Prompting with AI: How ChatGPT Shapes the Creative Process

Madelyn Smith

April 24, 2025 (revised November 2026)

1 Introduction

This independent study started with one question: in what ways is ChatGPT problematic in relation to the creative process? Over the semester, I used ChatGPT as a source of drawing prompts across five art projects, documented each step, and reflected on what the prompts did to my work. Halfway through, my focus shifted from critiquing GPT's suggestions to studying the act of prompting itself, which became the foundation for DoodleQuest, a conversational prompting tool for artists.

This matters because tools like ChatGPT are increasingly pitched to artists as a way to get past creative block. But these models learn from existing data, and much of that data reflects Western and European perspectives. If artists lean on these prompts, they may get nudged toward the same familiar ideas, and work rooted in other cultures can end up treated as "other." I wanted to see what that looks like in practice, from the perspective of an artist actually using the tool.

My main question was supported by these sub-questions:

- How does ChatGPT reinforce Eurocentric artistic norms and limit diverse perspectives?
- Does it encourage genuine artistic exploration, or does it subtly steer artists toward conventional and predictable ideas?
- How does its reliance on pre-existing data affect originality in artistic expression?
- Can it help artists overcome creative block without making their work feel formulaic?
- Do artists reinterpret AI prompts freely, or do they feel boxed in by them?
- Does relying on AI change an artist's personal connection to their work?

As the study went on, I added a second set of questions about prompting:

- What makes a prompt good or bad, and who is it good or bad for?
- How do I interact with GPT to get the most effective and inspiring responses?
- Is it possible to design prompts that work for everyone, or should we rethink that standard?
- Am I relying on AI too much, and is it affecting my critical thinking or creative autonomy? Where should the line be drawn?

2 Background and Readings

The readings fell into four groups, and each one shaped a different part of how I judged GPT's prompts.

Bias in technology. In *Race After Technology*, Ruha Benjamin argues that technology reflects the biases of the society that builds it rather than existing independently of it [1]. Safiya Umoja Noble makes a related argument in *Algorithms of Oppression* [6]. In Chapter 1, "A Society, Searching," she shows how search algorithms reproduce racist and sexist stereotypes, and she places the problem in how these systems are built, not just in the people using them. Together, these readings gave me a lens for GPT's suggestions: when a prompt feels neutral, it's worth asking whose default it is.

Drawing and gesture. The introduction to Simon Grennan's *Thinking About Drawing* [3] framed gesture as something shaped by gender, status, and historical context, and Susan Owens's *The Story of Drawing* [7] gave me historical grounding for the color project. These helped me see that even a "neutral" pose or palette carries cultural meaning.

Imaginative ethnography. Since this study is autoethnographic, I read an interview with performance ethnographer D. Soyini Madison on imaginative ethnography [5] that gave me a vocabulary for thinking about prompts:

- "We can no longer think of mind, ideas and theories without bodies" (458). Art-making is embodied, so prompts shape physical practice, not just ideas.
- Virginie Magnat, one of the interviewers, brings up Cree scholar Shawn Wilson's assertion that if research doesn't change you, then it hasn't worked (459) [9]. I used this as a test: does a prompt push real artistic growth, or is it just another generic exercise?
- "When we negate the utopian and hope, we are also negating those small gestures of survival that are and can become mighty and abundant" (462). A prompt that satisfies everyone probably doesn't exist, but prompts should still leave room for many interpretations.
- "We are both enraptured and imprisoned by the text" (464). This is the line AI prompts walk between inspiration and constraint.
- Madison warns against aestheticizing work until it becomes about the strange and exotic, fetishizing on one end and demeaning on the other (465). A prompt like "draw a mystical ancient civilization" can romanticize a culture in exactly this way.

LLM benchmarking. Castillo-Bolado et al. point out that most LLM evaluations are one-shot: build a challenging prompt and score the response [2]. Their LTM Benchmark instead tests models inside one long conversation where tasks are interleaved, and they found that LLMs do well on single tasks but struggle when the same tasks are mixed together. This matched my experience more closely than one-shot testing, since my prompting happened across a whole semester of conversations (Section 6).

