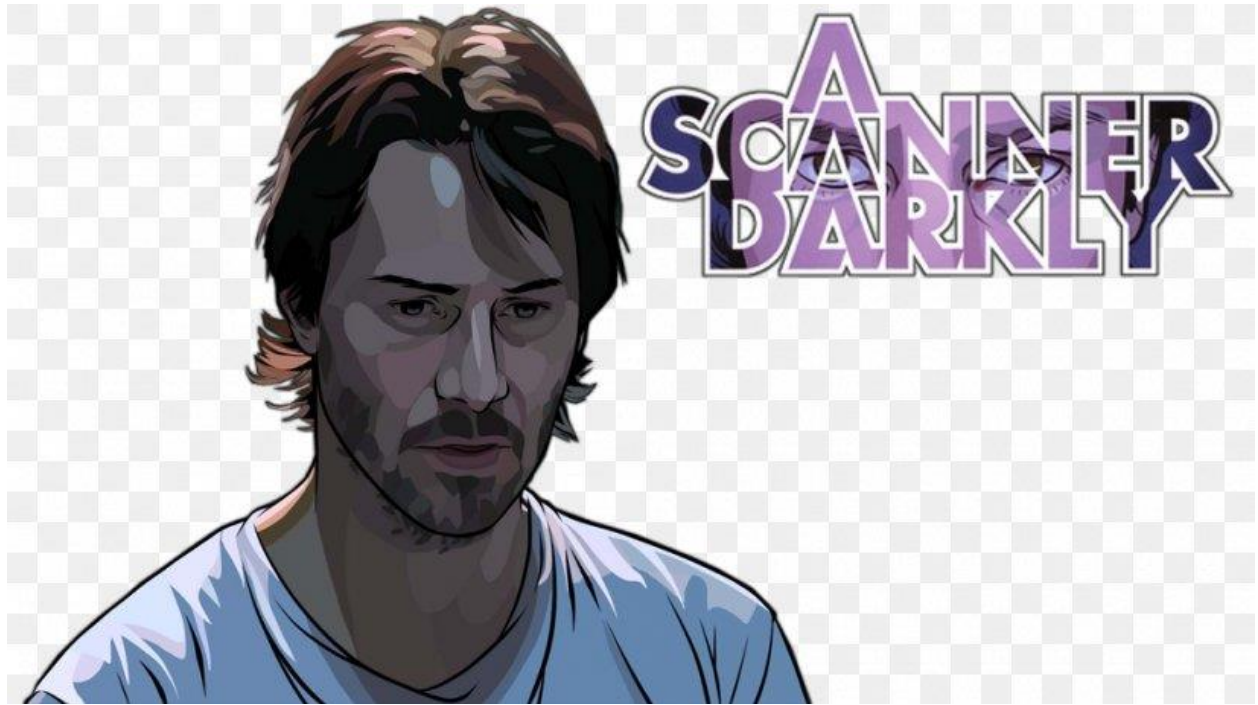


Addicted to Chatbots: ChatGPT as Substance D

Robert Hanna



(Favpng, 2017)

In Philip K. Dick's brilliant 1977 dystopian science fiction novel, *A Scanner Darkly*, and in the equally brilliant same-named 2006 film directed by Richard Linklater, we're presented with a near-future out-of-control epidemic in which 20% of the population has become addicted to an extremely powerful and ultimately deadly drug that causes severe hallucinations, schizophrenia, and long-term cognitive, affective, and practical impairment, *Substance D* (Wikipedia, 2023a, 2023b). And as the novel and the film (presented in glorious rotoscope animation) unfold, we also discover that the obscenely rich and successful company providing behavioral and medical treatment for Substance D addiction, *New-Path*, is actually the very same company that secretly manufactures and markets the drug. In this essay, I'm going to argue that there's a significant two-part analogy (i) between the highly addictive Large Language Model (LLM) or chatbot, *ChatGPT*, on the one hand, and Substance D on the other, and also (ii) between the obscenely rich and successful technocratic capitalist corporation that's the creator, producer, and marketer of ChatGPT, *OpenAI*, on the one hand, and *New-Path* on the other.

