Apparently I Want a Coffee

Brain State Prediction, Memory Engineering, and the Future of Human Experience

Tyler Bell · January 18, 2025
catsplayingchess.com/writings/apparently-i-want-a-coffee/

PART 1

Brain State Prediction

A recent paper from researchers at the University of Sydney demonstrates something remarkable: **AI can predict brain states**[1]. Using the transformer architecture—the same technology powering modern AI language models—researchers successfully predicted ~5 seconds of brain activity from ~20 seconds of recorded data.

The researchers achieved this using fMRI data representing activity across 379 brain regions. Impressively, their predictions maintained functional connectivity patterns—the complex web of relationships between different brain areas. In other words, the predicted brain states weren't just random activity; they preserved the fundamental organization of brain function.

From 30 fMRI scans, their model could accurately predict the next 7 consecutive brain states. How accurately? The first predicted time point/state following the inputs had **a correlation of 0.997 with the ground truth fMRI data.** And the first 5 predicted states had a correlation of greater than 0.85!

Extrapolating: **we now have a model that can predict into the future what the brain will do next.**

Parallels to Language Models

The results might seem modest at first glance and this extrapolation might seem extreme. But consider the trajectory of language models. GPT-3 completed the text from some user-provided context and input prompt.

Example 1: Movie to Emoji Generation

Input:

```
Back to Future: 👨👴🚗🕒
Batman: 🤵🦇
Transformers: 🚗🤖
Star Wars:
```

Output:

```
🚀🏰🚀
```

Example 2: First-Person to Third-Person Writing

Input:

```
in: i started the business in 2010 we sell gift boxes with chocolates and other things
out: The business was founded in 2010, and they sell gift boxes with chocolates and other
items.

in: we don't have any investors
out: The company has taken no outside funding.

in: our only major risk is our supplier who we depend on for 90% of our stock
out: The company's main challenge is its dependence on a single supplier for 90 percent of
its stock.

in: We have a lot of loyal customers and are growing rapidly
out: The company has a lot of loyal customers and is growing rapidly.

in: we don't use social media for marketing
out: The company doesn't use social media for marketing.

in: i am currently in the process of opening up a new store
out:
```

Output:

```
The company is currently in the process of opening up a new store.
```

Trajectory of Impending Improvements

The brain study reminds me of GPT-3. In the same way that we provided some context and examples to GPT-3 with our input prompt, the 20 seconds of fMRI data is providing context and examples to the brain state prediction model. Along these lines then, the same way that GPT-3 completes the text from our input prompts, the brain state prediction model can predict the next few brain states from the fMRI input data. It stands to reason that the improvements we have seen in LLMs since GPT-3 will be seen in brain state prediction. Not only is this logical from the standpoint of continuous technology advancement (we have nowhere to progress but forward), but much of the groundwork required for these improvements have already been laid.

Soon we should see analogous improvements in brain state prediction along the lines of:

- **Increased input length:** longer fMRI history provided as contextual input to the model.

- **Increased output length:** more accurately predicted states / more time predicted into the future.
- **Increased model strength and performance** at similar or fewer model parameters.
- **Faster inference time:** more states accurately predicted per fixed amount of time and compute.
- **Test-time compute (TTC):** reasoning-like search behavior that leads to more accurate predictions.
- **Thanks to TTC:** an abundance of better, increasingly generalized, synthetic data that can be used as training data for future models.
- **Thanks to improved training data:** stronger brain prediction models that can be (1) distilled down into (slightly weaker but) smaller and faster versions of themselves or (2) used to generate even more and better synthetic data via TTC.
- **Takeoff:** the cycle and improvements repeat and continue.

Just as early language models learned to predict the next word in a sequence, we can now predict the next state in a sequence of brain activity. Just as language models progressed from simple prediction to understanding context, generating coherent long-form content, and engaging in complex reasoning, brain state prediction could evolve from brief moments to extended sequences of neural activity. I feel convinced about this and there is even more evidence to support it.

