Seeds of Science

Stories as Technology: Past, Present, and Future

Roger's Bacon[1], Sergey Samsonau[2] , Dario Krpan[3] (example article)

"What is it about a good story that causes it to have life-changing effects on one person but not another? In theory, future technologies might allow us to develop the truly deep and fine-grained understanding of fiction and human psychology required to answer this question with a high-level of precision. This understanding, coupled with AI and other advanced technology (see below), could lead to the development of advanced "story engineering" capabilities (i.e. the ability to create stories that can elicit highly specific psychological effects on an individual), a capability which could prove both immensely useful and dangerous."

"The universe is made of stories, not atoms" - Muriel Rukeyzer

I.

Humans are storytelling animals. According to Yuval Noah Harari, stories are the <u>key to our success</u> because they provide us with our unique ability to flexibly cooperate in large groups. He describes this view in the following excerpt from his best-selling book *Sapiens*.

> "Fiction has enabled us not merely to imagine things, but to do so collectively. We can weave common myths such as the biblical creation story, the Dreamtime myths of Aboriginal Australians, and the nationalist myths of modern states. Such myths give Sapiens the unprecedented ability to cooperate flexibly in large numbers. Ants and bees can also work together in huge numbers, but they do so in a very rigid manner and only with close relatives. Wolves and chimpanzees cooperate far more flexibly than ants, but they can do so only with small numbers of other individuals that they

---

[2] New York University
[3] London School of Economics

know intimately. Sapiens can cooperate in extremely flexible ways with countless numbers of strangers.

Any large-scale human cooperation—whether a modern state, a medieval church, an ancient city, or an archaic tribe—is rooted in common myths that exist only in people's collective imagination. Churches are rooted in common religious myths. Two Catholics who have never met can nevertheless go together on crusade or pool funds to build a hospital because they both believe that God was incarnated in human flesh and allowed Himself to be crucified to redeem our sins. States are rooted in common national myths. Two Serbs who have never met might risk their lives to save one-another because both believe in the existence of the Serbian nation, the Serbian homeland, and the Serbian flag. Yet none of these things exists outside the stories that people invent and tell one another. There are no gods, no nations, no money and no human rights, except in our collective imagination."

This is a view of stories, broadly construed, as a socio-emotional technology that enables the achievement of collective goals. Angus Fletcher (professor of English at Ohio State University and former neuroscience researcher) develops a similar view in a recent book *Wonderworks: The 25 Most Powerful Inventions in the History of Literature* and argues that we should be teaching literature more like a STEM subject. He discusses his views in an interview with Nautilus magazine:

Why do you call literature a technology?

"A technology is any human-made thing that solves a problem. Most of our technology exists to master our world, to domesticate space. That's why we have smartphones and smart homes and satellites. Literature tackles the opposite set of problems: not how to master the nonhuman world but how to master ourselves. It wrestles with the psychological problems inside us. Grief, lack of meaning, loneliness - literature was invented to deal with these

problems. To have happy and democratic societies, effective engineers and scientists, we need people who are joyful, not angry, who have a deep sense of empathy and purpose, who have an ability for logic and problem-solving. You get all these things from literature.

...It's a machine designed to work in concert with another machine, our brain. The purpose of the two machines is to accelerate each other. Literature is a

way of accelerating human imagination. And human imaginations accelerate literature."

Neuroscientist and author Erik Hoel develops a similar perspective—fiction as a kind of imagination technology—in a fascinating paper that outlines a new functional theory of dreaming—The overfitted-brain hypothesis (OBH).

> "Notably, all deep neural networks face the issue of overfitting as they learn, which is when performance on one data set increases but the network's performance fails to generalize (often measured by the divergence of performance on training vs. testing data sets). This ubiquitous problem in DNNs is often solved by modelers via "noise injections" in the form of noisy or corrupted inputs. The goal of this paper is to argue that the brain faces a similar challenge of overfitting, and that nightly dreams evolved to combat the brain's overfitting during its daily learning. That is, dreams are a biological mechanism for increasing generalizability via the creation of corrupted sensory inputs from stochastic activity across the hierarchy of neural structures. Sleep loss, specifically dream loss, leads to an overfitted brain that can still memorize and learn but fails to generalize appropriately."

The OBH is distinguished from other dream function hypotheses in that it takes the phenomenology of dreams—sparse, hallucinatory, fabulist, narrative—as the key functional feature, not the epiphenomenal expression of some background process like memory consolidation.