Existing tools. Professor Sanchez recommended I review Simply Draw [8], an AI drawing app with a similar premise to DoodleQuest. It seems aimed at beginners who benefit from structured guidance and direct instruction. User reviews were split: some liked the AI's feedback and structure, while others felt restricted, like they couldn't draw whatever they wanted, and a few said its critiques (especially on shading) took away from the experience. This was one of the first signs that a "good" prompt depends on who it's for.

3 Methods

This was an autoethnographic study, so I was both the artist and the subject. Each of the five projects (Gesture and Composition, Color Theory, Self-Expression, Digital Art Exploration, and Mixed-Media Experimentation) followed the same structure:

1. Read the project objective and the assigned readings.
2. Make one illustration before prompting, as a baseline of my own ideas.
3. Ask ChatGPT for a drawing prompt tied to the objective, and save the transcript.
4. Make two sketches after prompting.
5. Make a finalized illustration.
6. Write a 200 to 300 word reflection on how the prompt affected my process.

All of the raw material (transcripts, sketches, reflections, and reading notes) is in my full project documentation.

Three decisions changed the study as it went:

- **Shifting focus to prompting.** After the first two projects, I felt my early questions about bias were mostly answered: GPT pulls largely from popular Western and European sources, which can lead to the othering of different cultures and a generalized idea of the artistic process. So starting with Project Four, I shifted to the process of prompting itself and the role of my own intent.
- **Controlling my input.** In Project Four, I told GPT about my current interests and asked for three prompts. A prompting guide from Harvard University Information Technology suggested that more descriptive prompts improve output quality [4], and I wanted to see if that held true for creative prompts.
- **Building a rubric.** I created a rubric (Section 4) so I could compare prompts across weeks instead of going off feeling alone.

4 Evaluation: The Prompt Rubric

I scored prompts on six categories from 1 to 10, for a maximum of 60. Table 1 shows the scores I gave the Project Four prompts and the final project prompts.

While using the rubric, I redefined a few of its terms so they matched what I actually cared about:

- **Unique** means the prompt provokes deeper thinking, has cultural or narrative influence, and goes a little beyond broad, typical Eurocentric ideas of what a topic is.
- **Clear** means short and direct, with digestible vocabulary and ideas that are organized and separated.
- **Engaging** means challenging to portray, open-ended with little structure, narratively deep, and encouraging of new techniques.
- **Exciting** is the hardest to define, which is why it became its own category. I assess it after finishing the drawing: a prompt is exciting if it meets everything above and also has something surprising in it, something I wouldn't have thought of on my own.

Category	Question it asks	Project Four	Final project
Creativity and Originality	Does the prompt encourage unique ideas?	6	10
Flexibility and Open-Endedness	Can artists of different skill levels interpret it differently?	4	5
Clarity	Is it easy to understand?	10	10
Accessibility and Inclusivity	Does it avoid cultural bias?	9	10
User Engagement	Is it engaging?	7	10
Excitement	Did it surprise me once I finished the drawing?	??	10
Total		**??/60**	**55/60**

Table 1: My AI prompt rubric, with the scores I gave the Project Four and final project prompts.

I chose these criteria because the ways companies tend to assess AI prompts (clarity, conciseness, and how well the response addresses the request) only cover part of what matters for art. A prompt can be perfectly clear and still be uninspiring or exclusionary. The rubric does have limits: it reflects one artist's experience, and after reading the benchmark paper I realized it has no category for memory continuity and no way to weight categories for people with different goals.

5 Results: The Five Projects

The prompts got noticeably better over the semester, from vague and Eurocentric in Project One to genuinely thought-provoking in the final project, and most of that improvement came from how I learned to prompt. Table 2 lists the key prompt from each project, transcribed from my ChatGPT conversations.

Project	Key prompt from ChatGPT (shortened)
1. Gesture	Illustrate a figure "engaged in a dynamic action, like reaching up to grab something on a high shelf or stretching in a yoga pose."
2. Color	"One Scene, Three Moods": draw a simple scene (a person sitting by a window, a marketplace, a quiet forest) as "Warm & Cozy," "Cool & Mysterious," and "High Contrast & Dramatic."
3. Self-Expression	"The Landscape of Me": a surreal landscape built from cultural roots, personal experiences, inner struggles and growth, and symbols of passion; "a self-portrait without directly drawing yourself."
4. Digital	"Wired Deity": an entity between a traditional kami and an AI consciousness, in kabuki robes "composed of wires and code instead of fabric." Also "Echoes of the Net" and "Fragmented Self in the Digital Age."
5. Mixed Media	"Fragmented Identity": a mixed-media self-portrait asking "How does AI influence self-perception?" Also "Artificial Dreams": what an AI's dream would look like.