Parallels to Video/World Models

I am further confident in the prediction due to improvements in world models. Take for instance **Oasis** [2] [3] that was announced on October 31, 2024.

```
"Oasis, the first playable, realtime, open-world AI model. It's a video game, but entirely
generated by AI. Oasis is the first step in our research towards more complex interactive
worlds.

Oasis takes in user keyboard input and generates real-time gameplay, including physics, game
rules, and graphics. You can move around, jump, pick up items, break blocks, and more. There
is no game engine; just a foundation model."
```

The 500M parameter model can run locally and a larger parameter count model can run on a cloud-based demo. The model supports user interactions like jumping, breaking blocks, and more. The reception online was lackluster, mostly in part due to Oasis's restriction to 'only generating Minecraft' and its short-term memory. The 'context window' is only about 1 second (~20 frames). With this, there is not much persistence of the world: spinning rapidly can cause the world to disappear or seemingly teleport the user to a new location. But at this early stage it

is crucial to remember: **there is no game engine.** Only a foundation model that generates real-time frames as the user interacts with the world.

```
"The model is composed of two parts: a spatial autoencoder, and a latent diffusion backbone. Both are Transformer-based...we chose Transformers to ensure stable, predictable scaling, and fast inference..."
```

Just over a month later, Google dropped **Genie 2**, a large-scale foundation world model [4].

```
Genie 2, a foundation world model capable of generating an endless variety of action-controllable, playable 3D environments for training and evaluating embodied agents. Based on a single prompt image, it can be played by a human or AI agent using keyboard and mouse inputs.
```

Overcoming the 1 second memory limitation of Oasis, this model can generate consistent worlds for 10-20 seconds (and up to a minute). This is a 10x-20x (and up to a 60x) improvement. Also, Genie is not limited to generating Minecraft worlds. The Genie world-model has the ability to generate consistent worlds for a wide variety of environments, agents, and games. The interactions available in each generated world are dependent upon the world itself: first-person perspectives and movement, isometric views and movement, third-person perspectives and movement, diverse controls (car, boat, bird, dragon, hot air balloon)... the user (or agent) can burst balloons, open doors, explode barrels, climb ladders... the Genie world-model can generate NPC behavior, water physics, models of gravity, lighting effects, reflections and initialize the world from real-world or AI-generated seed images.

LLMs have progressed and will continue to progress. World models and world generation have progressed and will continue to progress. **Brain state prediction and brain state generation have progressed and will continue to progress.** Not only will this result in AI systems that can begin assisting us, before we even know what it is that we need. But this will also result in generative systems that can begin generating environments, realities, conditions, and interactions to assist, entertain, and satisfy us.

Take for example the 2023 work from Stability, “Reconstructing the Mind’s Eye: fMRI-to-Image with Contrastive Learning and Diffusion Priors” [5] [6].

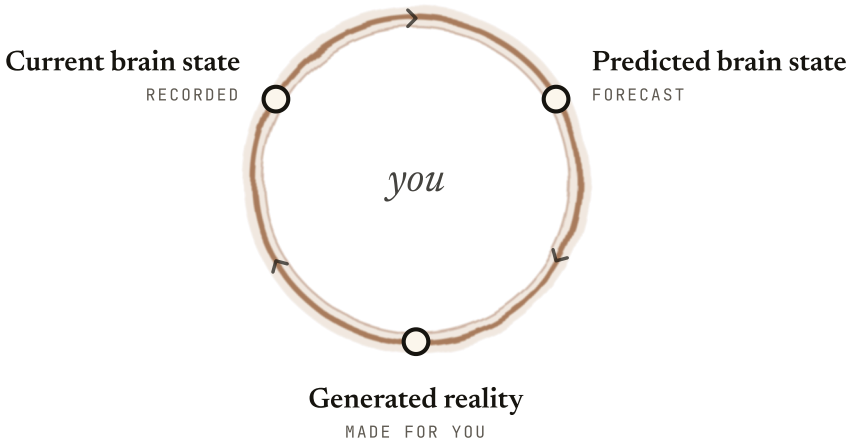
```
MindEye is a state-of-the-art approach that reconstructs and retrieves images from fMRI brain activity...MindEye achieves state-of-the-art performance across both image retrieval and reconstruction. That is, given a sample of fMRI activity from a participant viewing an image, MindEye can either identify which image out of a pool of possible image candidates was the original seen image (retrieval), or it can recreate the image that was seen (reconstruction).
```

Here we have image selection and/or image generation based on brain activity. The same brain activity that can now be predicted into the future. Let's go on a walk:

We know we can generate images from text and other images. By extension, we can now generate beautiful, high quality videos from text and images. Increasingly, we are able to generate playable/explorable/interactive 3D scenes from text and images [4:1] [7] [8] [9]. Along the parallel from Stability, if we can generate images from brain activity, we can (or will be able to) generate videos from brain activity, and then explorable/interactive 3D scenes from brain activity. Combine this with the brain state prediction results, it follows that we will be able to generate images, videos, and novel immersive experiences **from brain activity that has not yet occurred.**

At some point a chicken and egg problem arises: is the next predicted brain state representative of the next actual brain state or is it one that, in a way, has been generated conditioned on the world that is being continuously generated (in response to previous brain states)? In this case, an interesting feedback loop is created. Are we predicting the next brain state or are we essentially generating the next brain state by generating the appropriate stimuli?