Ever since ChatGPT was rolled out in November 2022, it has been informally widely noted that it's highly addictive in just the way that various kinds of mood-altering or hallucinogenic drugs are highly addictive. And indeed, even prior to ChatGPT's roll-out, at least one social scientific study had already provided good empirical evidence that chatbots generally are highly addictive in just this way (Xie and Pentina, 2022).

In a series of recent essays (Hanna, 2023a, 2023b, 2023c, 2023d, 2023e, 2023f), I've argued for two closely related claims. **First**, I've argued that the very idea of "artificial intelligence" is not only an *oxymoron*, in the sense that it's simply false and self-contradictory that any digital computing system or digital technology ("artificial") can equal or exceed the achievements of our rational but also "human, all-too-human" essentially embodied innate mental capacities, faculties, or powers ("intelligence"), but also a *pernicious* myth from which we urgently need to liberate ourselves: what I call *the myth of artificial intelligence*. The myth of artificial intelligence is pernicious precisely because our widespread contemporary dogmatic or at least uncritical acceptance of it leads us to depreciate, neglect, misuse, and even impair our own innate mental capacities, faculties, and powers, via our excessive use of, reliance on, and indeed addiction to, digital computing systems and digital technology. And **second**, I've argued that the primary problem posed by the recent invasion of Large Language Language Models (LLMs) or chatbots like ChatGPT isn't in fact cheating or plagiarism at colleges and universities, but instead the fact that a great many and indeed increasingly many, perhaps even a majority, of all students at contemporary social institutions of higher education, and indeed also at contemporary social institutions of primary and/or middle, and secondary education, are now simply refusing—and will increasingly refuse in the foreseeable future—to think and write for themselves, with grave and indeed tragic consequences, namely, depreciating, misusing, neglecting, and even impairing their innate mental capacities, faculties, and powers, especially those required for autonomous critical reasoning and authentic human creativity: hence I've called this *invasion of the mind snatchers*. And the same threat of mind snatching addiction generalizes over everyone who is using digital technology in our contemporary world.

How does this mind snatching addiction unfold? Leaving aside the obvious "human, all-too-human" dimension of sheer laziness, people say to themselves, roughly,

"Oh!, if only I didn't have to spend so much of my valuable time and energy on the difficult and exhausting acts of thinking and writing, and could simply turn all that painful and tedious stuff over to ChatGPT, then I'd be free to be *so* creative, happy, and productive!"

But in fact, authentic human creativity in thinking and writing arises only through undertaking the essentially embodied, living organismic, uncomputable, conscious, self-conscious, freely-willed, effortful, and enactive processes of *actually thinking and actually writing for yourself* (Hanna, 2023e). Therefore, turning those processes over to ChatGPT is just like taking mood-altering or hallucinogenic drugs in order to produce the cognitive, affective, and practical *illusion* that one is happy, instead of undertaking the essentially embodied, living organismic, uncomputable, conscious, self-conscious, freely-willed, effortful, and enactive processes of *actually becoming and being happy for yourself* (Hanna, 2023g). And when the dope runs out, you discover that you're in fact desperately unhappy – and then you crash-&-burn. Or in other words, using ChatGPT produces only the cognitive, affective, and practical *illusion* of actual thinking and actual writing. And then when you aren't logged into ChatGPT, you discover that you're in fact desperately cognitively impaired, and can't think or write your way out of a wet paper bag, and that the more you use ChatGPT, the worse it gets, and the more you're driven to use ChatGPT to get your work done on time – and then you crash-&-burn. So ChatGPT is significantly analogous to Substance D.

Now, what about the significant analogy between OpenAI and New-Path? There's one obvious disanalogy, which is that in *A Scanner Darkly*, New-Path is *secretly* producing and marketing the very drug whose addictive consequences it purports to treat, whereas, by a prima facie contrast, OpenAI is, well, *so* "open" about what they call the "limitations" of ChatGPT. For example, on OpenAI's "Introducing Chat GPT" page we read:

Limitations

- ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers. Fixing this issue is challenging, as: (1) during RL training, there's currently no source of truth; (2) training the model to be more cautious causes it to decline questions that it can answer correctly; and (3) supervised training misleads the model because the ideal answer [depends on what the model knows](#), rather than what the human demonstrator knows.
- ChatGPT is sensitive to tweaks to the input phrasing or attempting the same prompt multiple times. For example, given one phrasing of a question, the model can claim to not know the answer, but given a slight rephrase, can answer correctly.
- The model is often excessively verbose and overuses certain phrases, such as restating that it's a language model trained by OpenAI. These issues arise from biases in the training

data (trainers prefer longer answers that look more comprehensive) and well-known over-optimization issues.

- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.
- While we've made efforts to make the model refuse inappropriate requests, it will sometimes respond to harmful instructions or exhibit biased behavior. We're using the [Moderation API](#) to warn or block certain types of unsafe content, but we expect it to have some false negatives and positives for now. We're eager to collect user feedback to aid our ongoing work to improve this system. (Open AI, 2023a)

Mmm, right. And now let's click over to the page called "Developing Safe & Responsible AI," where we read that "[a]rtificial general intelligence has the potential to benefit nearly every aspect of our lives—so it must be developed and deployed responsibly" (OpenAI, 2023b). OK, thanks for that.

But why aren't they *also* telling us (i) that ChatGPT is now being widely used to cheat and plagiarize at primary and/or middle schools, high schools, colleges, universities, and other post-secondary educational social institutions, not to mention in business, journalism, law, and politics, and also (ii) that like other chatbots, ChatGPT is highly addictive and leads to our depreciating, misusing, neglecting, and even impairing our innate mental capacities, faculties, and powers, especially those required for autonomous critical reasoning and authentic human creativity? Moreover, no doubt, OpenAI will soon produce new digital technology, intended for use at primary and/or middle schools, high schools, colleges, universities, other post-secondary educational institutions, business, journalism, law, and politics, that supposedly reliably detects the use of chatbots like ChatGPT—initially free of charge, then eventually pay-walled and expensive enough to make them even more obscenely rich and successful; after which, users will quickly discover new ways of making their use undetectable; after which, OpenAI will produce a second generation of the supposedly chatbot-use-detecting digital technology; and so-on and so-forth, *ad infinitum*.

So the answer to my critical question—why aren't they *also* telling us all this?—is obvious. It's precisely because they want to play the Janus-faced satanic/angelic roles of being, on the one hand, the ruthless technocratic capitalist enabler of chatbot addiction and the supplier of the most addictive chatbot now available—namely, ChatGPT—and also being, on the other hand, the "open" and "safe & responsible," publicly-concerned corporate good citizen and potential fixer of the very problems that ChatGPT itself is causing. And then they're relying on our widespread belief in the myth of artificial

intelligence to postpone forever our critical recognition of the fact that what OpenAI is *actually* up to is essentially the same as the Dickian dystopian machinations and subterfuges of New-Path in *A Scanner Darkly*.

Let's assume that what I've just argued is sound. Now, what is to be done? In three of the recent essays that I mentioned above, I've proposed two different solutions to the problem posed by the mind snatching invasion of the LLMs or chatbots, and especially ChaGPT: **first**, *the hard solution*: through democratic processes, get the government to ban LLM or chatbot research and technology while it's still in its infancy, just as we should have, through democratic processes, gotten the government to ban nuclear weapons research and technology, while it was still in its infancy (Hanna, 2023b), and **second**, *the easy solution*: a simple but also radical solution according to which social institutions of higher education should not only shift backward to the required use of handwritten, in-class assignments for the purposes of undergraduate and graduate student evaluation and grading, but also shift forward to a professional academic higher education system in which all career advancement and the highest salaries for faculty members are based on teaching and other non-digital achievements, in which research-&-scholarship is done strictly for its own sake, and in which all publishing by means of hard-copy books or journals, or by means of digital technology, is done strictly for its own sake and for the sake of the general advancement of human knowledge (Hanna, 2023d, 2023e). So if you're interested in either or both of those proposed solutions, then you can take a look at those essays. But the primary purpose of the present essay has been simply to double-underline the importance and urgency of the problem we're all facing, by exploring significant parallels between, on the one hand, our collective current situation with respect to the mind snatching invasion of the chatbots, and on the other, the desperate, tragic situation of Keanu Reeves's schizophrenic and cognitively, affectively, and practically impaired character Fred-the-narc/Bob Arctor-the-addict, a situation also shared by all the other lost souls who are hopelessly addicted to Substance D in *A Scanner Darkly*.¹

¹ I'm grateful to Elma Berisha and Scott Heftler for thought-provoking conversations on and around the main topics of this essay.