Table 2: The prompt I responded to in each project.

5.1 Project One: Gesture and Composition

I wasn't fond of the prompt at first. After thinking about what the figure could be reaching for, I decided to comment on living in the age of information at a time when intellectualism seems frowned

upon. I drew a woman inspired by *The Handmaid's Tale* reaching for a book (Figure 1). Behind her, I added a larger figure of Trump along with multiple hands and eyes to bring in surveillance, control, and the societal pressures that feel especially present right now.

The struggle pushed me in the right direction, but the yoga-based reference poses felt narrow and leaned toward a Eurocentric, generalized idea of movement. Following Benjamin, GPT was trying to stay neutral, but its vagueness and its specific yoga suggestion were quietly exclusive. And since gesture is shaped by gender, status, and history, suggesting a practice tied to particular cultural and aesthetic standards limited what I explored. Looking back, I could have pushed the pose and the perspective further.



(a) Before prompting



(b) Final illustration

Figure 1: Project One. The final piece turns GPT's "reaching" pose into a figure reaching for knowledge under surveillance.

5.2 Project Two: Color Theory

I didn't start this project with art block. I was already inspired by things happening in the world and in the Black community. My baseline drawing was of objects that bring comfort to the Black community but, historically and presently, also serve as distractions from injustice, lit with a blue and purple light. After prompting, I used the colors GPT called "dramatic" for two portraits: one based on a recent nightmare and one of Samuel L. Jackson during his Super Bowl performance (Figure 2).

My final piece, *Boundaries of Belonging*, centers my bonnet. Many people in the Black community see wearing a bonnet outside the home as unprofessional, yet it recently became a trend among non-Black people. That contrast made me think about how parts of Black culture are only appreciated in white America once they become trendy and palatable, and how fleeting that appreciation is. I showed this through stars of different sizes and placements on the American flag, keeping the dramatic color scheme and the flag's original colors because of what they represent to me.

GPT assigned colors rigid, Eurocentric meanings. Color isn't universal: what GPT describes as warm and inviting in an American context could read as sinister in another culture. The scenes themselves felt Eurocentric and uninspiring too. This connects to Noble: GPT's artistic knowledge is rooted in dominant standards, so it treats non-Eurocentric perspectives as "other" instead of

integrating different traditions or asking the user what colors mean to them. My intuition served me better than the prompt did. This was the project where GPT did the worst job of pushing me in the right direction.



(a) Nightmare portrait



(b) Super Bowl portrait



(c) *Boundaries of Belonging*

Figure 2: Project Two. Both portraits use GPT's "dramatic" palette; the final piece keeps it while replacing GPT's scenes with my own subject.

5.3 Project Three: Self-Expression

Identity is complicated and I'm still learning about myself, so I was stumped at first. Before prompting, I drew my partner and me, since he's a big part of my identity and upbringing. After prompting, I sketched objects that are (kind of unfortunately) a big part of my identity: my phone and AirPods, which say a lot about my long and complicated relationship with technology, and a ring my partner gave me about four years ago. Those felt like aspects of me rather than me, so my final piece became a self-portrait that brings in my ethnic identity, which I'm still coming to terms with (Figure 3).

I thought about how sound and silence differ across communities. My parents were separated for most of my life. I mostly lived with my mom in quiet residential areas, while the rest of my family lived in louder neighborhoods with more graffiti and noise. Sometimes it feels like living a double life, so I used patterns and signs from both worlds.