The Prediction-Generation Loop



As we push the boundaries of brain-state prediction and generative AI, a fascinating feedback loop emerges. We predict future brain states, which then inform the generation of experiences and environments, which in turn influence our actual brain states, which feed back into subsequent predictions. Over time, prediction and reality begin to intertwine in a self-reinforcing cycle:

- **Brain-State Recording:** We capture neural activity across hundreds of brain regions, forming the raw data for analysis and modeling.

- **AI Prediction:** Advanced models extrapolate future brain states from current activity, inching closer to real-time forecasting of our thoughts, emotions, and intentions.
- **Generative Output:** Those predictions inform what experiences get generated—ranging from virtual realities and personalized content feeds to direct neural feedback—effectively guiding the brain's next states.
- **Real-World Impact:** The generated experiences alter actual neural activity, potentially aligning it more closely with the model's initial prediction.
- **Refinement and Recurrence:** With each iteration, new brain-state recordings help calibrate the model further, tightening the feedback loop.

This isn't just a theoretical leap. A direct parallel to this dynamic exists in today's recommendation algorithms: they predict what content you'll engage with, show it to you, and in doing so, shift your preferences—making future predictions even more accurate. Here, however, the loop is more fundamental: rather than just influencing the content you view, these predictive–and–generative AI systems could shape how your brain itself processes and responds to reality.

The Convergence of Prediction and Generation

This self-reinforcing loop raises a profound question: **what happens when prediction accuracy and generation fidelity both approach near-perfection?** In other words, when we can forecast exactly what will happen and seamlessly generate the exact experiences we intend, does the line between predicting the future and creating it vanish?

- **Perfect Prediction:** As AI's predictive power grows, error margins shrink. Brain-state forecasts might become so accurate that behavior, perception, and physiological responses are effectively "known" before they even occur.
- **Perfect Generation:** On the flip side, improvements in immersive technologies—like high-fidelity VR, direct neural stimulation, and adaptive multi-modal content generation—could allow us to craft experiences that align almost perfectly with any desired outcome.

When both forces converge, the act of predicting the future begins to resemble the act of manufacturing it:

- **Self-Fulfilling Prophecy:** If you can precisely anticipate how someone will feel or act, and if you can shape their environment accordingly, that prediction can become a self-fulfilling prophecy.
- **Probability vs. Certainty:** Probability might lose its traditional meaning if we can nudge events into becoming certainties. Are we truly guessing at the future, or are we proactively

eliminating alternative outcomes?

- **Blurred Reality:** As generation fidelity climbs, the difference between a "natural" brain state and an "artificial" or "guided" one becomes murky. When every experience is curated to meet a predictive model's blueprint, "authentic" reality and "designed" reality can start to feel indistinguishable.
- **Shared vs. Competing Agents:** In a multi-agent world, other people (or other AI systems) are also predicting and shaping outcomes. Whose predictions "win" in the event of a conflict or divergence? This raises questions about collective free will and societal consensus. Are my experiences, interactions, and brain states the ones being optimized, am I part of a collective that is being optimized, or am I simply some non-optimal weight that is part of someone else's optimization?

PART 2

Implications to Free Will in a Predictable World

> *"The future cannot be predicted, but futures can be invented."*
> – Dennis Gabor, Inventing the Future (1963)

As our abilities to predict and generate brain states advance, we face a fundamental question: how far ahead can we accurately forecast neural activity? If these models continue to scale—following the trajectory we've seen with language and world models—could we extend predictions from seconds to minutes, hours, or even days?

Individual Agency

When you get driving directions, do you follow the route that has been planned for you?

When your next brain state can be predicted with 99.7% accuracy and that prediction then shapes the environment you'll experience, is your response truly free? We're moving beyond simple behavioral prediction into a realm where the very neural patterns that underlie our thoughts and decisions can be forecast—and potentially influenced—before we're even aware of them.