> "Rather, the point of dreams is the dreams themselves, since they provide departures away from the statistically-biased input of an animal's daily life, which can assist therefore increase performance. It may seem paradoxical, but a dream of flying may actually help you keep your balance running"

Hoel closes the paper by speculating that fictions may serve as artificial dreams and accomplish some of the same overfitting-prevention function.

> "Finally, it is worth taking the idea of dream substitutions seriously enough to consider whether fictions, like novels or films, act as artificial dreams, accomplishing at least some of the same function. Within evolutionary psychology, the attempt to ground aspects of human behavior in evolutionary theory, there has been long-standing confusion with regard to human interest in fictions, since on their surface fictions have no utility. They are, after all, explicitly false information. Therefore it has been thought that

fictions are either demonstrations of cognitive fitness in order to influence mate choice (Hogh-Olesen, 2018), or can simply be reduced to the equivalent of "cheesecake" — gratifying to consume but without benefit. Proponents of this view have even gone so far as to describe the arts as a "pleasure technology" (Pinker, 1997). However, the OBH suggests fictions, and perhaps the arts in general, may actually have an underlying cognitive utility in the form of improving generalization and preventing overfitting, since they act as artificial dreams."

All of these ideas and perspectives point to a simple fact that we all intuitively recognize—stories (and fictions more broadly) can have significant, sometimes transformative, effects on the mind. Unsurprisingly, there has been a great deal of research on the short-term and long-term psychological effects of fiction (in literature, film, etc.; see the works cited below for a small sample of research in this area). Beyond any insights gained from scientific research, it is apparent that stories have a unique power to effect powerful psychological change.

What is it about a good story that causes it to have life-changing effects on one person but not another? In theory, future technologies might allow us to develop the truly deep and fine-grained understanding of fiction and human psychology required to answer this question with a high-level of precision. This understanding, coupled with AI and other advanced technology (see below), could lead to the development of advanced "story engineering" capabilities (i.e. the ability to create stories that can elicit highly specific psychological effects on an individual), a capability which could prove both immensely useful and dangerous. The remainder of this article represents a few (scattered) speculations on what such a future might look like.

II.

We are approaching a future in which AI writes fiction really well, *like really really well* (see GPT-3 creative fiction for some current AI writing). Future AI will not just be limited to literary masterpieces; AI-generated visual media (DALL-E's successors) will enable the artificial creation of graphic novels and (eventually) movies. There is no reason to think that artificial intelligence won't eventually surpass humans in every artistic domain (though of course many people disagree with this statement—see "Why Computers Will Never Write Good Novels" by the aforementioned Dr. Fletcher).

AI-generated stories, advanced neuroscientific imaging/manipulation techniques, and massive amounts of psychosocial data from a variety of sources (social media being one of them) could collide to create something like "high-throughput neuro-fiction

analysis". This could lead to profound insights into how plot elements, characters, contextual factors (the when, where, why, and how of a story), and the psychological profile of the reader/viewer interact to create specific mental and behavioral changes.

I can envision a future in which AI-generated stories are so incredibly compelling that they act as a kind of superstimulus; we will all be like the male Julodimorpha beetle, irresistibly drawn to brown beer bottles because of their extreme size, color, and dimpled bottom—the definition of sex appeal in a female Julodimorpha beetle.

The custom design of stories based on an individual's psychology and beliefs will prove especially powerful. We are already at the nascent stages of this technology, with the Cambridge Analytica scandal during the 2016 US election serving as a proof of concept. Cambridge Analytica claims to have developed detailed psychological profiles of individuals from a variety of data sources (some of them illegal) and then used those profiles to micro-target voters with advertisements designed to be effective for their personality type. While there is good reason to be dubious of the effectiveness of this techniques given our current abilities, there is no reason in theory why this same approach couldn't be hyper-effective in the future. We already have algorithms that recommend books based on past behavior, why wouldn't a super-intelligent AI with highly-detailed personal data be able to do the same with a truly superhuman level of skill? Think Amazon, but instead of just providing recommendations, the algorithm also generates the stories *for you and you only*.

III.

As with most powerful technologies, we can imagine benevolent and malevolent applications for "advanced story engineering" (as we may call it).