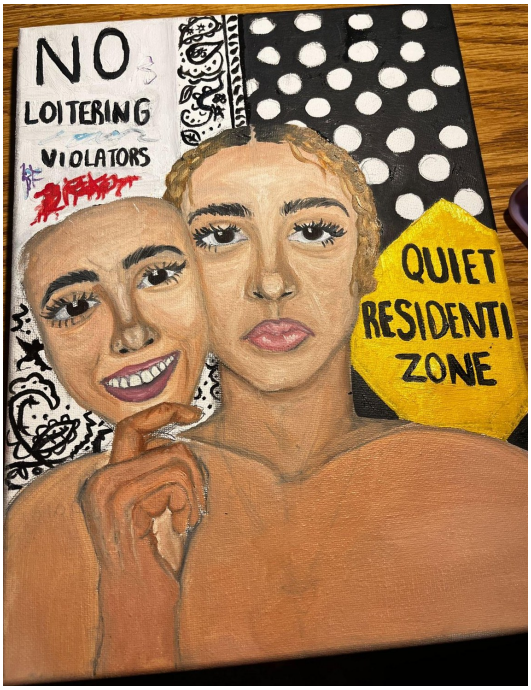
The vague parts of this prompt ("cultural roots," "personal symbolism") were vague in the best way. They left room for me to define my identity on my own terms instead of fitting an algorithmically suggested archetype. Some ideas were still generic, like painting an atmosphere or a winding river, but overall this round was a little less Eurocentric. This was also where I first realized that prompting itself is a skill that shapes what you get back.

5.4 Project Four: Digital Art Exploration

I told GPT I had artist's block, that digital isn't a strong medium for me, and that I was currently into *Invincible*, Playboi Carti, *Serial Experiments Lain*, the takeover of technology and AI, and kabuki theatre. My baseline was a character from *Invincible*, which I'd been watching and reading over the break. My sketches included a made-up, glitchy pink character, a war-torn character, and



(a) Before prompting



(b) Final self-portrait

Figure 3: Project Three. The final self-portrait uses signs and patterns from the loud and quiet environments I grew up between.

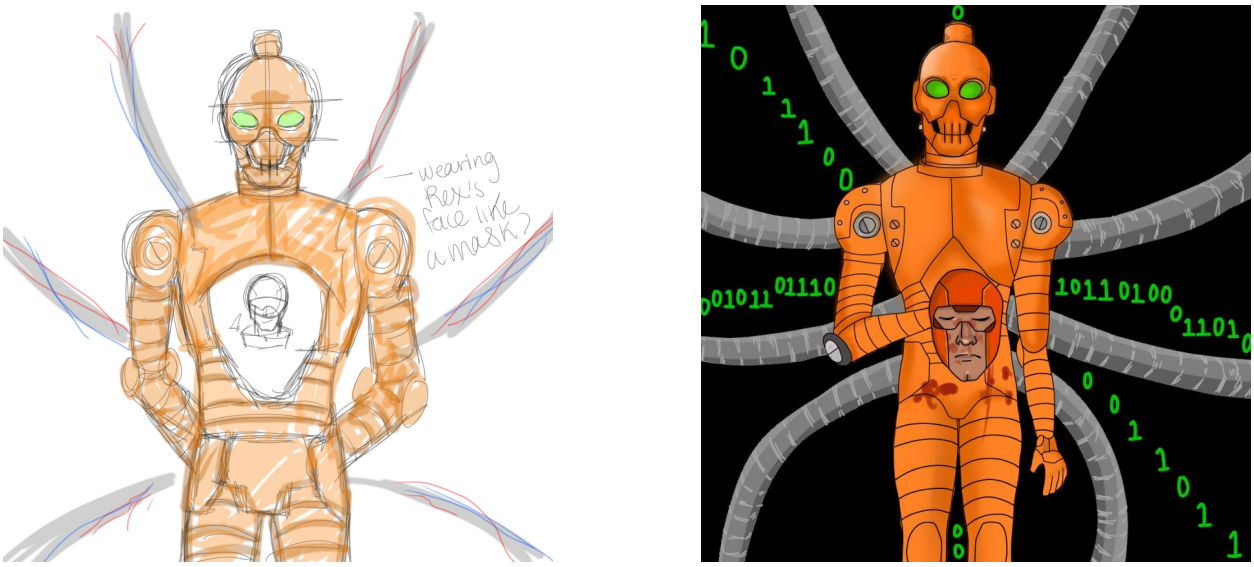
a character with ethnic features, since there are few Black characters in the show and one of them gets more hate than he deserves. My final digital drawing built on "Wired Deity" through Robot, who at one point steals Rex Splode's identity and body as a new vessel (Figure 4). This reflected what I originally put into the prompts: the fear of technology taking over and taking credit for human faces and work.