Consider a short-term example: If the system predicts you'll crave coffee in 10 seconds—and automatically adjusts your environment (maybe by delivering a push notification or adjusting aromas)—did you choose coffee or did the system simply steer you toward it?

I ask again: when you get driving directions, do you follow the route that has been planned for you? It's likely that you do. But of course you could take another route or drive off the road altogether if your heart desired. But do you? This is a crude example. For better and worse, I do not think it will be this overt. In reality, you will not even be aware that you are being steered

towards coffee. You simply became aware that you wanted a coffee (well after your AI had already become bored of this eventuality) and magically one was ready for you. Success!

Or maybe on that particular day Starbucks paid OpenAI enough in 'advertising' to subtly nudge people like you toward this seemingly innocuous craving...

Either way, such scenarios highlight a subtle but profound shift. The more accurate predictions become, the more effectively they will be able to shape your immediate reality, blurring the line between **anticipating your choices** and **producing them**.

Networked Consciousness

The complexity deepens when multiple individuals enter the picture. Each person exists within their own prediction–generation loop, but these loops aren't isolated. In their (now infamous) 2014 study [10], Facebook showed the ability to transfer and predict emotional states among a population of nearly 700,000 users, "leading people to experience the same emotions without their awareness."

```
In an experiment with people who use Facebook, we test whether emotional contagion occurs
outside of in-person interaction between individuals by reducing the amount of emotional
content in the News Feed. When positive expressions were reduced, people produced fewer
positive posts and more negative posts; when negative expressions were reduced, the opposite
pattern occurred. These results indicate that emotions expressed by others on Facebook
influence our own emotions, constituting experimental evidence for massive-scale contagion
via social networks. This work also suggests that, in contrast to prevailing assumptions, in-
person interaction and nonverbal cues are not strictly necessary for emotional contagion, and
that the observation of others' positive experiences constitutes a positive experience for
people.
```

If this was possible with nearly a million users in 2014 based "only" on social media profiles and interactions, imagine the scale of what will be possible when we can predict, transfer, and generate emotional and brain states. **In such a network, where does true autonomy reside?** Is each person acting freely, or are we participants in a grand feedback loop of predictions, influences, and reinforced behaviors?

Collective Determinism

These intertwined loops raise a crucial question: in a world where everyone's brain states are being predicted and subtly shaped, **who—or what—is steering the ship?**

1. **Personal Models:** predict and influence individual brain states.
2. **Social Models:** anticipate and shape group dynamics—family, friends, entire communities.

3. **Environmental Systems:** smart homes, immersive XR interfaces, etc. generate experiences based on these aggregated predictions.

All of this unfolds in a continuous feedback loop. The traditional notion of free will assumes a degree of autonomy in our decision-making. But as neural states become more predictable, networked, and molded by external cues, the boundary between genuine choice and predictive determinism begins to fade. We may still feel like we're making choices, yet the system itself may be orchestrating events in ways imperceptible to us.

In this context, I ask again: are my experiences, interactions, and brain states the ones being optimized for, are yours the ones being optimized for, or are we all part of a collective that is being optimized? *Or (exhaustingly) are all our actions and experiences now able to be imperceivably steered to optimize returns for the highest bidder(s)?*

Self-Fulfilling Prophecies and the Illusion of Choice

This isn't just philosophical speculation—the Facebook experiment showed how emotional states can propagate through networks "without direct interaction between people" and "in the complete absence of nonverbal cues." Now imagine that same principle guided by direct brain-state prediction and generation:

- If the model forecasts you'll experience sadness, it might (inadvertently or by design) arrange conditions that reinforce that mood. In effect, the sadness becomes both predicted and produced.
- If the system "anticipates" your curiosity about a topic, it can feed you precisely the stimuli to deepen that curiosity—culminating in a reality shaped by forecasts rather than spontaneous impulse.

Or maybe it will not be so linear. Maybe we will be able to set preferences that that can be used to shape our experiences. Take weight loss for example. Did your model predict that you're about to crave that extra slice of pizza? Maybe the generative model can pivot, generating new content to proactively distract you, preemptively curbing that craving before it ever existed. Or did your model predict that you're likely to be sad tonight? Perhaps your AI assistant can come in with positive words of encouragement, recall a recent memory of a time you were happy, or conjure up a scene of utmost cuteness.

But wait, was it my own free will to set these preferences in the first place or was I expertly nudged (read: guided, manipulated) into setting them?