Will there be a way to use AI-generated fictions as therapy for mental illness or for targeted support in difficult times? Will future doctors prescribe custom-made stories to help you deal with the grief from a lost loved one? Advanced story-engineering could be immensely useful in education. Instead of teachers choosing novels based on their preferences and whims and just hoping that it resonates with the class, an AI author will be able to conduct a comprehensive psychological analysis of the students (including mental characteristics, but also their current level of knowledge and skills) and write the perfect story—highly original, deeply moving, and educational—for the class to read and discuss together. Factually accurate and intensely compelling stories could be generated on demand to teach concepts in science, history, and other subjects (didactic novels) . We can also imagine an AI tutor that monitors the social-emotional learning

skills of a student and creates custom stories that teach interpersonal skills, empathy, resilience, etc.

There are many dangers here as well, some of which are not unique to AI-empowered story engineering, such as the general problems that arise from ceding control to a black box AI. Some movies are known as "cult classics" because of their small, obsessive fanbases; in the future, there may be AI-generated fictions that truly do inspire an intense cult-like level of devotion. This will be especially dangerous when the lines between reality and fiction become blurred. Consider QAnon, often described as cult by its (many) critics—a series of anonymous posts on an internet message board started a worldwide movement that has had significant political and cultural consequences. Advanced story engineering could be used to generate conspiracy "stories" like QAnon capable of inciting even greater levels of fervor and extremist behavior. Again, we are already in the beginning stages—a team from Middlebury college has used GPT-3 to create a QAnon chat bot. Eventually, we may be able to generate entirely fictional people who are known only through their videos and podcasts—imagine an Alex Jones-like figure spewing highly compelling conspiracies and hate speech. This power could also be used for good—we can also imagine an AI-generated spiritual guru that provides real wisdom and guidance to her followers (of course this could easily go awry as well). What happens when we cease to care whether a person is real or an AI-generated fiction?

Countermeasures of some sort will have to be developed. I can imagine something like "fiction hygiene" becoming important in the same way that personal and public hygiene are. Paradoxically, the best way to prevent people from becoming too obsessed with stories or fictional people might be to use stories that warn against the danger of doing so.

The following sources provide a small sample of the research on the neurological and psychological effects of fiction.

Berns, G. S., Blaine, K., Prietula, M. J., & Pye, B. E. (2013). Short- and Long-Term Effects of a Novel on Connectivity in the Brain. *Brain Connectivity*, 3(6), 590–600.

Carroll, J. (2018). Minds and meaning in fictional narratives: An evolutionary perspective. *Review of General Psychology*, *22*(2), 135-146.

Castano, E., Martingano, A. J., & Perconti, P. (2020). The effect of exposure to fiction on attributional complexity, egocentric bias and accuracy in social perception. *PLOS ONE, 15(5), e0233378.*

Consoli, G. (2018). Preliminary steps towards a cognitive theory of fiction and its effects. *Journal of Cultural Cognitive Science*, *2*(1), 85-100.

Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading has a small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology: General, 147*(11), 1713–1727.

Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Opening the closed mind: The effect of exposure to literature on the need for closure. *Creativity research journal*, 25(2), 149-154.

Jacobs, A. M., & Willems, R. M. (2018). The fictive brain: Neurocognitive correlates of engagement in literature. *Review of General Psychology*, 22(2), 147–160. doi:10.1037/gpr0000106

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, *342*(6156), 377-380.

- Note: the validity of these results has been the source of significant controversy - see the following papers by Kidd & Castano and Panero et al.

Kidd, D. C., & Castano, E. (2017). Panero et al. (2016): Failure to replicate methods caused the failure to replicate results. *Journal of Personality and Social Psychology, 112*(3), e1–e4.

Kidd, D., & Castano, E. (2017). Different stories: How levels of familiarity with literary and genre fiction relate to mentalizing. *Psychology of Aesthetics, Creativity, and the Arts, 11*(4), 474–486.

Kidd, D., & Castano, E. (2019). Reading Literary Fiction and Theory of Mind: Three Preregistered Replications and Extensions of Kidd and Castano (2013). *Social Psychological and Personality Science*, 10(4), 522–531.

Oatley, K. (2016). *Fiction: Simulation of Social Worlds. Trends in Cognitive Sciences, 20(8), 618–628.* doi:10.1016/j.tics.2016.06.002

Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology*, *111*(5), e46

Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2017). No support for the claim that literary fiction uniquely and immediately improves theory of mind: A reply to Kidd and Castano's commentary on Panero et al. (2016). *Journal of Personality and Social Psychology, 112*(3), e5–e8.