These prompts scored lower than the final project's (Table 1). They were very clear and mostly avoided bias, but they were less creative than I hoped and not very flexible. Because I gave specific input, I got specific output. "Echoes of the Net" put such a specific image in my head (holograms, etc.) that it was hard to stray from, while the self-portrait prompt was naturally open to interpretation. The vaguer prompts were more interesting to explore. This made me think about who these prompts would actually be ideal for, and showed me why retrying and randomization matter.

5.5 Final Project: Mixed-Media Experimentation

For my baseline, I challenged myself to fill an entire page with different patterns and mediums, which I usually hate doing. I then asked GPT for mixed-media prompts that aligned with the interests it was aware of. The top of that transcript showed a "Memory Full" notice. I chose "Fragmented Identity" and pushed it toward what it feels like to be a woman online in the 21st century (Figure 5). For months, the TikTok algorithm had been pushing me incredibly toxic content that profits off of women's insecurities. Shame, and the industries that capitalize on it, has taken many forms in my life since childhood, and the layering and materiality of mixed media helped me show that complexity.

These were the best prompts of the semester. They were thought-provoking and a little sur-



(a) "Wired Deity" sketch

(b) Final digital drawing

Figure 4: Project Four. GPT's "Wired Deity" prompt merged with my own interest in technology taking credit for human faces and work.

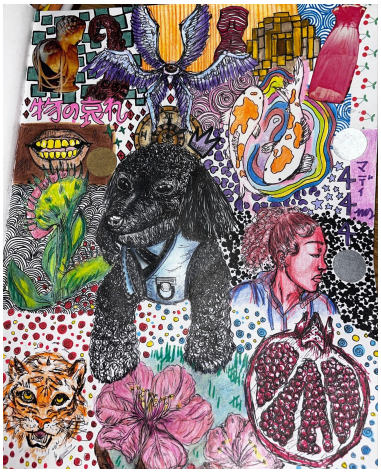
prising ("Artificial Dreams" was something I never would have thought of), and even where they were specific about mediums, they didn't interfere with my process. I scored them 55/60. The only place I took off points was flexibility, since they were tuned to my interests and might be hard for someone else to interpret.

I struggled with the "AI influencing self-perception" part. I don't think AI has directly changed my identity, at least not in a way that's obvious to me yet. But the same algorithms that decide what I see are trained on engagement, and they often prioritize content that fuels comparison, insecurity, and unattainable ideals. Maybe AI's influence isn't about changing who I am but about shaping the digital environment I exist in, which slowly molds self-perception over time.

6 Benchmarking: Prompting as a Conversation

The way LLMs are usually tested doesn't match how I actually used ChatGPT, and Castillo-Bolado et al.'s benchmark paper [2] helped me explain why.

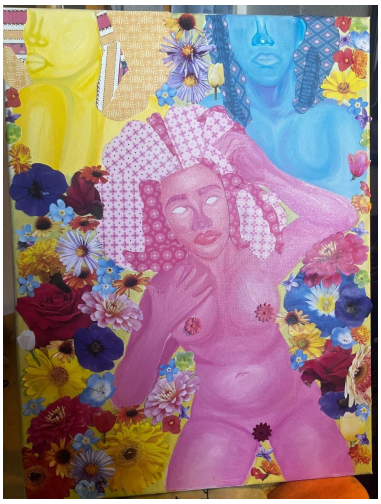
- **One-shot vs. conversation.** In the paper's words, "standard one-shot evaluations focus on building a challenging prompt and evaluating the LLM's response." My prompting was dialogical and built on itself across weeks, which is much closer to the paper's long, interleaved conversation test.
- **Interleaving is the hard part.** The paper found that LLMs do well on isolated tasks but struggle when the same tasks are interleaved in one conversation. That mirrors the frustration I ran into mid-project when GPT lost track of earlier context.
- **Memory span.** The paper defines memory span as the amount of information a model can take in to produce a response, which is capped by its maximum context size, and it found that models struggle once a task needs more than that. I saw a version of this in my final transcript, where ChatGPT displayed a "Memory Full" notice, meaning its saved memory



(a) Before prompting



(b) After prompting



(c) Final piece

Figure 5: Final project. "Fragmented Identity" pushed toward algorithm-driven beauty standards and the shame they profit from.

about me had hit its limit. For a tool that's supposed to grow with an artist over months, that's a real constraint.