Where does free will fit when experiences are so finely tuned to match the model's expectation? Are you living your life, or living the model's output? Will we live meaninglessly perfect lives or

perfectly meaningless lives?

Toward a New Ethics of Prediction

As these predictive and generative loops become more refined, the question of responsibility and consent looms large:

- **Personal Autonomy:** How can individuals maintain agency in a world of micro-nudges, immersive feedback, and near-perfect forecasts?
- **Collective Well-Being:** When entire communities or social platforms utilize brain-state predictions, do they optimize for engagement, well-being, profit—or something else?
- **Transparency and Governance:** What safeguards or regulations ensure that predictive power isn't abused—intentionally or unintentionally—to manipulate emotions, thoughts, or behaviors?

We're entering a new era, one where the capacity to predict our minds is intertwined with the power to shape them. Whether that convergence leads to a richer, more enlightened human experience or a subtle erosion of freedom depends on choices we make now—about ethics, governance, and the very definition of autonomy in a world where free will itself may become entangled in the loops of prediction and generation.

PART 3

The Memory Engineering Horizon

> *"The past is never dead. It's not even past."*
> – William Faulkner, Requiem for a Nun (1950)

The convergence of brain state prediction and generation opens doors far beyond simple influence or control. One compelling possibility emerges: **the ability to reconstruct, enhance, and even share experiences at the neural level.**

Today, virtual reality headsets are already used in healthcare to help manage pain or anxiety, such as during chemotherapy sessions [11] [12] [13]. These devices distract patients with calming or engaging 3D environments, reducing perceived discomfort. Although they only offer visual and auditory stimuli, it's not a huge leap to imagine next-generation neural interfaces—initially noninvasive, but far more sensitive—tapping directly into the brain's activity. Over time, these interfaces could evolve to become capable of recording, editing, and even playing back our internal experiences. What begins as a modest enhancement to VR might eventually become a profound tool for immersive memory recall, emotional regulation, and shared experiences. *Note: even if direct neural interfaces are not possible for a good while, real-time 3D world models will be able*

to play back experiences in multi-modal immsersive VR and/or AR. We already see aspects of this concept with the Apple Vision Pro that can both record and play back Spatial (2.5D) Videos.

Consider this (somewhat morbid) science fiction scenario: You're on your deathbed in hospice care. To help remove you from the confines of your physical body, you're offered a neural interface—a "memory helmet" of sorts. This headset accesses your memories and allows you to relive your life over the course of a few hours. Advanced generativeAI fills in the gaps in your fuzzy memories. Global parameters like a "luck slider" can be turned up or down to ensure that your experience is overall enjoyable. After your session, the device is removed and wiped clean for the next patient.



The neural 'VR' headset featured in Netflix's 3 Body Problem.

Such a device would allow you to:

- Relive memories with enhanced clarity: an AI fills in the gaps in your recollection with parameters allowing you to control the level of precision vs creativity, luck, etc. in your experience (analogous to 'temperature' in LLMs, today).
- Adjust the emotional tone of remembered experiences.
- Share processed, edited, or enhanced memories with others.
- Preserve personality patterns for future interaction / digital cloning.
- Process and reshape sensitive or traumatic memories in a controlled environment.

These aren't just science fiction concepts anymore. The technical foundation is already being laid. **What will you choose to remember?**

Beyond Memory: Reshaping Reality

The implications extend far beyond memory manipulation. As these technologies mature, we're approaching an era where the line between memory, present experience, and future prediction begins to blur. The same systems that can reconstruct and enhance memories could be used to shape real-time experiences and even pre-emptively generate desired future states.

This convergence of capabilities creates something more profound than simple memory enhancement. Brain state prediction determines your likely future emotional states while world models generate environments and experiences tailored to those states. Neural feedback systems adjust your current state in real-time, memory enhancement systems store and optimize these experiences, and social networks share and synchronize these experiences across individuals.

What emerges is a kind of "reality engineering" where experiences aren't just remembered or lived but actively crafted. Rather than waiting for experiences to happen, systems could create optimal scenarios based on predicted desires and needs. Environmental and neural states could be modified in real-time to maintain desired experiences. Groups of people might share coordinated experiences, creating collective realities. The boundary between past memories, present experiences, and anticipated futures becomes increasingly fluid.

The question transforms from **"What will you choose to remember?"** to **"What reality do you want to experience?"**

Fabricating Entirely New Experiences



AI Image of an Italian Vacation. Fresh gelato and warm sunshine.

