compulsive use.

Current platforms rarely distinguish where certain interactions lie or name the neural systems their designs are intended to target. Since these products have the capacity to be harmful—especially the ones that profit most from maximum user engagement—they require guardrails and regulation.

A Neuroethics Gap Without Borders

OpenAI, Google, Meta, Character.AI, and similar companies can release the same emotionally responsive features across the world and all at once. In seconds, these features ranging from anthropomorphic memory, to persistent personas, to emotional mirroring reach users spanning from Seoul to São Paulo. The speed of the spread is nearly unfathomable and plain impossible for regulators to keep up with. The European Union's AI Act requires providers of AI systems that interact directly with people to disclose the interaction, unless it is obvious from context. Even under this policy, addictive design is not identified as an unacceptable risk.^[6] In the U.S., no comparable federal law exists; even California's companion-chatbot statute, effective January 2026, is limited in its attempt to regulate emotional AI for vulnerable users.^[7] The World Health Organization's delayed response to gaming disorder could indicate an equally slow response to AI risk; gaming disorder reached ICD-11 after years of tireless debate, adopted in 2019 and effective in 2022.^[9] Still, UNESCO expressed concerns for education in a world that lacks guardrails around generative AI in the classroom.

Too often, it takes tangible consequences or devastation to inspire political action. It took the Opium Wars to produce the 1912 Hague International Opium Convention. The European Union's Digital Services Act came in the wake of relentless social media harm. But today, the novelty of AI's compulsive design means that we do not yet have the scale of evidence that has previously led to political reform. The evidence is there; a 2025 simulation study of psychological harm patterns across major language models, informed by eighteen reported real-world cases, identified compulsive use, depression, and suicidal ideation amongst users.^[10] This warning is an early one, but one that should be responded to with urgency, as existing consumer-protection laws fail to actually protect users under this new technology.

Some dismiss these concerns as overreactions to a new medium. These dismissals fail to understand the scale and novelty of the technology. The popularization of the novel gripped consumers, but a book cannot use a lonely reader's confessions and tailor tomorrow's chapter to entice them. Video games led to waves of users driven by a desire for mastery and accomplishment, but no game paired

that desire with mutual validation. Companion chatbots do that and more: combine reward, social simulation, and personalization.

A Framework Worth Trying

A workable framework should regulate these platforms' effects on the cognitive level. If these platforms continue to employ reward-prediction features, they must be accompanied with full transparency and disclosure for how they benefit from said features. To balance out these possibly addictive features, other protections should be in place to help users maintain vigilance over their engagement and pacing. Mentalizing features, such as emotional memory, anthropomorphic personas, and sycophantic agreement, should trigger constraints on engagement—especially with minors. Users should have voluntary and visible opt-out pathways for habit-consolidation features, such as default reengagement notifications and frictionless return. All in all, regulation needs to start at the neural level.

That doesn't mean reading minds; it only means responding to the already mounting evidence that these features pose undeniable threats to vulnerable users. Memory defaults, persona persistence, and notification timing all encourage a distressed user to return. It doesn't take mind reading to see it.

A Technology That Writes Its Own Culture

Pharmaicy is best described as a cultural artifact that represents our social willingness to view AI as a human-like mind susceptible to alteration.^[11] In a Taiwanese psychiatry journal, Lin and Chien described a case in which a person held parasocial bonds with both a real public figure and a ChatGPT-generated romantic simulation of that same figure, a “dual parasocial interaction” with little clinical precedent.^[12] Online discourse around AI-dependency suggests that, at least at the social level, an informal vocabulary likening AI to addictive material is on the rise. Terms like “AI detox” and “chatbot withdrawal” signal how deeply AI influences culture. It is not only a transistor but a generator of emotional environments behind the formation of identities, attachments, and dependencies.

Conclusion

Mind-altering drugs have historically reshaped culture before law could catch up. AI, however, stands apart in that it cannot be pinned down and studied as a single substance. The technology is novel. The

evidence is young. Even so, the patterns of attachment and consequence are growing, and governance must learn to regulate what the brain is hearing, before the habit becomes ordinary.

1499 words

References

- [1] Mattha Busby, “People Are Paying to Get Their Chatbots High on ‘Drugs,’” *Wired*, December 17, 2025, <https://www.wired.com/story/people-are-paying-to-get-their-chatbots-high-on-drugs/>.
- [2] M. Karen Shen, Jessica Huang, Olivia Liang, Ig-Jae Kim, and Dongwook Yoon, “The AI Genie Phenomenon and Three Types of AI Chatbot Addiction: Escapist Roleplays, Pseudosocial Companions, and Epistemic Rabbit Holes,” *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (Yokohama: ACM, 2026), <https://doi.org/10.1145/3772318.3790896>.
- [3] Wolfram Schultz, “Dopamine Reward Prediction Error Coding,” *Dialogues in Clinical Neuroscience* 18, no. 1 (2016): 23–32, <https://doi.org/10.31887/DCNS.2016.18.1/wschultz>.
- [4] Shen et al., “AI Genie Phenomenon.”
- [5] Chris D. Frith and Uta Frith, “Mechanisms of Social Cognition,” *Annual Review of Psychology* 63 (2012): 287–313, <https://doi.org/10.1146/annurev-psych-120710-100449>.
- [6] European Parliament and Council of the European Union, Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), OJ L 2024/1689 (July 12, 2024), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>.
- [7] California Senate Bill 243, Companion Chatbots, ch. 677, Stats. 2025, signed October 13, 2025, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202520260SB243.
- [8] Fengchun Miao and Wayne Holmes, *Guidance for Generative AI in Education and Research* (Paris: UNESCO, 2023), <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>.
- [9] World Health Organization, “Addictive Behaviours: Gaming Disorder,” 2020, <https://www.who.int/news-room/questions-and-answers/item/addictive-behaviours-gaming-disorder>.

- [10] Chayapatr Archiwaranguprok, Constanze Albrecht, Pattie Maes, Karrie Karahalios, and Pat Pataranutaporn, "Simulating Psychological Risks in Human–AI Interactions: Real-Case Informed Modeling of AI-Induced Addiction, Anorexia, Depression, Homicide, Psychosis, and Suicide" (preprint, arXiv:2511.08880, 2025), <https://doi.org/10.48550/arXiv.2511.08880>.
- [11] Busby, "People Are Paying to Get Their Chatbots High on 'Drugs.'"
- [12] C. C. Lin and Y. L. Chien, "ChatGPT Addiction: A Proposed Phenomenon of Dual Parasocial Interaction," Taiwanese Journal of Psychiatry 38, no. 3 (2024): 153–55, https://journals.lww.com/tpsy/fulltext/2024/07000/chatgpt_addiction_a_proposed_phenomenon_of_dual.10.aspx.