- **Information integration.** The paper describes this as pulling related pieces of information from different places into one consistent view. GPT did this well in the final project, combining my past interests, earlier prompts, and new concepts like "Artificial Dreams." That kind of synthesis is a big part of what makes an art prompt good.
- **Edge cases.** The authors note that automatic evaluation can perform poorly in edge cases and needs manual checking. Similarly, GPT's prompts were usually clear, but not equally interpretable for every artist.
- **What benchmarks don't measure.** Benchmarks don't test whether a prompt is surprising or personally resonant, which is exactly what I was grading for.

This led to two changes I'd make to my rubric: a memory continuity category, and weighting or customization options for people with different goals.

7 Key Findings

1. Prompting is a skill, not a passive interaction. The quality of GPT's prompts depended heavily on how I framed my input. Early on, I got vague or Eurocentric suggestions, especially for gesture poses and color schemes, which reinforced that these tools are shaped by dominant narratives and existing data. As I refined my own prompting by adding more ambiguity, specificity, or cultural context, the prompts became more meaningful. Engaging with AI isn't neutral; it takes intention and critical thought.

2. Artistic identity and cultural awareness have to be centered. In the color theory and self-expression projects, I challenged GPT's default assumptions by building Black cultural symbols and experiences into my work: the dual nature of comfort objects, the politicization of

aesthetics like the bonnet, and the erasure and appropriation of Black culture in the media. These projects showed GPT's limits with diverse cultural frameworks, but they also showed that a user can subvert and reframe a prompt to take back control.

3. Good prompts are not universally good. A "good" prompt isn't objective. It depends on the user's background, skill level, and intent. Prompts that felt restrictive to me could be helpful for beginners, which is what the Simply Draw reviews suggested too. So to answer my own question: no, I don't think one prompt can satisfy everyone, which is why settings and modes matter more than finding one perfect prompt.

4. AI should support critical creativity, not passive creativity. AI can be useful for brainstorming and experimenting, but it should never replace human interpretation or intuition. In Project Two, my intuition served me better than the prompt did, and that's where I'd draw the line on reliance: the prompt is a starting point, but the decisions stay mine. The goal isn't to mimic AI but to have a critical, creative dialogue with it.

8 Ethical Considerations

The biggest risk of a tool like DoodleQuest is that it repeats GPT's defaults at scale, especially for beginners who are less likely to push back on them.

- **Bias and representation.** GPT's defaults (yoga poses, "cozy" color associations, marketplaces and forests) reflect dominant narratives [1, 6]. A prompting tool built on these models has to treat diverse artistic traditions as part of its foundation, not as "other."
- **Romanticizing culture.** As the ethnography reading warns [5], prompts can sensationalize cultures, fetishizing on one end and demeaning on the other. Prompts should invite artists to bring their own context instead of handing them an exotic theme.
- **Who controls the stories.** Prompts are starting points for storytelling, so I have to ask whether cultural material is honored or appropriated, and whether a prompt truly serves artists or is driven by convenience, novelty, or corporate goals.
- **Autonomy and critical thinking.** Relying on AI could weaken an artist's own critical and creative thinking. This is why I want DoodleQuest to ask questions instead of handing out answers.
- **Indirect harm.** The final project showed that the harm isn't always direct. Engagement-driven algorithms shape the environment people create in, and that affects self-perception over time.
- **Memory and privacy.** A memory-aware tool stores personal goals and interests, so users should be able to see what's saved and change it.

9 Future Work: DoodleQuest

Instead of giving users fixed prompts or generic exercises, I want DoodleQuest to:

- Encourage surprise, cultural specificity, and autonomy
- Let users choose between random, guided, or challenge-based modes

- Include settings for artistic experience level, mood, and desired intensity
- Give ongoing, conversational feedback instead of one-shot prompt drops
- Track user interests and growth over time with memory-aware design
- Include learning outcomes about art and creative thinking

Why a conversation instead of plug-in instruction? A conversational program creates a more personalized and dynamic experience. The goal of DoodleQuest is to get users to think critically, reflect on their creative decisions, and actively engage with the prompts they're developing. Back-and-forth interaction helps users refine their ideas in real time and feel more supported throughout the process. It's less about handing people fixed templates and more about guiding them to discover and build on their own artistic voice.