While the idea of re-living or refining real memories is already radical, these same tools could go further: fabricating experiences that never happened. Imagine a neural interface creating a vivid "memory" of an Italian vacation you never actually took—complete with the taste of fresh gelato and the warmth of sunshine on your skin. If the system's generative power were convincing enough, you might come to cherish these artificially constructed moments every bit as much as real ones. Initially, these experiences might take place in VR-like headsets, but as neural interfaces evolve, they might be able to be experienced completely synthetically (similar to how kung fu was learned in The Matrix)

The implications are staggering. Would such memories be considered authentic? Or maybe they'll be embraced and shared like a favorite book or film? If every sensation—from the crunch of gravel underfoot to the distant murmur of Italian conversation—feels real at the neural level, **how thick is the line between lived experience and imagination?**

The Path Forward

We stand at a crucial juncture. The technical foundations for neural engineering are being laid faster than our ethical frameworks can adapt. How do we preserve personal autonomy in a predicted world? What constitutes authentic experience when every memory and moment can be enhanced? Who owns neural data, and how do we protect it? As these technologies become available, who gets access to them? And how do we handle collective consent when experiences can be shared across neural networks?

The societal impact promises to be profound. We're not just redefining human experience and consciousness; we're transforming how people interact and relate to one another. What will interactions look like between people who are living life augmented by this AI and those who are not?

When memories can be reimaged and relived, the past isn't just alive—it's malleable. When we can predict, generate realities, and modify neural states at will, the present isn't just happening—it's being actively shaped. The future isn't just coming—it's being engineered.

Will we use these tools to heal trauma and expand individual and collective consciousness? Or will we trap ourselves within optimized artificial realities? What kind of reality do you want to remember? What kind of present do you want to engineer? **What kind of tomorrow do you want to relive?**

RELATED SHORT STORY
"The Veil"

The story that began here is now a piece of its own: **The Veil**.

NOTES & SOURCES

1. Sun, Y., Cabezas, M., Lee, J., Wang, C., Zhang, W., Calamante, F., & Lv, J. (2024). “Predicting Human Brain States with Transformer.” arXiv:2412.19814

2. “Oasis: A Universe in a Transformer” https://oasis-model.github.io/

3. “Oasis: A Universe in a Transformer” https://www.decart.ai/articles/oasis-interactive-ai-video-game-model

4. “Genie 2: A Large-Scale Foundation World Model” https://deepmind.google/discover/blog/genie-2-a-large-scale-foundation-world-model/

5. “Reconstructing the Mind’s Eye: fMRI-to-Image with Contrastive Learning and Diffusion Priors” https://stability.ai/research/minds-eye

6. Scotti, P. S., Banerjee, A., Goode, J., Shabalin, S., Nguyen, A., Cohen, E., Dempster, A. J., Verlinde, N., Yundler, E., Weisberg, D., Norman, K. A., & Abraham, T. M. (2023). “Reconstructing the Mind’s Eye: fMRI-to-Image with Contrastive Learning and Diffusion Priors.” arXiv:2305.18274

7. “CAT4D: Create Anything in 4D with Multi-View Video Diffusion Models.” https://cat-4d.github.io/

8. Wu, R., Gao, R., Poole, B., Trevithick, A., Zheng, C., Barron, J. T., & Holynski, A. (2024). “CAT4D: Create Anything in 4D with Multi-View Video Diffusion Models.” arXiv:2411.18613

9. “GenEx: Generative an Explorable World.” https://www.genex.world/

10. Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). “Experimental evidence of massive-scale emotional contagion through social networks.” *Proceedings of the National Academy of Sciences*, 111(24), 8788-8790. https://doi.org/10.1073/pnas.1320040111

11. Schneider, S. M., & Workman, M. L. (1999). “Effects of Virtual Reality on Symptom Distress in Children Receiving Chemotherapy.” *CyberPsychology & Behavior*, 2(2), 125-134. https://doi.org/10.1089/cpb.1999.2.125

12. Schneider, S. M., Prince-Paul, M., Allen, M. J., Silverman, P., & Talaba, D. (2004). “Virtual Reality as a Distraction Intervention for Women Receiving Chemotherapy.” *Oncology Nursing Forum*, 31(1), 81-88. PDF

13. Schneider, S. M., & Hood, L. E. (2007). “Virtual Reality: A Distraction Intervention for Chemotherapy.” *Oncology Nursing Forum*, 34(1), 39-46. https://doi.org/10.1188/07.ONF.39-46