REFERENCES

(Favpng, 2017). "Keanu Reeves: A Scanner Darkly." *Favpng*. Available online at URL = https://favpng.com/png_view/scanner-keanu-reeves-a-scanner-darkly-film-trailer-png/5RZXxusw.

(Hanna, 2023a). Hanna, R. "How and Why ChatGPT Failed The Turing Test." Unpublished MS. Available online at URL = https://www.academia.edu/94870578/How_and_Why_ChatGPT_Failed_The_Turing_Test_January_2023_version.

(Hanna, 2023b). Hanna, R. "Hinton & Me: Don't Pause Giant AI Experiments, Ban Them." Unpublished MS. Available online at URL = https://www.academia.edu/97882365/Hinton_and_Me_Don_t_Pause_Giant_AI_Experiments_Ban_Them_May_2023_version.

(Hanna, 2023c). Hanna, R. "The Myth of Artificial Intelligence and Why It Persists." Unpublished MS. Available online at URL = https://www.academia.edu/101882789/The_Myth_of_Artificial_Intelligence_and_Why_It_Persists_May_2023_version.

(Hanna, 2023d). Hanna, R. "Invasion of the Mind Snatchers, Or, The Easy Solution to the Problem of Chatbots in Higher Education." Unpublished MS. Available online at URL = https://www.academia.edu/102995569/Invasion_of_the_Mind_Snatchers_Or_The_Easy_Solution_to_the_Problem_of_Chatbots_in_Higher_Education_June_2023_version.

(Hanna, 2023e). Hanna, R. "Creative Rage Against the Computing Machine: Necessary and Sufficient Conditions for Authentic Human Creativity." Unpublished MS. Available online at URL = https://www.academia.edu/101633470/Creative_Rage_Against_the_Computing_Machine_Necessary_and_Sufficient_Conditions_for_Authentic_Human_Creativity_June_2023_version.

(Hanna, 2023f). Hanna, R. "Further Thoughts on The Myth of Artificial Intelligence, the Mind Snatching Invasion of the Chatbots, and How to Save Higher Education." Unpublished MS. Available online at URL = https://www.academia.edu/103442312/Further_Thoughts_on_The_Myth_of_Artificial_Intelligence_the_Mind_Snatching_Invasion_of_the_Chatbots_and_How_to_Save_Higher_Education_June_2023_version >.

(Hanna, 2023g). Hanna, R. "'It's a Human Thing. You Wouldn't Understand.' Computing Machinery and Affective Intelligence." Unpublished MS. Available online at URL = https://www.academia.edu/102664604/Its_a_Human_Thing_You_Wouldnt_Understand_Computing_Machinery_and_Affective_Intelligence_June_2023_version >.

(OpenAI, 2023a). OpenAI Blog. "Introducing ChatGPT." Available online at URL = <https://openai.com/blog/chatgpt/>.

(OpenAI, 2023b). OpenAI Blog. "Developing Safe & Responsible AI." Available online at URL = <https://openai.com/safety>.

(Wikipedia, 2023a). Wikipedia. "A Scanner Darkly." Available online at URL = https://en.wikipedia.org/wiki/A_Scanner_Darkly.

(Wikipedia, 2023b). Wikipedia. "A Scanner Darkly (Film)." Available online at URL = [https://en.wikipedia.org/wiki/A_Scanner_Darkly_\(film\)](https://en.wikipedia.org/wiki/A_Scanner_Darkly_(film)).

(Xie and Pentina, 2022). Xie, T. and Pentina, I. "Attachment Theory as a Framework to Understand Relationships with Social Chatbots: A Case Study of Replika." *Proceedings of the 55th Hawaii International Conference on System Sciences, 2022*. Pp. 2046-2055. Available online at URL = <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/69a4e162-d909-4bf4-a833-bd5b370dbeca/content>.