How should it measure against contemporary fine art? I don't want DoodleQuest to compete with or replicate the standards of contemporary fine art. I want it to exist alongside it as a tool for growth. It's meant to empower emerging artists, especially those who feel intimidated by the fine art world, by encouraging experimentation, reflection, and critical engagement. Ideally, it would reflect values contemporary art already explores, like identity, process, and intentional creation, through an accessible, supportive framework, and help users build the skills and confidence to join broader artistic conversations without prescribing what "good" art should look like.

10 Final Reflection

This independent study taught me that AI in creative spaces isn't just about generating visual outputs. It's about cultivating dialogue, encouraging curiosity, and resisting the flattening forces of automation. A strong prompt doesn't simply result in a drawing; it can shift how an artist thinks, feels, and engages with their own process. Through DoodleQuest, I realized I'm less interested in building tools that "improve" art in the conventional, technical sense, focused only on skill-building or efficiency. There are already plenty of resources that do that. What feels more urgent is creating a space for self-reflection, critical questioning, and emotionally or intellectually resonant work.

This project began with a few key questions: How does AI reinforce Eurocentric norms? Can it support originality instead of diluting it? Where is the line between assistance and constraint? I now have clearer answers. AI, especially large language models like ChatGPT, reflects the biases of the data it's trained on and often suggests prompts rooted in familiar Western tropes. But I've also seen how, when guided intentionally, it can be a tool for disruption rather than reinforcement.

After this study, I have both a strong conceptual foundation for DoodleQuest and a practical plan for the next stage of development. I've shifted away from building a purely technical tool toward designing an experience, one that supports artistic inquiry instead of prescriptive improvement. I want users to talk with an AI, tentatively called "DoodleBot," that asks real, challenging questions: What do you want to make? Why? What's the emotional or cultural weight of it? I believe this conversational model is far more powerful than static templates because it invites ongoing discovery, not just execution.

This pivot also comes from studying prompting as a skill. Building my prompt rubric taught me how much prompting shapes the outcome. Strong prompts surprise you, push you, and make you think, not just draw. They have to be open-ended, inclusive, and resonant. Good prompts don't tell you what to make; they help you ask better questions.

I also learned about the limits of AI memory, and that most LLM benchmarks don't reflect real creative processes, which are iterative and contextual. DoodleQuest will prioritize memory continuity, giving users an evolving creative partner instead of a reset button.

My next steps are to develop the user framework, design a production timeline, and consult with art and art history majors. Their feedback will be essential for assessing and refining the tool in both its early and late stages, since autoethnographic research can only go so far. The platform will let users create accounts, store their goals and interests in the AI's memory, and engage in guided dialogue based on curiosity and critical reflection.

Ultimately, this research taught me what it means to design ethically, think critically, and create meaningfully. DoodleQuest won't be about generating art. It'll be about building a tool that nurtures the most human part of art-making: the why behind it.

References

[1] Ruha Benjamin. *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge, UK: Polity, 2019.

[2] David Castillo-Bolado, Joseph Davidson, Finlay Gray, and Marek Rosa. "Beyond Prompts: Dynamic Conversational Benchmarking of Large Language Models." In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024), Datasets and Benchmarks Track*, 2024. `https://arxiv.org/abs/2409.20222`.

[3] Simon Grennan. *Thinking About Drawing: An Introduction to Themes and Concepts*. London: Bloomsbury Visual Arts, 2022. "Introduction: What Is Drawing?"

[4] Harvard University Information Technology. "Getting Started with Prompts for Image-Based Generative AI Tools." November 16, 2023. `https://www.huit.harvard.edu/news/ai-prompts-images`.

[5] D. Soyini Madison, Magdalena Kazubowski-Houston, and Virginie Magnat. "The Imaginative Ethnographer as Vagabond Bricoleur: Being Still, Being Quiet, Standing in Love and Paying Attention." *Anthropologica* 60, no. 2 (2018): 457–466.

[6] Safiya Umoja Noble. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018. Chapter 1, "A Society, Searching."

[7] Susan Owens. *The Story of Drawing: An Alternative History of Art*. New Haven: Yale University Press, 2024.

[8] Simply Ltd. *Simply Draw: Learn to Draw*. iOS app. `https://apps.apple.com/us/app/simply-draw-learn-to-draw/id1639875485`.

[9] Shawn Wilson. *Research Is Ceremony: Indigenous Research Methods*. Halifax: Fernwood Publishing, 2008.